SAS® 9.2 Language Reference
Dictionary
Fourth Edition

SAS® Documentation

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2011. *SAS® 9.2 Language Reference: Dictionary, Fourth Edition*. Cary, NC: SAS Institute Inc.

SAS® 9.2 Language Reference: Dictionary, Fourth Edition

ISBN 978-1-60764-882-6

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st electronic book, January 2011
2nd electronic book, August 2011

1st printing, March 2011

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at **support.sas.com/publishing** or call 1-800-727-3228.

Contents

What's New

Overview

The SAS 9.2 Base new features, language elements, and enhancements to the language elements continue to expand the capabilities of SAS:

- SAS now supports the next generation Internet Protocol, IPv6, as well as IPv4.
- The DATA step component Java object enables instantiation of Java classes and accessing fields and methodsChapter 10, "Java Object Language Elements," on page 2145 on resultant objects.
- The SAS logging facility is a new logging subsystem that can be used to collect, categorize, and filter log events and write them to various output devices. The logging facility can be used to log SAS server events or events that are initiated from SAS programs. This feature is new for SAS 9.2 Phase 2.
- In addition to SAS Monospace and SAS Monospace Bold TrueType fonts, new TrueType fonts are available when you install SAS.
- Universal Printing now supports Scalable Vector Graphics (SVG), Portable Network Graphics (PNG), and PDFA-1b print output formats.
- You can access remote files by using the Secure File Transfer Protocol (SFTP) access method.
- SAS now reads and writes ISO 8601 dates, time, and intervals.
- In support of batch programming, if a program terminates without completion, the new checkpoint mode enables programs to be resubmitted in restart mode, resuming with the DATA or PROC step that was executing when the program terminated.
- In the "Functions and CALL Routines "section there are several new and enhanced functions as well as functions that were previously in other products and that are now part of Base SAS. The functions that moved from the Risk Dimensions product calculate the call and put prices from European options on futures, based on various models. The functions that moved from SAS/ETS return information about various date and time intervals. The functions from SAS High-Performance Forecasting return specific dates.
- The documentation for string functions and CALL routines now has a restriction that identifies whether theses functions and CALL routines support Single Byte

Character Sets (SBCS), Double Byte Character Sets (DBCS), or Multi-Byte Character Sets (MBCS). This distinction is important because improper use of these functions and CALL routines can result in unexpected behavior in programs that are written in a non-English language. The description for the restrictions is located in the Function Compatibility with DBCS, MBCS, and SBCS Character Sets section of the documentation.

- In a DATA step, you can track the execution of code within a DO group. The DATA statement has an optional argument for you to write a note to the SAS log when the DO statement begins and ends.
- New SAS system options enable you to set a default record length, specify options for accessing PDF files, specify values for Scalable Vector Graphics, support the checkpoint mode and the restart mode, and support fonts.
- Some of the new features for the DATA step object attributes, operators, and methods remove all items from the hash object without deleting the instance of the hash object, consolidate the FIND and ADD methods into a single method call, return the number of items in the hash object, and specifies a starting key item for iteration.
- In previous versions of *SAS Language Reference: Dictionary*, references to language elements in other publications were included in their respective dictionary for each language element type. For example, you could find a reference for the $BIDI format in the format dictionary entries. You can now find references to language elements that are documented in other publications within each section for the language element types. Online, this section appears just before the dictionary entries for each language element type. In the PDF or print copy, this section appears as the last topic for each language element type.

A section that describes how SAS syntax is written has been added. This section contains examples of how to interpret the syntax.

SAS System Features

Checkpoint Mode and Restart Mode

If a batch program terminates before it completes and it was started in checkpoint mode, the program can be resubmitted in restart mode, resuming with the DATA or PROC step that was executing when the program terminated. DATA and PROC steps that have already completed do not need to be rerun. See "Checkpoint Mode and Restart Mode" in *SAS Language Reference: Concepts*.

Support for ISO 8601 Basic and Extended Time Notations

In SAS 9.1.3, the formats and informats that support the ISO 8601 basic and extended time notations were documented in the *SAS 9.1.3 XML LIBNAME: User's Guide*. These formats and informats have been renamed and are now documented in *SAS Language Reference: Dictionary*.

The new names clearly distinguish the basic and extended formats and informats. You can see the renamed formats and informats in their respective sections in the topics that follow. In addition, a new CALL routine, IS8601_CONVERT, converts ISO 8601 intervals to datetime and duration values, and datetime and duration values to an ISO 8601 interval.

Support for IPv6

SAS 9.2 introduces support for the "next generation" of Internet Protocol, IPv6, which is the successor to the current Internet Protocol, IPv4. Rather than replacing IPv4 with IPv6, SAS 9.2 supports both protocols. A primary reason for the new protocol is that the limited supply of 32-bit IPv4 address spaces is being depleted. IPv6 uses a 128-bit address scheme, which provides more IP addresses than IPv4 did.

For more information, see Internet Protocol Version 6 (IPV6) in *SAS Language Reference: Concepts*.

Universal Printing and New TrueType Fonts

In SAS 9.2, all Universal Printers and many SAS/GRAPH devices use the FreeType engine to render TrueType fonts for output in all of the operating environments that SAS software supports. In addition, by default, many SAS/GRAPH device drivers and all Universal Printers generate output using ODS styles, and these ODS styles use TrueType fonts.

In addition to SAS Monospace and SAS Monospace Bold, 40 additional fonts (TrueType) are available when you install SAS:

- Three Latin fonts compatible with Microsoft
- Ten graphic symbol fonts
- Eight multilingual Unicode fonts
- Nineteen monolingual Asian fonts

In the third maintenance release for SAS 9.2, the MingLiU_HKSCS TruType font is new. In addition, the HeiT, MingLiU, MingLiU_HKSCS, and PMingLiu fonts support the HKSCS2004 (Hong Kong Supplemental Character Set) characters.

New Universal printers include the following:

PDFA	produces an archivable PDF compliant with PDF/A-1b .
PNG	produces Portable Network Graphics, which is a raster image format that is designed to replace the older simple GIF and the more complex TIFF format.
PNGt	produces transparent Portable Network Graphics.
SVG	produces Scalable Vector Graphics, which is a language for describing two-dimensional graphics and graphical applications in XML.
SVGt	produces transparent Scalable Vector Graphics.
SVGnotip	produces Scalable Vector Graphics without tooltips.
SVGView	produces Scalable Vector Graphics with controls to navigate through multi-page SVG documents.
SVGZ	produces compressed Scalable Vector Graphics.

For more information, see Printing with SAS in *SAS Language Reference: Concepts*.

SAS Logging Facility Language Elements

The SAS logging facility is a flexible, configurable logging subsystem that you can use to collect, categorize, and filter log events and write them to a variety of output

devices. The SAS language now includes autocall macros, functions, and DATA step component objects for creating logging facility components that categorize log events. The logging facility and the SAS log are two separate logging systems. For more information, including the reference documentation for the logging facility language elements, see *SAS Logging: Configuration and Programming Reference*. This feature is new for SAS 9.2 Phase 2.

WHERE-Expression Processing

In a WHERE expression, the LIKE operator now supports an escape character. The escape character enables you to search for the percent sign (%) and the underscore (_) characters in values. For more information, see "Syntax of WHERE Expression" in *SAS Language Reference: Concepts*.

DATA Step Java Object

The DATA step component Java object enables you to instantiate Java classes and access fields and methods on the resultant objects. Although the documentation for the DATA step component Java object for SAS 9.2 Phase 1 has been available on **`http://support.sas.com`**, the documentation is available in SAS Help and Documentation for SAS 9.2 Phase 2.

Viewing Help and ODS Output in the Remote Browser

The remote browser has been used in some operating environments in prior releases of SAS to view SAS Help and ODS HTML output. You can now view SAS Help and ODS HTML output, and PDF and RTF output under z/OS, OpenVMS, UNIX, and Windows 64-bit environments. Windows 32-bit environments use the SAS browser to view Help and ODS output.

You enable remote browsing by configuring these system options:

HELPBROWSER= specifies whether you want to use the remote browser or the SAS browser.

HELPHOST= specifies the name of the computer where the remote browser sends Help and ODS output.

HELPPORT= specifies the port number for the remote browser client.

For more information about remote browsing, see the Help documentation for your operating environment: OpenVMS, UNIX, Windows, z/OS

SAS Language Elements

Data Set Options

The DLDMGACTION=NOINDEX data set option has a new argument.

The NOINDEX argument automatically repairs the data set without the indexes and integrity constraints, deletes the index file, updates the data file to reflect the disabled indexes and integrity constraints, and limits the data file to be opened only in INPUT mode.

Formats

- The following formats are new:

  $BASE64X
  converts character data to ASCII text using Base 64 encoding.

  $N8601B
  writes ISO 8601 duration, datetime, and interval forms using the basic notations P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyymmdd*T*hhmmss*.

  $N8601BA
  writes ISO 8601 duration, datetime, and interval forms using the basic notations P*yyyymmdd*T*hhmmss* and *yyyymmdd*T*hhmmss*.

  $N8601E
  writes ISO 8601 duration, datetime, and interval forms using the extended notations P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyy-mm-dd*T*hh:mm:ss*.

  $N8601EA
  writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*.

  $N8601EH
  writes ISO 8601 duration, datetime, and interval forms for the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using a hyphen (-)for omitted components.

  $N8601EX
  writes ISO 8601 duration, datetime, and interval forms for the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using an x for each digit of an omitted component.

  $N8601H
  writes ISO 8601 duration, datetime, and interval forms P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using a hyphen (-)for omitted components in datetime values.

  $N8601X
  writes ISO 8601 duration, datetime, and interval forms P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using an x for each digit of an omitted component in datetime values.

  B8601DA
  writes date values using the IOS 8601 base notation *yyyymmdd*.

  B8601DN
  writes the date from a datetime value using the ISO 8601 basic notation *yyyymmdd*.

  B8601DT
  writes datetime values in the ISO 8601 basic notation *yyyymmdd*T*hhmmssffffff*.

  B8601DZ
  writes datetime values in the Coordinated Universal Time (UTC) time scale using the ISO 8601 datetime and time zone basic notation *yyyymmdd*T*hhmmss*+|–*hhmm*.

B8601LZ
writes time values as local time by appending a time zone offset difference between the local time and UTC, using the ISO 8601 basic time notation *hhmmss+|–hhmm*.

B8601TM
writes time values using the ISO 8601 basic notation *hhmmssffff*.

B8601TZ
adjusts time values to the Coordinated Universal Time (UTC) and writes them using the ISO 8601 basic time notation *hhmmss+|–hhmm*.

BESTD
prints numeric values, lining up decimal places for values of similar magnitude, and prints integers without decimals.

E8601DA
writes date values using the ISO 8601 extended notation *yyyy-mm-dd*.

E8601DN
writes the date from a SAS datetime value using the ISO 8601 extended notation *yyyy-mm-dd*.

E8601DT
writes datetime values in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.

E8601DZ
writes datetime values in the Coordinated Universal Time (UTC) time scale using the ISO 8601 datetime and time zone extended notation *yyyy-mm-dd*T*hh:mm:ss+|–hh:mm*.

E8601LX
writes time values as local time, appending the Coordinated Universal Time (UTC) offset for the local SAS session, using the ISO 8601 extended time notation *hh:mm:ss+|–hh:mm*.

E8601TM
writes time values using the ISO 8601 extended notation *hh:mm:ss.ffffff*.

E8601TZ
adjusts time values to the Coordinated Universal Time (UTC) and writes the values using the ISO 8601 extended notation *hh:mm:ss+|–hh:mm*.

MDYAMPM
writes datetime values in the form *mm/dd/yy<yy> hh:mm* AM|PM. The year can be either two or four digits. This feature is new for SAS 9.2 Phase 2 and later.

PERCENTN
produces percentages, using a minus sign for negative values.

VMSZN
generates VMS and MicroFocus COBOL zoned numeric data.

- The following formats were previously documented in other publications and are now part of this document:

WEEKUw.
writes a week number in decimal format by using the U algorithm.

WEEKVw.
writes a week number in decimal format by using the V algorithm.

WEEKWw.
writes a week number in decimal format by using the W algorithm.

- The following format is enhanced:

  DATEw.
  In addition to writing dates in the form *ddmmmyy* or *ddmmmyyyy*, the DATE*w*. format now writes dates in the form *dd-mmm-yyyy*.

- The following formats are no longer supported and have been removed from the documentation:
  - SIZEK
  - SIZEKB
  - SIZEKMG

Functions and CALL Routines

- In the second maintenance release for SAS 9.2, best practices for custom interval names for date and time functions is new.
- The following functions and CALL routines are new:

  ALLCOMB
  generates all combinations of the values of *n* variables taken *k* at a time in a minimal change order.

  ALLPERM
  generates all permutations of the values of several variables in a minimal change order.

  ARCOSH
  returns the inverse hyperbolic cosine.

  ARSINH
  returns the inverse hyperbolic sine.

  ARTANH
  returns the inverse hyperbolic tangent.

  CALL ALLCOMB
  generates all combinations of the values of *n* variables taken *k* at a time in a minimal change order.

  CALL ALLCOMBI
  generates all combinations of the indices of *n* objects taken *k* at a time in a minimal change order.

  CALL GRAYCODE
  generates all subsets of *n* items in a minimal change order.

  CALL ISO8601_CONVERT
  converts an ISO 8601 interval to datetime and duration values, or converts datetime and duration values to an ISO 8601 interval.

  CALL LEXCOMB
  generates all distinct combinations of the nonmissing values of *n* variables taken *k* at a time in lexicographic order.

  CALL LEXCOMBI
  generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order.

CALL LEXPERK
generates all distinct permutations of the nonmissing values of *n* variables taken *k* at a time in lexicographic order.

CALL LEXPERM
generates all distinct permutations of the nonmissing values of several variables in lexicographic order.

CALL SORTC
sorts the values of character arguments.

CALL SORTN
sorts the values of numeric arguments.

CATQ
concatenates character or numeric values by using a delimiter to separate items and by adding quotation marks to strings that contain the delimiter.

CHAR
returns a single character from a specified position in a character string.

CMISS
counts the number of missing arguments.

COUNTW
counts the number of words in a character expression.

DIVIDE
returns the result of a division that handles special missing values for ODS output.

ENVLEN
returns the length of an environment variable.

EUCLID
returns the Euclidean norm of the nonmissing arguments.

FINANCE
computes financial calculations such as deprecation, maturation, accrued interest, net present value, periodic savings, and internal rates of return.

FINDW
searches a character string for a word.

FIRST
returns the first character in a character string.

GCD
returns the greatest common divisor for one or more integers.

GEODIST
returns the geodetic distance between two latitude and longitude coordinates.

GRAYCODE
generates all subsets of *n* items in a minimal change order.

INTFIT
returns a time interval that is aligned between two dates.

INTGET
returns an interval based on three date or datetime values.

INTSHIFT
returns the shift interval that corresponds to the base interval.

INTTEST
returns 1 if a time interval is valid, and returns 0 if a time interval is invalid.

LCM
returns the smallest multiple that is exactly divisible by every number in a set of numbers.

LCOMB
computes the logarithm of the COMB function—that is, the logarithm of the number of combinations of *n* objects taken *r* at a time.

LEXCOMB
generates all distinct combinations of the nonmissing values of *n* variables taken *k* at a time in lexicographic order.

LEXCOMBI
generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order.

LEXPERK
generates all distinct permutations of the nonmissing values of *n* variables taken *k* at a time in lexicographic order.

LEXPERM
generates all distinct permutations of the nonmissing values of several variables in lexicographic order.

LFACT
computes the logarithm of the FACT (factorial) function.

LOG1PX
returns the log of 1 plus the argument.

LPERM
computes the logarithm of the PERM function—that is, the logarithm of the number of permutations of *n* objects, with the option of including *r* number of elements.

LPNORM
returns the L_p norm of the second argument and subsequent nonmissing arguments.

MD5
returns the result of the message digest of a specified string.

MODEXIST
determines whether a software image exists in the version of SAS that you have installed.

MSPLINT
returns the ordinate of a monotonicity-preserving interpolating spline.

RENAME
renames a member of a SAS library, an external file, or a directory.

SUMABS
returns the sum of the absolute values of the nonmissing arguments.

TRANSTRN
removes or replaces all occurrences of a substring in a character string.

WHICHC
searches for a character value that is equal to the first argument, and returns the index of the first matching value.

WHICHN
searches for a numeric value that is equal to the first argument, and returns the index of the first matching value.

ZIPCITYDISTANCE
returns the geodetic distance between two ZIP code locations.

- The descriptions of the arguments in the following functions are enhanced:

  DOPEN
  opens a directory, and returns a directory identifier value.

  EXIST
  verifies the existence of a SAS library member.

  FOPEN
  opens an external file and returns a file identifier value.

  FEXIST
  verifies the existence of an external file that is associated with a fileref.

  FILENAME
  assigns or deassigns a fileref to an external file, a directory, or an output device.

  FILEREF
  verifies whether a fileref has been assigned for the current SAS session.

  LIBNAME
  assigns or deassigns a libref for a SAS library.

  LIBREF
  verifies that a libref has been assigned.

  MOPEN
  opens a file by directory ID and member name, and returns either the file identifier or a 0.

  PATHNAME
  returns the physical name of a SAS library or an external file, or returns a blank.

- The following functions were previously in Risk Dimensions, and are now in Base SAS:

  BLACKCLPRC
  calculates the call price for European options on futures, based on the Black model.

  BLACKPTPRC
  calculates the put price for European options on futures, based on the Black model.

  BLKSHCLPRT
  calculates the call price for European options, based on the Black-Scholes model.

BLKSHPTPRT
calculates the put price for European options, based on the Black-Scholes model.

GARKHCLPRC
calculates the call price for European options on stocks, based on the Garman-Kohlhagen model.

GARKHPTPRC
calculates the put price for European options on stocks, based on the Garman-Kohlhagen model.

MARGRCLPRC
calculates the call price for European options on stocks, based on the Margrabe model.

MARGRPTPRC
calculates the put price for European options on stocks, based on the Margrabe model.

- The following functions were previously in SAS/ETS, and are now in Base SAS:

  INTCINDEX
  returns the cycle index, given a date, time, or datetime value.

  INTCYCLE
  returns the date, time, or datetime interval at the next higher seasonal cycle, given a date, time, or datetime interval.

  INTFMT
  returns a recommended format, given a date, time, or datetime interval.

  INTINDEX
  returns the seasonal index, given a date, time, or datetime interval and value.

  INTSEAS
  returns the length of the seasonal cycle, given a date, time, or datetime interval.

- The following functions were previously in SAS High-Performance Forecasting, and are now in Base SAS:

  HOLIDAY
  returns the date of the specified holiday for the specified year.

  NWKDOM
  returns the date for the *n*th occurrence of a weekday for the specified month and year.

- The following functions were moved from *SAS Language Reference: Dictionary* to the SAS/IML documentation:

  MODULEIC
  calls an external routine and returns a character value (in the IML environment only).

  MODULEIN
  calls an external routine and returns a numeric value (in the IML environment only).

  CALL MODULEI
  calls an external routine without any return code (in the IML environment only).

- The following functions and CALL routines are enhanced:

CALL POKE
can now write floating-point numbers directly into memory on a 32–bit platform.

CALL POKELONG
can now write floating-point numbers directly into memory on 32-bit and 64-bit platforms.

CALL SCAN
returns the position and length of a given word from a character expression.

DATDIF
now has new values for the *basis* argument, and has a reference to a document that is published by the Securities Industry Association.

FSEP
now has an optional argument for a hexadecimal character delimiter.

INDEX
now has an example that shows how leading and trailing spaces are handled.

INDEXW
can now have alternate delimiters. If you use an alternate delimiter, then INDEXW does not recognize the end of the text as the end data. Another example has also been added to the function.

INTCK
now has a fifth argument in the syntax. Retail calendar intervals that are ISO 8601 compliant, and custom intervals have been added.

INTNX
can now use retail calendar intervals that are ISO 8601 compliant.

INTCINDEX, INTCYCLE, INTFIT, INTFMT, INTGET, INTINDEX, INTSEAS, INTSHIFT, and INTTEST
are now able to use retail calendar intervals that are ISO 8601 compliant.

LAG
now has more information about memory limits.

LIBNAME
now has sections that explain how to use the LIBNAME function with one, two, three, and four arguments.

OPEN
has a new fourth argument. This argument specifies whether the first argument is a two-level name (data set name) or a filename.

SCAN
returns the *n*th word from a character expression.

TRANSTRN
has been rewritten.

TRANWRD
has an updated Comparisons section and a new example.

WEEK
now has enhanced documentation for the U, V, and W descriptors.

ZIPSTATE
now has information about Army Post Office (APO) and Fleet Post Office (FPO) codes.

- The RX set of functions and CALL routines have been removed from the documentation. They have been replaced by a set of PRX functions and CALL routines, which have been available in previous versions of SAS, and which provide superior functionality.

  The following table lists the RX functions and CALL routines and their PRX replacements:

RX Function	PRX Replacement
CALL RXCHANGE	CALL PRXCHANGE
CALL RXFREE	CALL PRXFREE
CALL RXSUBSTR	CALL PRXSUBSTR
RXMATCH	PRXMATCH
RXPARSE	PRXPARSE

- The SCANQ function and the CALL SCANQ routine have been removed from the documentation and replaced by the superior functionality of the SCAN function and CALL SCAN routine.

Informats

- The following informats are new:

  $BASE64X
  converts ASCII text to character data by using Base 64 encoding.

  $N8601B
  reads complete, truncated, and omitted forms of ISO 8601 duration, datetime, and interval values that are specified in either the basic or extended notations.

  $N8601E
  reads ISO 8601 duration, datetime, and interval values that are specified in the extended notation.

  B8601DA
  reads date values that are specified in the ISO 8601 basic notation *yyyymmdd*.

  B8601DN
  reads date values that are specified the ISO 8601 basic notation *yyyymmdd* and returns SAS datetime values where the time portion of the value is 000000.

  B8601DT
  reads datetime values that are specified in the ISO 8601 basic notation *yyyymmdd*T*hhmmssffffff*.

  B8601DZ
  reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using the ISO 8601 datetime basic notation *yyyymmdd*T*hhmmss*+|–*hhmm* or *yyyymmdd*T*hhmmssffffff*Z.

B8601TM
reads time values that are specified in the ISO 8601 basic notation *hhmmssffffff*.

B8601TZ
reads time values that are specified in the ISO 8601 basic time notation *hhmmssfffff*+|–*hhmm* or *hhmmssffffff*Z.

E8601DA
reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*.

E8601DN
reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd* and returns SAS datetime values where the time portion of the value is 000000.

E8601DT
reads datetime values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.

E8601DZ
reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using the ISO 8601 datetime extended notation *hh:mm:ss*+|–*hh:mm.fffff* or*hh:mm:ss.fffff*Z.

E8601LZ
reads Coordinated Universal Time (UTC) values that are specified in the ISO 8601 extended notation *hh:mm:ss*+|–*hh:mm.fffff* or *hh:mm:ss.fffff*Z and converts them to the local time.

E8601TM
reads time values that are specified in the ISO 8601 extended notation *hh:mm:ss.ffffff*.

E8601TZ
reads time values that are specified in the ISO 8601 extended time notation *hh:mm:ss*+|–*hh:mm.ffffff* or *hh:mm:ss*Z.

S3270FZDB
reads zoned decimal data in which zeros have been left blank. This feature is new for SAS 9.2 Phase 2 and later.

VMSZN
reads VMS and MicroFocus COBOL zoned numeric data.

- The following informat is enhanced:

  TRAILSGN
  In addition to reading trailing plus (+) and minus (–) signs, the TRAILSGN informat now reads values that contain commas.

- The following informats were previously documented in other publications and are now part of this document:

  WEEKUw.
  reads the format of the number-of-week value within the year and returns a SAS date value using the U algorithm.

  WEEKVw.
  reads the format of the number-of-week value within the year and returns a SAS date value using the V algorithm.

WEEKWw.
reads the format of the number-of-week value within the year and returns a SAS date value using the W algorithm.

- The SIZEKMG informat is no longer supported and has been removed from the documentation.

Statements

- The following statements are new:

  CHECKPOINT EXECUTE_ALWAYS
  enables you to execute the DATA or PROC step that immediately follows without considering the checkpoint-restart data.

  FILENAME, SFTP Access Method
  enables you to access remote files by using the SFTP protocol.

  SYSECHO
  enables IOM clients to manually track the progress of a segment of a submitted SAS program.

- The following statements are enhanced:

  %INCLUDE
  - The filename of a file that is located in an aggregate storage location and does not have a valid SAS name can be used as a fileref if the filename is enclosed in quotation marks.
  - The maximum line limit is now 6K.

  ABORT
  Two new optional arguments enable you to do the following:
  - cause the execution of the submitted statements to be canceled.
  - suppress the output of all variables to the SAS log.

  ATTRIB
  The TRANSCODE=NO attribute is not supported by some SAS Workspace Server clients. In SAS 9.2, if the attribute is not supported, variables with TRANSCODE=NO are replaced (masked) with asterisks (*). Before SAS 9.2, variables with TRANSCODE=NO were transcoded.

  BY
  The BY statement honors the linguistic collation of data that is sorted by using the SORT procedure with the SORTSEQ=LINGUISTIC option.

  DATA
  Three new optional arguments enable you to do the following:
  - write a note to the SAS log for the beginning and end of each level of nesting DO statements.
  - specify the maximum number of nested LINK statements.
  - suppress the output of all variables to the SAS log.

  DECLARE
  - Data set options can now be used with the dataset: argument tag.
  - Three new argument tags enable you to do the following:
    - maintain a summary count of hash object keys.
    - ignore duplicate keys when loading a data set into the hash object.
    - specify whether multiple data items are allowed for each key.

FILE

- The filename of a file that is located in an aggregate storage location and does not have a valid SAS name can be used as a fileref if the filename is enclosed in quotation marks.
- A new option enables you to specify a character string as an alternate delimiter (other than a blank) to be used for LIST output.

FILENAME, CATALOG Access Method

You can now specify RECFM=S (stream–record format).

FILENAME, EMAIL (SMTP) Access Method

- You can now specify a file attachment without an extension.
- A new option enables you to specify the priority of the e-mail message.

FILENAME, FTP Access Method

Seven new FTP options enable you to do the following:

- specify the name of an authentication domain metadata object that references credentials (user ID and password) in order to connect to the FTP server without your having to explicitly specify the credentials.
- specify that the member type of DATA is automatically appended to the member name when you use the DIR option.
- enable autocall macro retrieval of lowercase directory or member names from FTP servers.
- save the user ID and password after the user ID and password prompt are successfully executed.
- specify the line delimiter to use for variable-record formats: carriage return followed by a line feed, a line feed only, or a NULL character.
- specify the length of the FTP server response message.
- in the second maintenance release for SAS 9.2, specify an FTP response wait time in milliseconds.

FILENAME, SFTP Access Method

In SAS 9.2 Phase 2 and later, two new SFTP options enable you to do the following:

- specify the fully qualified pathname and the filename of the batch file that contains the SFTP commands. These commands are submitted when the SFTP access method is executed.
- specify an SFTP response wait time in milliseconds.

FILENAME, URL Access Method

- N can now be used as an alias for a stream-record format (RECFM=S).
- Five new URL options enable you to do the following:
  - specify the name of an authentication domain metadata object that references credentials (user ID and password) in order to connect to the proxy or Web server without your having to explicitly specify the credentials.
  - specify a fileref to which the header information is written when a file is opened using the URL access method. The header information is the same information that is written to the SAS log.
  - specify a user name with which you can access the proxy server.
  - specify a password with which you can access the proxy server.
  - specify the line delimiter to use when RECFM=V.

FILENAME, WebDAV Access Method

- For SAS 9.2 Phase 2 and later, the FILENAME statement, WebDAV Access Method is available for use in the z/OS operating environment.
- The SASBAMW keyword in the FILENAME statement syntax has been changed to WEBDAV.
- Three new WebDAV options enable you to do the following:
  - access directory files.
  - specify that a file extension is automatically appended to the filename when you use the DIR option.
  - retry lowercase directory or member names from WebDAV servers by using an autocall macro.

FOOTNOTE

a new argument enables you to specify formatting options for the ODS HTML, RTF, and PRINTER(PDF) destinations.

INFILE

- The filename of a file that is located in an aggregate storage location and does not have a valid SAS name can be used as a fileref if the filename is enclosed in quotation marks.
- A new option enables you to specify a character string as an alternate delimiter (other than a blank) to be used for LIST input.
- A new optional argument specifies the type of device or the access method that is used if the fileref points to an input or output device or location that is not a physical file.

LIBNAME for WebDAV Server Access

- When you assign a libref to a file on a WebDAV server, the path (URL location), user ID, and password are associated with that libref. After the first libref is assigned, the user ID and password will be validated on subsequent attempts to assign another libref to the same library.
- SAS will honor a lock request on a file on a WebDAV server only if the file is already locked by another user.
- Two new WebDAV options enable you to do the following:
  - specify the name of an authentication domain metadata object that references credentials (user ID and password) in order to connect to the WebDAV server without your having to explicitly specify the credentials.
  - prompt the user for an ID and password.

MERGE

a new argument enables you to specify at least two existing SAS data sets by using either a numbered range list or a named prefix list.

SET

- a new argument creates and names a variable that stores the name of the SAS data set from which the current observation is read. The stored name can be a data set name or a physical name. The physical name is the name by which the operating environment recognizes the file.
- a new argument enables you to specify at least two existing SAS data sets by using either a numbered range list or a named prefix list.

TITLE
added an argument that enables you to specify formatting options for the ODS HTML, RTF, and PRINTER(PDF) destinations.

System Options

- The following system options are new:

CGOPTIMIZE
specifies the level of optimization to perform during code optimization. This feature is new for SAS 9.2 Phase 2 and later.

CMPMODEL=
specifies the output model type for the MODEL procedure.

DEFLATION=
specifies the level of compression for device drivers that support the Deflate compression algorithm.

DMSPGMLINESIZE=
specifies the maximum number of characters in a Program Editor line.

EMAILFROM
when sending an e-mail that uses SMTP, specifies whether the e-mail option FROM is required in either the FILE or the FILENAME statement.

FILESYNC=
specifies when operating system buffers that contain contents of permanent SAS files are written to disk.

FONTEMBEDDING
specifies whether font embedding is enabled in Universal Printer and SAS/GRAPH printing.

FONTRENDERING=
specifies whether SAS/GRAPH devices that are based on the SASGDGIF, SASGDTIF, and SASGDIMG modules render fonts by using the operating system or by using the FreeType font engine.

GSTYLE
specifies whether ODS styles can be used in the generation of graphs that are stored as GRSEG catalog entries.

HELPBROWSER=
specifies the browser to use for SAS Help and ODS output. This feature is new for SAS 9.2 Phase 2 and later.

HELPHOST=
specifies the name of the computer where the remote browser is to send Help and ODS output. This feature is new for SAS 9.2 Phase 2 and later.

HELPPORT=
specifies the port number for the remote browser client. This feature is new for SAS 9.2 Phase 2 and later.

HTTPSERVERPORTMAX=
specifies the highest port number that can be used by the SAS HTTP server for remote browsing. This feature is new for SAS 9.2 Phase 2 and later.

HTTPSERVERPORTMIN=
specifies the lowest port number that can be used by the SAS HTTP server for remote browsing. This feature is new for SAS 9.2 Phase 2 and later.

IBUFNO=
specifies an optional number of extra buffers to be allocated for navigating an index file. SAS automatically allocates a minimal number of buffers in order to navigate the index file. Typically, you do not need to specify extra buffers. However, using IBUFNO= to specify extra buffers could improve execution time by limiting the number of input/output operations that are required for a particular index file.

INTERVALDS=
specifies a SAS data set that contains user-supplied holidays that can be used by the INTNX and INTCK functions. This feature is new for SAS 9.2 Phase 2 and later.

JPEGQUALITY
specifies the JPEG quality factor that determines the ratio of image quality to the level of compression for JPEG files processed by the SAS/GRAPH JPEG device driver.

LRECL=
specifies the default logical record length to use for reading and writing external files.

PDFACCESS
specifies whether text and graphics from PDF documents can be read by screen readers for the visually impaired.

PDFASSEMBLY
specifies whether PDF documents can be assembled.

PDFCOMMENT
specifies whether PDF document comments can be modified.

PDFCONTENT
specifies whether the contents of a PDF document can be changed.

PDFCOPY
specifies whether text and graphics from a PDF document can be copied.

PDFFILLIN
specifies whether PDF forms can be filled in.

PDFPAGELAYOUT
specifies the page layout for PDF documents.

PDFPAGEVIEW
specifies the page viewing mode for PDF documents.

PDFPASSWORD
specifies the password to use to open a PDF document and the password used by a PDF document owner.

PDFPRINT
specifies the resolution to print PDF documents.

PDFSECURITY
specifies the printing permissions for PDF documents.

PRIMARYPROVIDERDOMAIN=
specifies the domain name of the primary authentication provider. This feature is new for SAS 9.2 Phase 2 and later.

RLANG
specifies whether SAS executes R language statements.

S2V
specifies the starting position to begin reading a file specified in a %INCLUDE statement, an autoexec file, or an autocall macro file with a variable length format.

SORTVALIDATE
specifies whether the SORT procedure verifies that a data set is sorted according to the variables in the BY statement when the sort indicator metadata indicates a user-specified sort order.

SQLCONSTDATETIME
specifies whether the SQL procedure replaces references to the DATE, TIME, DATETIME, and TODAY functions in a query with their equivalent constant values before the query executes.

SQLREDUCEPUT
for the SQL procedure, specifies the engine type that a query uses for which optimization is performed by replacing a PUT function in a query with a logically equivalent expression.

SQLREDUCEPUTOBS
for the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the minimum number of observations that must be in a table for PROC SQL to consider optimizing the PUT function in a query.

SQLREDUCEPUTVALUES=
for the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the minimum number of SAS format values that can exist in a PUT function expression in order for PROC SQL to consider optimizing the PUT function in a query.

SQLREMERGE
specifies whether the SQL procedure can process queries that use remerging of data.

SQLUNDOPOLICY=
specifies whether the SQL procedure keeps or discards updated data if errors occur while the data is being updated.

STEPCHKPT
specifies whether to run a batch program in checkpoint-restart mode. In checkpoint-restart mode, if a batch program terminates during execution, the program can be restarted beginning with the DATA or PROC step that was executing when the program terminated.

STEPCHKPTLIB
specifies the libref which identifies the library that contains the checkpoint-restart data.

STEPRESTART
specifies whether to start a batch program using the checkpoint data.

SVGCONTROLBUTTONS
specifies whether to display the paging control buttons and an index in a multi-page SVG document.

SVGHEIGHT
specifies the height of the viewport unless the SVG output is embedded in another SVG output; specifies the value of the **HEIGHT** attribute of the outermost **<svg>** element in the SVG file.

SVGPRESERVEASPECTRATIO
specifies whether to force uniform scaling of SVG output; sets the **preserveAspectRatio** attribute on the outermost **<svg>** element.

SVGTITLE
specifies the title in the title bar of the SVG output; specifies the value of the **<title>** element in the SVG file.

SVGVIEWBOX
specifies the coordinates, width, and height that are used to set the **viewBox** attribute on the outermost **<svg>** element, which enables SVG output to scale to the viewport.

SVGWIDTH
specifies the width of the viewport unless the SVG output is embedded in another SVG output; specifies the value of the **width** attribute of the outermost **<svg>** element in the SVG file.

SVGX
specifies the x-axis coordinate of one corner of the rectangular region into which an embedded **<svg>** element is placed; specifies the **x** attribute on the outermost **<svg>** element of the SVG file.

SVGY
specifies the y-axis coordinate of one corner of the rectangular region into which an embedded **<svg>** element is placed; specifies the **y** attribute on the outermost **<svg>** element of the SVG file.

UPRINTCOMPRESSION
specifies whether to enable compression of Universal Printer and SAS/GRAPH print files.

VARLENCHK=
specifies the type of message to write to the SAS log if the length of a variable is increased when the input data set is read using the SET, MERGE, UPDATE, or MODIFY statements. This option is new for SAS 9.2 Phase 2.

- The following system options have a new argument:

DLDMGACTION=NOINDEX
For data sets, automatically repairs the data set without the indexes and integrity constraints, deletes the index file, updates the data file to reflect the disabled indexes and integrity constraints, and limits the data file to be opened only in INPUT mode.

CMPOPT=FUNCDIFFERENCING
specifies whether analytic derivatives are computed for user-defined functions.

- The following system options are enhanced:

ECHOAUTO
SAS writes the autoexec file statements to the SAS log.

EMAILHOST
You can now specify multiple Simple Mail Transfer Protocol (SMTP) mail servers.

EMAILPW
In the third maintenance release for SAS 9.2, you can use encoded e-mail passwords. These passwords are encoded with PROC PWENCODE.

E-mail system options
All e-mail system options can now be set at any time. They are no longer restricted to being set when SAS starts.

OVP
The default value for the OVP system option is now NOOVP.

SYSPRINTFONT=
You can specify the name of a Universal Printer to which the SYSPRINTFONT system option setting applies.

- The syntax for the following system options is different when these system options are used after SAS starts, as compared to the syntax that is used when SAS starts. For the syntax to use when SAS starts, see the documentation for your operating environment. This feature is new for SAS 9.2 Phase 2:

  APPEND=
  Appends a value to the existing value of the specified system option.

  INSERT=
  Inserts the specified value as the first value of the specified system option.

- The following system options are no longer supported and have been removed from the documentation:

  BATCH
  no longer has an impact on the settings for the LINESIZE, OVP, PAGESIZE, and SOURCE system options when SAS executes.

  GISMAPS
  SAS 9.2 no longer supplies U.S. Census Tract maps for SAS/GIS.

- The following system option has been deleted from SAS:

  CONSOLELOG=

DATA Step Object Attributes, Operators, and Methods

- For SAS 9.2 Phase 2 and later, the Java object language elements in Chapter 10, "Java Object Language Elements," on page 2145 are now documented in *SAS Language Reference: Dictionary*.
- The following hash and hash iterator methods are new:

  CLEAR
  removes all items from the hash object without deleting the hash object instance.

  EQUALS
  determines whether two hash objects are equal.

  FIND_NEXT
  sets the current list item to the next item in the current key's multiple item list and sets the data for the corresponding data variables.

  FIND_PREV
  sets the current list item to the previous item in the current key's multiple item list and sets the data for the corresponding data variables.

  HAS_NEXT
  determines whether there is a next item in the current key's multiple data item list.

HAS_PREV
: determines whether there is a previous item in the current key's multiple data item list.

REF
: consolidates the FIND and ADD methods into a single method call.

REMOVEDUP
: removes the data that is associated with the specified key's current data item from the hash object.

REPLACEDUP
: replaces the data that is associated with the current key's current data item with new data.

SETCUR
: specifies a starting key item for iteration.

SUM
: retrieves the summary value for a given key from the hash table and stores the value in a DATA step variable.

SUMDUP
: retrieves the summary value for the current data item of the current key and stores the value in a DATA step variable.

- The following hash object method is enhanced:

  DEFINEDONE
  : added an optional argument that enables recovery from memory failure when loading a data set into a hash object.

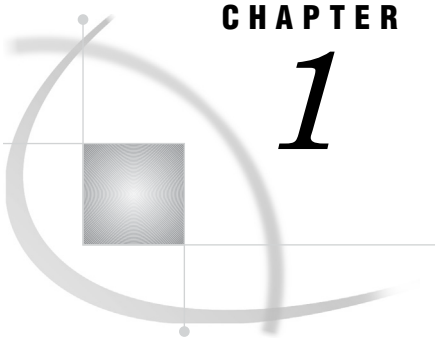
- The following hash object attribute is new:

  ITEM_SIZE
  : returns the number of items in the hash object.

- The _NEW_ statement has been reclassified as an operator.
- For SAS 9.2 Phase 2 and later, the items in a multiple data item list are now maintained in the order in which you insert them.

PART 1

Dictionary of Language Elements



CHAPTER

1

Introduction to the SAS 9.2 Language Reference: Dictionary

The SAS Language Reference: Dictionary

SAS Language Reference: Dictionary provides detailed reference information for the major language elements of Base SAS software:

- data set options
- formats
- functions and CALL routines
- informats
- statements
- SAS system options.
- hash and hash iterator DATA step component object attributes and methods
- Java DATA step component object attributes and methods

It also includes the following four appendixes:

- DATA step debugger
- Perl Regular Expression (PRX) Metacharacters
- SAS utility macro
- Recommended reading.

For extensive conceptual information about the SAS System and the SAS language, including the DATA step, see *SAS Language Reference: Concepts*.

Syntax Conventions for the SAS Language

Overview of Syntax Conventions for the SAS Language

SAS uses standard conventions in the documentation of syntax for SAS language elements. These conventions enable you to easily identify the components of SAS syntax. The conventions can be divided into these parts:

- syntax components
- style conventions
- special characters
- references to SAS libraries and external files

Syntax Components

The components of the syntax for most language elements include a keyword and arguments. For some language elements only a keyword is necessary. For other language elements the keyword is followed by an equal sign (=).

keyword
: specifies the name of the SAS language element that you use when you write your program. Keyword is a literal that is usually the first word in the syntax. In a CALL routine, the first two words are keywords.

  In the following examples of SAS syntax, the keywords are the first words in the syntax:

  CHAR (*string, position*)

  CALL RANBIN (*seed, n, p, x*);

  ALTER (*alter-password*)

  BEST *w*.

  REMOVE <*data-set-name*>

  In the following example, the first two words of the CALL routine are the keywords:

  CALL RANBIN(*seed*, *n*, *p*, *x*)

  The syntax of some SAS statements consists of a single keyword without arguments:

  DO;
  ... *SAS code* ...
  END;

  Some system options require that one of two keyword values be specified:

  DUPLEX | NODUPLEX

argument
: specifies a numeric or character constant, variable, or expression. Arguments follow the keyword or an equal sign after the keyword. The arguments are used by SAS to process the language element. Arguments can be required or optional. In the syntax, optional arguments are enclosed between angle brackets.

In the following example, *string* and *position* follow the keyword CHAR. These arguments are required arguments for the CHAR function:

CHAR (*string, position*)

Each argument has a value. In the following example of SAS code, the argument *string* has a value of 'summer', and the argument *position* has a value of 4:

```
x=char('summer', 4);
```

In the following example, *string* and *substring* are required arguments, while *modifiers* and *startpos* are optional.

FIND(*string, substring* <,*modifiers*> <,*startpos*>

Note: In most cases, example code in SAS documentation is written in lowercase with a monospace font. You can use uppercase, lowercase, or mixed case in the code that you write. △

Style Conventions

The style conventions that are used in documenting SAS syntax include uppercase bold, uppercase, and italic:

UPPERCASE BOLD identifies SAS keywords such as the names of functions or statements. In the following example, the keyword ERROR is written in uppercase bold:

ERROR<message>;

UPPERCASE identifies arguments that are literals.

In the following example of the CMPMODEL= system option, the literals include BOTH, CATALOG, and XML:

CMPMODEL = BOTH | CATALOG | XML

italics identifies arguments or values that you supply. Items in italics represent user-supplied values that are either one of the following:

- nonliteral arguments

  In the following example of the LINK statement, the argument *label* is a user-supplied value and is therefore written in italics:

  LINK *label*;

- nonliteral values that are assigned to an argument

  In the following example of the FORMAT statement, the argument DEFAULT is assigned the variable *default-format*:

  FORMAT = *variable-1* <, ..., *variable-n format* ><DEFAULT = *default-format*>;

Items in italics can also be the generic name for a list of arguments from which you can choose (for example, *attribute-list*). If more than one of an item in italics can be used, the items are expressed as *item-1, ..., item-n*.

Special Characters

The syntax of SAS language elements can contain the following special characters:

= an equal sign identifies a value for a literal in some language elements such as system options.

In the following example of the MAPS system option, the equal sign sets the value of MAPS:

MAPS = *location-of-maps*

< > angle brackets identify optional arguments. Any argument that is not enclosed in angle brackets is required.

In the following example of the CAT function, at least one item is required:

CAT (*item-1* <, ..., *item-n*>)

| a verticle bar indicates that you can choose one value from a group of values. Values that are separated by the vertical bar are mutually exclusive.

In the following example of the CMPMODEL= system option, you can choose only one of the arguments:

CMPMODEL = BOTH | CATALOG | XML

... an ellipsis indicates that the argument or group of arguments following the ellipsis can be repeated. If the ellipsis and the following argument are enclosed in angle brackets, then the argument is optional.

In the following example of the CAT function, the ellipsis indicates that you can have multiple optional items:

CAT (*item-1* <, ..., *item-n*>)

'*value*' or "*value*" indicates that an argument enclosed in single or double quotation marks must have a value that is also enclosed in single or double quotation marks.

In the following example of the FOOTNOTE statement, the argument *text* is enclosed in quotation marks:

FOOTNOTE <*n*> <*ods-format-options 'text'* | "*text*">;

; a semicolon indicates the end of a statement or CALL routine.

In the following example each statement ends with a semicolon:

```
data namegame;
   length color name $8;
   color = 'black';
   name = 'jack';
   game = trim(color) || name;
run;
```

References to SAS Libraries and External Files

Many SAS statements and other language elements refer to SAS libraries and external files. You can choose whether to make the reference through a logical name (a

libref or fileref) or use the physical filename enclosed in quotation marks. If you use a logical name, you usually have a choice of using a SAS statement (LIBNAME or FILENAME) or the operating environment's control language to make the association. Several methods of referring to SAS libraries and external files are available, and some of these methods depend on your operating environment.

In the examples that use external files, SAS documentation uses the italicized phrase *file-specification*. In the examples that use SAS libraries, SAS documentation uses the italicized phrase *SAS-library*. Note that *SAS-library* is enclosed in quotation marks:

```
infile file-specification obs = 100;
libname libref 'SAS-library';
```

CHAPTER

2

SAS Data Set Options

Definition of Data Set Options

Data set options specify actions that apply only to the SAS data set with which they appear. They let you perform the following operations:

- renaming variables
- selecting only the first or last *n* observations for processing
- dropping variables from processing or from the output data set
- specifying a password for a data set

Syntax

Specify a data set option in parentheses after a SAS data set name. To specify several data set options, separate them with spaces.

(*option-1=value-1<...option-n=value-n>*)

These examples show data set options in SAS statements:

```
data scores(keep=team game1 game2 game3);
data mydata(index=(b k) label='label for my data set' drop=p read=secret);
data new(drop=i n index=(j combo=(x1 a1 a20 b1 b50 )));
data idxdup2(compress=yes index=(ok1 ok2 ssn/unique ok3));
proc print data=new(drop=year);
set old(rename=(date=Start_Date));
```

Using Data Set Options

Using Data Set Options with Input or Output SAS Data Sets

Most SAS data set options can apply to either input or output SAS data sets in DATA steps or procedure (PROC) steps. If a data set option is associated with an input data set, the action applies to the data set that is being read. If the option appears in the DATA statement or after an output data set specification in a PROC step, SAS applies the action to the output data set. In the DATA step, data set options for output data sets must appear in the DATA statement, not in any OUTPUT statements that might be present.

Some data set options, such as COMPRESS=, are meaningful only when you create a SAS data set because they set attributes that exist for the duration of the data set. To change or cancel most data set options, you must re-create the data set. You can change other options (such as PW= and LABEL=) with PROC DATASETS. For more information, see the "DATASETS Procedure" in *Base SAS Procedures Guide*.

When data set options appear on both input and output data sets in the same DATA or PROC step, first SAS applies data set options to input data sets. Then SAS evaluates programming statements or applies data set options to output data sets. Likewise, data set options that are specified for the data set being created are applied after programming statements are processed. For example, when using the RENAME= data set option, the new names are not associated with the variables until the DATA step ends.

In some instances, data set options conflict when they are used in the same statement. For example, you cannot specify both the DROP= and KEEP= data set options for the same variable in the same statement. Timing can also be an issue in some cases. For example, if using KEEP= and RENAME= on a data set specified in the SET statement, KEEP= needs to use the original variable names. SAS processes KEEP= before the data set is read. The new names specified in RENAME= apply to the programming statements that follow the SET statement.

How Data Set Options Interact with System Options

Many system options and data set options share the same name and have the same function. System options remain in effect for all DATA and PROC steps in a SAS job or session unless they are respecified.

The data set option overrides the system option for the data set in the step in which it appears. In this example, the OBS= system option in the OPTIONS statement specifies that only the first 100 observations are processed from any data set within the SAS job. The OBS= data set option in the SET statement, however, overrides the system option for data set TWO and specifies that only the first five observations are read from data set TWO. The PROC PRINT step prints the data set FINAL. This data set contains the first 5 observations from data set TWO, followed by the first 100 observations from data set THREE:

```
options obs=100;

data final;
   set two(obs=5) three;
run;

proc print data=final;
run;
```

Data Set Options by Category

Table 2.1 Categories and Descriptions of SAS Data Set Options

Category	SAS Data Set Options	Description
Data Set Control	"ALTER= Data Set Option" on page 14	Assigns an ALTER password to a SAS file that prevents users from replacing or deleting the file, and enables access to a read- and write-protected file.
	"BUFNO= Data Set Option" on page 15	Specifies the number of buffers to be allocated for processing a SAS data set.
	"BUFSIZE= Data Set Option" on page 16	Specifies the size of a permanent buffer page for an output SAS data set.
	"CNTLLEV= Data Set Option" on page 18	Specifies the level of shared access to a SAS data set.
	"COMPRESS= Data Set Option" on page 19	Specifies how observations are compressed in a new output SAS data set.
	"DLDMGACTION= Data Set Option" on page 21	Specifies the action to take when a SAS data set in a SAS library is detected as damaged.
	"ENCRYPT= Data Set Option" on page 23	Specifies whether to encrypt an output SAS data set.
	"GENMAX= Data Set Option" on page 27	Requests generations for a new data set, modifies the number of generations for an existing data set, and specifies the maximum number of versions.
	"GENNUM= Data Set Option" on page 28	Specifies a particular generation of a SAS data set.
	"INDEX= Data Set Option" on page 34	Defines an index for a new output SAS data set.
	"LABEL= Data Set Option" on page 37	Specifies a label for a SAS data set.
	"OBSBUF= Data Set Option" on page 44	Determines the size of the view buffer for processing a DATA step view.
	"OUTREP= Data Set Option" on page 46	Specifies the data representation for the output SAS data set.
	"PW= Data Set Option" on page 49	Assigns a READ, WRITE, and ALTER password to a SAS file, and enables access to a password-protected SAS file.
	"PWREQ= Data Set Option" on page 50	Specifies whether to display a dialog box to enter a SAS data set password.
	"READ= Data Set Option" on page 51	Assigns a READ password to a SAS file that prevents users from reading the file, unless they enter the password.
	"REPEMPTY= Data Set Option" on page 54	Specifies whether a new, empty data set can overwrite an existing SAS data set that has the same name.

Category	SAS Data Set Options	Description
	"REPLACE= Data Set Option" on page 55	Specifies whether a new SAS data set that contains data can overwrite an existing data set that has the same name.
	"REUSE= Data Set Option" on page 56	Specifies whether new observations can be written to freed space in compressed SAS data sets.
	"SORTEDBY= Data Set Option" on page 57	Specifies how a data set is currently sorted.
	"SPILL= Data Set Option" on page 59	Specifies whether to create a spill file for non-sequential processing of a DATA step view.
	"TOBSNO= Data Set Option" on page 66	Specifies the number of observations to send in a client/server transfer.
	"TYPE= Data Set Option" on page 66	Specifies the data set type for a specially structured SAS data set.
	"WRITE= Data Set Option" on page 71	Assigns a WRITE password to a SAS file that prevents users from writing to a file, unless they enter the password.
Miscellaneous	"FILECLOSE= Data Set Option" on page 24	Specifies how a tape is positioned when a SAS data set is closed.
Observation Control	"FIRSTOBS= Data Set Option" on page 25	Specifies the first observation that SAS processes in a SAS data set.
	"IN= Data Set Option" on page 33	Creates a Boolean variable that indicates whether the data set contributed data to the current observation.
	"OBS= Data Set Option" on page 39	Specifies the last observation that SAS processes in a data set.
	"POINTOBS= Data Set Option" on page 48	Specifies whether SAS creates compressed data sets whose observations can be randomly accessed or sequentially accessed.
	"WHERE= Data Set Option" on page 67	Specifies specific conditions to use to select observations from a SAS data set.
	"WHEREUP= Data Set Option" on page 69	Specifies whether to evaluate new observations and modified observations against a WHERE expression.
User Control of SAS Index Usage	"IDXNAME= Data Set Option" on page 30	Directs SAS to use a specific index to match the conditions of a WHERE expression.
	"IDXWHERE= Data Set Option" on page 31	Specifies whether SAS uses an index search or a sequential search to match the conditions of a WHERE expression.
Variable Control	"DROP= Data Set Option" on page 22	For an input data set, excludes the specified variables from processing; for an output data set, excludes the specified variables from being written to the data set.

Category	SAS Data Set Options	Description
	"KEEP= Data Set Option" on page 36	For an input data set, specifies the variables to process; for an output data set, specifies the variables to write to the data set.
	"RENAME= Data Set Option" on page 52	Changes the name of a variable.

Dictionary

ALTER= Data Set Option

Assigns an ALTER password to a SAS file that prevents users from replacing or deleting the file, and enables access to a read- and write-protected file.

Valid in: DATA step and PROC steps

Category: Data Set Control

See: ALTER= Data Set Option in the documentation for your operating environment.

Syntax

ALTER=*alter-password*

Syntax Description

alter-password
must be a valid SAS name. See "Rules for Words and Names in the SAS Language" in *SAS Language Reference: Concepts*.

Details

The ALTER= option applies to all types of SAS files except catalogs. You can use this option to assign a password to a SAS file or to access a read-protected, write-protected, or alter-protected SAS file.

When replacing a SAS data set that is protected with an ALTER password, the new data set inherits the ALTER password. To change the ALTER password for the new data set, use the MODIFY statement in the DATASETS procedure.

Note: A SAS password does not control access to a SAS file beyond the SAS system. You should use the operating system-supplied utilities and file-system security controls in order to control access to SAS files outside of SAS. △

See Also

Data Set Options:

"ENCRYPT= Data Set Option" on page 23
"PW= Data Set Option" on page 49
"READ= Data Set Option" on page 51
"WRITE= Data Set Option" on page 71

"File Protection" in *SAS Language Reference: Concepts*

"Manipulating Passwords" in "The DATASETS Procedure" in *Base SAS Procedures Guide*

BUFNO= Data Set Option

Specifies the number of buffers to be allocated for processing a SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

See: BUFNO= Data Set Option in the documentation for your operating environment.

Syntax

BUFNO= *n* | *n*K | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K**
specifies the number of buffers in multiples of 1 (bytes); 1,024 (kilobytes). For example, a value of **`8`** specifies 8 buffers, and a value of **`1k`** specifies 1024 buffers.

***hex*X**
specifies the number of buffers as a hexadecimal value. You must specify the value beginning with a number (0-9), followed by an X. For example, the value **`2dx`** sets the number of buffers to 45 buffers.

MIN
sets the minimum number of buffers to 0, which causes SAS to use the minimum optimal value for the operating environment. This is the default.

MAX
sets the number of buffers to the maximum possible number in your operating environment, up to the largest four-byte, signed integer, which is 2^{31}-1, or approximately 2 billion.

Details

The buffer number is not a permanent attribute of the data set; it is valid only for the current SAS session or job.

BUFNO= applies to SAS data sets that are opened for input, output, or update.

A larger number of buffers can speed up execution time by limiting the number of input and output (I/O) operations that are required for a particular SAS data set. However, the improvement in execution time comes at the expense of increased memory consumption.

To reduce I/O operations on a small data set as well as speed execution time, allocate one buffer for each page of data to be processed. This technique is most effective if you read the same observations several times during processing.

Operating Environment Information: The default value for BUFNO= is determined by your operating environment and is set to optimize sequential access. To improve performance for direct (random) access, you should change the value for BUFNO=. For the default setting and possible settings for direct access, see the BUFNO= data set option in the SAS documentation for your operating environment. △

Comparisons

- If the BUFNO= data set option is not specified, then the value of the BUFNO= system option is used. If both are specified in the same SAS session, the value specified for the BUFNO= data set option overrides the value specified for the BUFNO= system option.
- To request that SAS allocate the number of buffers based on the number of data set pages and index file pages, use the SASFILE global statement.

See Also

Data Set Options:

"BUFSIZE= Data Set Option" on page 16

System Options:

"BUFNO= System Option" on page 1851

Statements:

"SASFILE Statement" on page 1755

BUFSIZE= Data Set Option

Specifies the size of a permanent buffer page for an output SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with output data sets only.

See: BUFSIZE= Data Set Option in the documentation for your operating environment.

Syntax

BUFSIZE= *n* | *n*K | *n*M | *n*G | *hex*X | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G**
specifies the page size in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); or 1,073,741,824 (gigabytes). For example, a value of **`8`** specifies a page size of 8 bytes, and a value of **`4k`** specifies a page size of 4096 bytes.

The default is 0, which causes SAS to use the minimum optimal page size for the operating environment.

***hex*X**
specifies the page size as a hexadecimal value. You must specify the value beginning with a number (0-9), followed by an X. For example, the value **`2dx`** sets the page size to 45 bytes.

MAX
sets the page size to the maximum possible number in your operating environment, up to the largest four-byte, signed integer, which is 2^{31}-1, or approximately 2 billion bytes.

Details

The page size is the amount of data that can be transferred for a single I/O operation to one buffer. The page size is a permanent attribute of the data set and is used when the data set is processed.

A larger page size can speed up execution time by reducing the number of times SAS has to read from or write to the storage medium. However, the improvement in execution time comes at the cost of increased memory consumption.

To change the page size, use a DATA step to copy the data set and either specify a new page or use the SAS default. To reset the page size to the default value in your operating environment, use BUFSIZE=0.

Note: If you use the COPY procedure to copy a data set to another library that is allocated with a different engine, the specified page size of the data set is not retained. △

Operating Environment Information: The default value for BUFSIZE= is determined by your operating environment and is set to optimize sequential access. To improve performance for direct (random) access, you should change the value for BUFSIZE=. For the default setting and possible settings for direct access, see the BUFSIZE= data set option in the SAS documentation for your operating environment. △

See Also

Data Set Options:

"BUFNO= Data Set Option" on page 15

System Options:

"BUFSIZE= System Option" on page 1853

CNTLLEV= Data Set Option

Specifies the level of shared access to a SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Specify for input data sets only.

Syntax

CNTLLEV=LIB | MEM | REC

Syntax Description

LIB
specifies that concurrent access is controlled at the library level. Library-level control restricts concurrent access to only one update process to the library.

MEM
specifies that concurrent access is controlled at the SAS data set (member) level. Member-level control restricts concurrent access to only one update or output process to the SAS data set. If the data set is open for an update or output process, then no other operation can access the data set. If the data set is open for an input process, then other concurrent input processes are allowed but no update or output process is allowed.

REC
specifies that concurrent access is controlled at the observation (record) level. Record-level control allows more than one update access to the same SAS data set, but it denies concurrent update of the same observation.

Details

The CNTLLEV= option specifies the level at which shared update access to a SAS data set is denied. A SAS data set can be opened concurrently by more than one SAS session or by more than one statement, window, or procedure within a single session. By default, SAS procedures permit the greatest degree of concurrent access possible while they guarantee the integrity of the data and the data analysis. Therefore, you do not typically use the CNTLLEV= data set option.

Use this option when

- your application controls the access to the data, such as in SAS Component Language (SCL), SAS/IML software, or DATA step programming
- you access data through an interface engine that does not provide member-level control of the data.

If you use CNTLLEV=REC and the SAS procedure needs member-level control for integrity of the data analysis, SAS prints a warning to the SAS log that inaccurate or unpredictable results can occur if the data are updated by another process during the analysis.

Examples

Example 1: Changing the Shared Access Level In the following example, the first SET statement includes the CNTLLEV= data set option in order to override the default level of shared access from member-level control to record-level control. The second SET statement opens the SAS data set with the default member-level control.

```
set datalib.fuel (cntllev=rec) point=obsnum;
   .
   .
   .
set datalib.fuel;
   by area;
```

COMPRESS= Data Set Option

Specifies how observations are compressed in a new output SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with output data sets only.

Syntax

COMPRESS=NO | YES | CHAR | BINARY

Syntax Description

NO
specifies that the observations in a newly created SAS data set are uncompressed (fixed-length records).

YES | CHAR
specifies that the observations in a newly created SAS data set are compressed (variable-length records) by SAS using RLE (Run Length Encoding). RLE compresses observations by reducing repeated consecutive characters (including blanks) to two-byte or three-byte representations.

Alias: ON

Tip: Use this compression algorithm for character data.

Note: COMPRESS=CHAR is accepted by Version 7 and later versions. △

BINARY
specifies that the observations in a newly created SAS data set are compressed (variable-length records) by SAS using RDC (Ross Data Compression). RDC combines run-length encoding and sliding-window compression to compress the file.

Tip: This method is highly effective for compressing medium to large (several hundred bytes or larger) blocks of binary data (numeric variables). Because the compression function operates on a single record at a time, the record length needs to be several hundred bytes or larger for effective compression.

Details

Compressing a file is a process that reduces the number of bytes required to represent each observation. Advantages of compressing a file include reduced storage requirements for the file and fewer I/O operations necessary to read or write to the data during processing. However, more CPU resources are required to read a compressed file (because of the overhead of uncompressing each observation), and there are situations where the resulting file size might increase rather than decrease.

Use the COMPRESS= data set option to compress an individual file. Specify the option for output data sets only—that is, data sets named in the DATA statement of a DATA step or in the OUT= option of a SAS procedure. Use the COMPRESS= data set option only when you are creating a SAS data file (member type DATA). You cannot compress SAS views, because they contain no data.

After a file is compressed, the setting is a permanent attribute of the file, which means that to change the setting, you must re-create the file. That is, to uncompress a file, specify COMPRESS=NO for a DATA step that copies the compressed file.

Comparisons

The COMPRESS= data set option overrides the COMPRESS= option in the LIBNAME statement and the COMPRESS= system option.

The data set option POINTOBS=YES, which is the default, determines that a compressed data set can be processed with random access (by observation number) rather than sequential access. With random access, you can specify an observation number in the FSEDIT procedure and the POINT= option in the SET and MODIFY statements.

When you create a compressed file, you can also specify REUSE=YES (as a data set option or system option) in order to track and reuse space. With REUSE=YES, new observations are inserted in space freed when other observations are updated or deleted. When the default REUSE=NO is in effect, new observations are appended to the existing file.

POINTOBS=YES and REUSE=YES are mutually exclusive—that is, they cannot be used together. REUSE=YES takes precedence over POINTOBS=YES. That is, if you set REUSE=YES, SAS automatically sets POINTOBS=NO.

The TAPE engine supports the COMPRESS= data set option, but the engine does not support the COMPRESS= system option.

The XPORT engine does not support compression.

See Also

Data Set Options:

"POINTOBS= Data Set Option" on page 48

"REUSE= Data Set Option" on page 56

Statements:

"LIBNAME Statement" on page 1656

System Options:

"COMPRESS= System Option" on page 1872

"REUSE= System Option" on page 1983

"Compressing Data Files" in *SAS Language Reference: Concepts*

DLDMGACTION= Data Set Option

Specifies the action to take when a SAS data set in a SAS library is detected as damaged.

Valid in: DATA step and PROC steps

Category: Data Set Control

Syntax

DLDMGACTION=FAIL | ABORT | REPAIR | NOINDEX | PROMPT

Syntax Description

FAIL
stops the step, issues an error message to the log immediately. This is the default for batch mode.

ABORT
terminates the step, issues an error message to the log, and terminates the SAS session.

REPAIR
automatically repairs and rebuilds indexes and integrity constraints, unless the data file is truncated. You use the REPAIR statement in PROC DATASETS to restore a truncated data set. It issues a warning message to the log. This is the default for interactive mode.

NOINDEX
automatically repairs the data file without the indexes and integrity constraints, deletes the index file, updates the data file to reflect the disabled indexes and integrity constraints, and limits the data file to be opened only in INPUT mode. A warning is written to the SAS log instructing you to execute the PROC DATASETS REBUILD statement to correct or delete the disabled indexes and integrity constraints.

See also: "REBUILD Statement" in the "DATASETS Procedure" in *Base SAS Procedures Guide*

"Recovering Disabled Indexes and Integrity Constraints" in *SAS Language Reference: Concepts*

PROMPT
displays a dialog box that asks you to select the FAIL, ABORT, REPAIR, or NOINDEX action.

DROP= Data Set Option

For an input data set, excludes the specified variables from processing; for an output data set, excludes the specified variables from being written to the data set.

Valid in: DATA step and PROC steps

Category: Variable Control

Syntax

DROP=*variable-1 <...variable-n>*

Syntax Description

variable-1 <...variable-n>
lists one or more variable names. You can list the variables in any form that SAS allows.

Details

If the option is associated with an input data set, the variables are not available for processing. If the DROP= data set option is associated with an output data set, SAS does not write the variables to the output data set, but they are available for processing.

Comparisons

- The DROP= data set option differs from the DROP statement in these ways:
  - In DATA steps, the DROP= data set option can apply to both input and output data sets. The DROP statement applies only to output data sets.
  - In DATA steps, when you create multiple output data sets, use the DROP= data set option to write different variables to different data sets. The DROP statement applies to all output data sets.
  - In PROC steps, you can use only the DROP= data set option, not the DROP statement.
- The KEEP= data set option specifies a list of variables to be included in processing or to be written to the output data set.

Examples

Example 1: Excluding Variables from Input In this example, the variables SALARY and GENDER are not included in processing and they are not written to either output data set:

```
data plan1 plan2;
   set payroll(drop=salary gender);
   if hired<'01jan98'd then output plan1;
   else output plan2;
run;
```

You cannot use SALARY or GENDER in any logic in the DATA step because DROP= prevents the SET statement from reading them from PAYROLL.

Example 2: Processing Variables without Writing Them to a Data Set In this example, SALARY and GENDER are not written to PLAN2, but they are written to PLAN1:

```
data plan1 plan2(drop=salary gender);
   set payroll;
   if hired<'01jan98'd then output plan1;
   else output plan2;
run;
```

See Also

Data Set Options:

"KEEP= Data Set Option" on page 36

Statements:

"DROP Statement" on page 1499

ENCRYPT= Data Set Option

Specifies whether to encrypt an output SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with output data sets only.

Syntax

ENCRYPT=YES | NO

Syntax Description

YES
encrypts the file. This encryption uses passwords that are stored in the data set. At a minimum, you must specify the READ= or the PW= data set option at the same time that you specify ENCRYPT=YES. Because the encryption method uses passwords, you cannot change *any* password on an encrypted data set without re-creating the data set.

CAUTION:
Record all passwords when using ENCRYPT=YES. If you forget the passwords, you cannot reset it without assistance from SAS. The process is time-consuming and resource-intensive. △

NO
does not encrypt the file.

Details

When using ENCRYPT=YES, the following list applies:

- To copy an encrypted data file, the output engine must support the encryption. Otherwise, the data file is not copied.
- Encrypted files work only in SAS 6.11 or later.
- You cannot encrypt SAS views, because they contain no data.
- If the data file is encrypted, all associated indexes are also encrypted.
- Encryption requires approximately the same amount of CPU resources as compression.
- You cannot use PROC CPORT on SAS Proprietary encrypted data files.

Using the ENCRYPT=YES Option

This example creates an encrypted SAS data set using encryption:

```
data salary(encrypt=yes read=green);
   input name $ yrsal bonuspct;
   datalines;
Muriel    34567  3.2
Bjorn     74644  2.5
Freda     38755  4.1
Benny     29855  3.5
Agnetha   70998  4.1
;
```

To use this data set, specify the read password:

```
proc contents data=salary(read=green);
run;
```

See Also

Data Set Options:

"ALTER= Data Set Option" on page 14
"PW= Data Set Option" on page 49
"READ= Data Set Option" on page 51
"WRITE= Data Set Option" on page 71

"SAS Data File Encryption" in *SAS Language Reference: Concepts*

FILECLOSE= Data Set Option

Specifies how a tape is positioned when a SAS data set is closed.

Valid in: DATA step and PROC steps

Category: Miscellaneous

See: FILECLOSE= Data Set Option in the documentation for your operating environment.

Syntax

FILECLOSE=DISP | LEAVE | REREAD | REWIND

Syntax Description

DISP
positions the tape volume according to the disposition specified in the operating environment's control language.

LEAVE
positions the tape at the end of the file that was just processed. Use FILECLOSE=LEAVE if you are not repeatedly accessing the same files in a SAS program but you are accessing one or more subsequent SAS files on the same tape.

REREAD
positions the tape volume at the beginning of the file that was just processed. Use FILECLOSE=REREAD if you are accessing the same SAS data set on tape several times in a SAS program.

REWIND
rewinds the tape volume to the beginning. Use FILECLOSE=REWIND if you are accessing one or more previous SAS files on the same tape, but you are not repeatedly accessing the same files in a SAS program.

Operating Environment Information: These values are not recognized by all operating environments. Additional values are available on some operating environments. See the appropriate sections of the SAS documentation for your operating environment for more information about using SAS libraries that are stored on tape. △

FIRSTOBS= Data Set Option

Specifies the first observation that SAS processes in a SAS data set.

Valid in: DATA step and PROC steps
Category: Observation Control
Restriction: Valid for input (read) processing only.
Restriction: Cannot use with PROC SQL views.

Syntax

FIRSTOBS= *n* | *n*K | *n*M | *n*G | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G**
specifies the number of the first observation to process in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); or 1,073,741,824 (gigabytes). For example, a value of `8` specifies the 8th observation, and a value of `3k` specifies 3,072.

***hex*X**
specifies the number of the first observation to process as a hexadecimal value. You must specify the value beginning with a number (0-9), followed by an X. For example, the value `2dx` sets the 45th observation as the first observation to process.

MIN
sets the number of the first observation to process to 1. This is the default.

MAX
sets the number of the first observation to process to the maximum number of observations in the data set, up to the largest eight-byte, signed integer, which is 2^{63}-1, or approximately 9.2 quintillion observations.

Details

The FIRSTOBS= data set option affects a single, existing SAS data set. Use the FIRSTOBS= system option to affect all steps for the duration of your current SAS session.

FIRSTOBS= is valid for input (read) processing only. Specifying FIRSTOBS= is not valid for output or update processing.

You can apply FIRSTOBS= processing to WHERE processing. For more information, see "Processing a Segment of Data That Is Conditionally Selected" in *SAS Language Reference: Concepts*.

Comparisons

- The FIRSTOBS= data set option overrides the FIRSTOBS= system option for the individual data set.
- While the FIRSTOBS= data set option specifies a starting point for processing, the OBS= data set option specifies an ending point. The two options are often used together to define a range of observations to be processed.
- When external files are read, the FIRSTOBS= option in the INFILE statement specifies which record to read first.

Examples

This PROC step prints the data set STUDY beginning with observation 20:

```
proc print data=study(firstobs=20);
run;
```

This SET statement uses both FIRSTOBS= and OBS= to read-only observations 5 through 10 from the data set STUDY. Data set NEW contains six observations.

```
data new;
  set study(firstobs=5 obs=10);
run;
```

See Also

Data Set Options:

"OBS= Data Set Option" on page 39

Statements:

"INFILE Statement" on page 1591

"WHERE Statement" on page 1792

System Options:

"FIRSTOBS= System Option" on page 1908

GENMAX= Data Set Option

Requests generations for a new data set, modifies the number of generations for an existing data set, and specifies the maximum number of versions.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with output data sets only.

Syntax

GENMAX=*number-of-generations*

Syntax Description

number-of-generations
requests generations for a data set and specifies the maximum number of versions to maintain. The value can be from 0 to 1000. The default is GENMAX=0, which means that no generation data sets are requested.

Details

You use GENMAX= to request generations for a new data set and to modify the number of generations for an existing data set. The first time the data set is replaced, SAS keeps the replaced version and appends a four-character version number to its member name, which includes # and a three-digit number. For example, for a data set named A, a historical version would be A#001.

After generations of a data set are requested, the member name is limited to 28 characters (rather than 32), because the last four characters are reserved for the appended version number. When the GENMAX= data set option is set to 0, the member name can be up to 32 characters.

If you reduce the number of generations for an existing data set, SAS deletes the oldest versions above the new limit.

Examples

Example 1: Requesting Generations When You Create a Data Set This example shows how to request generations for a new data set. The DATA step creates a data set named WORK.A that can have as many as 10 generations (one current version and nine historical versions):

```
data a(genmax=10);
   x=1;
   output;
run;
```

Example 2: Modifying the Number of Generations on an Existing Data Set This example shows how to change the number of generations on the data set MYLIB.A to 4:

```
proc datasets lib=mylib;
   modify a(genmax=4);
run;
```

See Also

Data Set Option:

"GENNUM= Data Set Option" on page 28

"Understanding Generation Data Sets" in "SAS Data Files" in *SAS Language Reference: Concepts*

GENNUM= Data Set Option

Specifies a particular generation of a SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with input data sets only.

Syntax

GENNUM=*integer*

Syntax Description

integer

is a number that references a specific version from a generation group. Specifying a positive number is an absolute reference to a specific generation number that is appended to a data set's name. Specifying a negative number is a relative reference to a historical version in relation to the base version, from the youngest to the oldest. Typically, a value of 0 refers to the current (base) version.

Note: The DATASETS procedure provides a variety of statements for which specifying GENNUM= has additional functionality:

- For the DATASETS and DELETE statements, GENNUM= supports the additional values ALL, HIST, and REVERT.
- For the CHANGE statement, GENNUM= supports the additional value ALL.
- For the CHANGE statement, specifying GENNUM=0 refers to all versions rather than just the base version.

△

Details

After generations for a data set have been requested using the GENMAX= data set option, use GENNUM= to request a specific version. For example, specifying GENNUM=3 refers to the historical version #003, while specifying GENNUM=-1 refers to the youngest historical version.

Note that after 999 replacements, the youngest version would be #999. After 1,000 replacements, SAS rolls over the youngest version number to #000. Therefore, if you want the historical version #000, specify GENNUM=1000.

Both an absolute reference and a relative reference refer to a specific version. A relative reference does not skip deleted versions. Therefore, when working with a generation group that includes one or more deleted versions, using a relative reference results in an error if the version being referenced has been deleted. For example, if you have the base version AIR and three historical versions (AIR#001, AIR#002, and AIR#003) and you delete AIR#002, the following statements return an error, because AIR#002 does not exist. SAS does not assume you mean AIR#003:

```
proc print data=air (gennum= -2);
run;
```

Examples

Example 1: Requesting a Version Using an Absolute Reference This example prints the historical version #003 for data set A, using an absolute reference:

```
proc print data=a(gennum=3);
run;
```

Example 2: Requesting A Version Using a Relative Reference The following PRINT procedure prints the data set three versions back from the base version:

```
proc print data=a(gennum=-3);
run;
```

See Also

Data Set Option:

"GENMAX= Data Set Option" on page 27

"Understanding Generation Data Sets" in "SAS Data Files" in *SAS Language Reference: Concepts*

"The DATASETS Procedure" in *Base SAS Procedures Guide*

IDXNAME= Data Set Option

Directs SAS to use a specific index to match the conditions of a WHERE expression.

Valid in: DATA step and PROC steps
Category: User Control of SAS Index Usage
Restriction: Use with input data sets only
Restriction: Mutually exclusive with IDXWHERE= data set option

Syntax

b IDXNAME=*index-name*

Syntax Description

index-name
specifies the name (up to 32 characters) of a simple or composite index for the SAS data set. SAS does not attempt to determine whether the specified index is the best one or whether a sequential search might be more resource efficient.

Interaction: The specification is not a permanent attribute of the data set and is valid only for the current use of the data set.

Tip: To request that IDXNAME= usage be noted in the SAS log, specify the system option MSGLEVEL=I.

Details

By default, to satisfy the conditions of a WHERE expression for an indexed SAS data set, SAS identifies zero or more candidate indexes that could be used to optimize the WHERE expression. From the list of candidate indexes, SAS determines the one that provides the best performance, or rejects all of the indexes if a sequential pass of the data is expected to be more efficient.

Because the index SAS selects might not always provide the best optimization, you can direct SAS to use one of the candidate indexes by specifying the IDXNAME= data set option. If you specify an index that SAS does not identify as a candidate index, then IDXNAME= does not process the request. That is, IDXNAME= does not allow you to specify an index that would produce incorrect results.

Comparisons

IDXWHERE= enables you to override the SAS decision about whether to use an index.

Example

This example uses the IDXNAME= data set option in order to direct SAS to use a specific index to optimize the WHERE expression. SAS then disregards the possibility that a sequential search of the data set might be more resource efficient and does not attempt to determine whether the specified index is the best one. (Note that the EMPNUM index was not created with the NOMISS option.)

```
data mydata.empnew;
   set mydata.employee (idxname=empnum);
   where empnum < 2000;
run;
```

See Also

Data Set Option:

"IDXWHERE= Data Set Option" on page 31

"Using an Index for WHERE Processing" in *SAS Language Reference: Concepts*

"WHERE-Expression Processing" in *SAS Language Reference: Concepts*

IDXWHERE= Data Set Option

Specifies whether SAS uses an index search or a sequential search to match the conditions of a WHERE expression.

Valid in: DATA step and PROC steps

Category: User Control of SAS Index Usage

Restriction: Use with input data sets only.

Restriction: Mutually exclusive with IDXNAME= data set option

Syntax

IDXWHERE=YES | NO

Syntax Description

YES
tells SAS to choose the best index to optimize a WHERE expression, and to disregard the possibility that a sequential search of the data set might be more resource-efficient.

NO
tells SAS to ignore all indexes and satisfy the conditions of a WHERE expression with a sequential search of the data set.

Note: You cannot use IDXWHERE= to override the use of an index to process a BY statement. △

Details

By default, to satisfy the conditions of a WHERE expression for an indexed SAS data set, SAS decides whether to use an index or to read the data set sequentially. The software estimates the relative efficiency and chooses the method that is more efficient.

You might need to override the software's decision by specifying the IDXWHERE= data set option because the decision is based on general rules that might occasionally not produce the best results. That is, by specifying the IDXWHERE= data set option, you are able to determine the processing method.

Note: The specification is not a permanent attribute of the data set and is valid only for the current use of the data set. △

Note: If you issue the system option MSGLEVEL=I, you can request that IDXWHERE= usage be noted in the SAS log if the setting affects index processing. △

Comparisons

IDXNAME= enables you to direct SAS to use a specific index.

Examples

Example 1: Specifying Index Usage This example uses the IDXWHERE= data set option to tell SAS to decide which index is the best to optimize the WHERE expression. SAS then disregards the possibility that a sequential search of the data set might be more resource-efficient:

```
data mydata.empnew;
   set mydata.employee (idxwhere=yes);
   where empnum < 2000;
```

Example 2: Specifying No Index Usage This example uses the IDXWHERE= data set option to tell SAS to ignore any index and to satisfy the conditions of the WHERE expression with a sequential search of the data set:

```
data mydata.empnew;
   set mydata.employee (idxwhere=no);
   where empnum < 2000;
```

See Also

Data Set Option:

"IDXNAME= Data Set Option" on page 30

"Understanding SAS Indexes" in the "SAS Data Files" section in *SAS Language Reference: Concepts*

"WHERE-Expression Processing" in *SAS Language Reference: Concepts*

IN= Data Set Option

Creates a Boolean variable that indicates whether the data set contributed data to the current observation.

Valid in: DATA step

Category: Observation Control

Restriction: Use with the SET, MERGE, MODIFY, and UPDATE statements only.

Syntax

IN=*variable*

Syntax Description

variable
names the new variable whose value indicates whether that input data set contributed data to the current observation. Within the DATA step, the value of the variable is 1 if the data set contributed to the current observation, and 0 otherwise.

Details

Specify the IN= data set option in parentheses after a SAS data set name in the SET, MERGE, MODIFY, and UPDATE statements only. Values of IN= variables are available to program statements during the DATA step, but the variables are not included in the SAS data set that is being created, unless they are assigned to a new variable.

When you use IN= with BY–group processing, and when a data set contributes an observation for the current BY group, the IN= value is 1. The value remains as long as that BY group is still being processed and the value is not reset by programming logic.

Examples

In this example, IN= creates a new variable, OVERSEAS, that denotes international flights. The variable I has a value of 1 when the observation is read from the NONUSA data set. Otherwise, it has a value of 0. The IF-THEN statement checks the value of I to determine whether the data set NONUSA contributed data to the current observation. If I=1, the variable OVERSEAS receives an asterisk (*) as a value.

```
data allflts;
   set usa nonusa(in=i);
   by fltnum;
   if i then overseas='*';
run;
```

See Also

Statements:

"BY Statement" on page 1452
"MERGE Statement" on page 1679
"MODIFY Statement" on page 1684
"SET Statement" on page 1764
"UPDATE Statement" on page 1787

"BY-Group Processing" in *SAS Language Reference: Concepts*

INDEX= Data Set Option

Defines an index for a new output SAS data set.

Valid in: DATA step and PROC steps
Category: Data Set Control
Restriction: Use with output data sets only.

Syntax

INDEX=(*index-specification-1 ...*<*index-specification-n*>)

Syntax Description

index-specification
names and describes a simple or a composite index to be built. *Index-specification* has this form:

index= (*variable(s)* </UNIQUE> </NOMISS>)

index
is the name of a variable that forms the index or the name you choose for a composite index.

variable or variables
is a list of variables to use in making a composite index.

UNIQUE
specifies that the values of the key variables must be unique. If you specify UNIQUE for a new data set and multiple observations have the same values for the index variables, the index is not created. A slash (/) must precede the UNIQUE option.

NOMISS
excludes all observations with missing values from the index. Observations with missing values are still read from the data set but not through the index. A slash (/) must precede the NOMISS option.

Examples

Example 1: Defining a Simple Index The following INDEX= data set option defines a simple index for the SSN variable:

```
data new(index=(ssn));
```

Example 2: Defining a Composite Index The following INDEX= data set option defines a composite index named CITYST that uses the CITY and STATE variables:

```
data new(index=(cityst=(city state)));
```

Example 3: Defining a Simple and a Composite Index The following INDEX= data set option defines a simple index for SSN and a composite index for CITY and STATE:

```
data new(index=(ssn cityst=(city state)));
```

Example 4: Defining a Simple Index with the UNIQUE Option The following INDEX= data set option defines a simple index for the SSN variable with unique values:

```
data new(index=(ssn /unique));
```

Example 5: Defining a Simple Index with the NOMISS Option The following INDEX= data set option defines a simple index for the SSN variable, excluding all observations with missing values from the index:

```
data new(index=(ssn /nomiss));
```

Example 6: Defining Multiple Indexes Using the UNIQUE and NOMISS Options The following INDEX= data set option defines a simple index for the SSN variable and a composite index for CITY and STATE. Each variable must have a UNIQUE and NOMISS option:

```
data new(index=(ssn /unique/nomiss cityst=(city state)/unique/nomiss));
```

See Also

INDEX CREATE statement in "The DATASETS Procedure" in *Base SAS Procedures Guide*

CREATE INDEX statement in "The SQL Procedure" in *Base SAS Procedures Guide*

"Understanding SAS Indexes" in the "SAS Data Files" section of *SAS Language Reference: Concepts*

KEEP= Data Set Option

For an input data set, specifies the variables to process; for an output data set, specifies the variables to write to the data set.

Valid in: DATA step and PROC steps

Category: Variable Control

Syntax

KEEP=*variable-1 <...variable-n>*

Syntax Description

variable-1 <...variable-n>
lists one or more variable names. You can list the variables in any form that SAS allows.

Details

If the KEEP= data set option is associated with an input data set, only those variables that are listed after the KEEP= data set option are available for processing. If the KEEP= data set option is associated with an output data set, only the variables listed after the option are written to the output data set, but all variables are available for processing.

Comparisons

- The KEEP= data set option differs from the KEEP statement in the following ways:
  - In DATA steps, the KEEP= data set option can apply to both input and output data sets. The KEEP statement applies only to output data sets.
  - In DATA steps, when you create multiple output data sets, use the KEEP= data set option to write different variables to different data sets. The KEEP statement applies to all output data sets.
  - In PROC steps, you can use only the KEEP= data set option, not the KEEP statement.
- The DROP= data set option specifies variables to omit during processing or to omit from the output data set.

Example

In this example, only IDNUM and SALARY are read from PAYROLL, and they are the only variables in PAYROLL that are available for processing:

```
data bonus;
   set payroll(keep=idnum salary);
   bonus=salary*1.1;
run;
```

See Also

Data Set Options:

"DROP= Data Set Option" on page 22

Statements:

"KEEP Statement" on page 1648

LABEL= Data Set Option

Specifies a label for a SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

Syntax

LABEL='*label*'

Syntax Description

'label'
specifies a text string of up to 256 characters. If the label text contains single quotation marks, use double quotation marks around the label, or use two single quotation marks in the label text and surround the string with single quotation marks. To remove a label from a data set, assign a label that is equal to a blank that is enclosed in quotation marks.

Details

You can use the LABEL= option on both input and output data sets. When you use LABEL= on input data sets, it assigns a label for the file for the duration of that DATA or PROC step. When it is specified for an output data set, the label becomes a permanent part of that file and can be printed using the CONTENTS or DATASETS procedure, and modified using PROC DATASETS.

A label assigned to a data set remains associated with that data set when you update a data set in place, such as when you use the APPEND procedure or the MODIFY statement. However, a label is lost if you use a data set with a previously assigned label to create a new data set in the DATA step. For example, a label previously assigned to data set ONE is lost when you create the new output data set ONE in this DATA step:

```
data one;
     set one;
run;
```

Comparisons

- The LABEL= data set option enables you to specify labels only for data sets. You can specify labels for the variables in a data set using the LABEL statement.
- The LABEL= option in the ATTRIB statement also enables you to assign labels to variables.

Examples

These examples assign labels to data sets:

```
data w2(label='1976 W2 Info, Hourly');


data new(label='Peter''s List');


data new(label="Hillside's Daily Account");


data sales(label='Sales For May(NE)');
```

See Also

Statements:

"ATTRIB Statement" on page 1448
"LABEL Statement" on page 1650
"MODIFY Statement" on page 1684

"The CONTENTS Procedure" in *Base SAS Procedures Guide*

"The DATASETS Procedure" in *Base SAS Procedures Guide*

OBS= Data Set Option

Specifies the last observation that SAS processes in a data set.

Valid in: DATA step and PROC steps

Category: Observation Control

Default: MAX

Restriction: Use with input data sets only

Restriction: Cannot use with PROC SQL views

Syntax

OBS= *n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies a number to indicate when to stop processing observations, with *n* being an integer. Using one of the letter notations results in multiplying the integer by a specific value. That is, specifying K (kilo) multiplies the integer by 1,024, M (mega) multiplies by 1,048,576, G (giga) multiplies by 1,073,741,824, or T (tera) multiplies by 1,099,511,627,776. For example, a value of **`20`** specifies 20 observations, while a value of **`3m`** specifies 3,145,728 observations.

***hex*X**
specifies a number to indicate when to stop processing as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the hexadecimal value F8 must be specified as **`0F8x`** in order to specify the decimal equivalent of 248. The value **`2dx`** specifies the decimal equivalent of 45.

MIN
sets the number to indicate when to stop processing to 0. Use OBS=0 in order to create an empty data set that has the structure, but not the observations, of another data set.

Interaction: If OBS=0 and the NOREPLACE option is in effect, then SAS can still take certain actions because it actually executes each DATA and PROC step in the

program, using no observations. For example, SAS executes procedures, such as CONTENTS and DATASETS, that process libraries or SAS data sets.

MAX
sets the number to indicate when to stop processing to the maximum number of observations in the data set, up to the largest 8-byte, signed integer, which is 2^{63}-1, or approximately 9.2 quintillion. This is the default.

Details

OBS= tells SAS when to stop processing observations. To determine when to stop processing, SAS uses the value for OBS= in a formula that includes the value for OBS= and the value for FIRSTOBS=. The formula is

```
(obs - firstobs) + 1 = results
```

For example, if OBS=10 and FIRSTOBS=1 (which is the default for FIRSTOBS=), the result is ten observations, that is `(10 - 1) + 1 = 10`. If OBS=10 and FIRSTOBS=2, the result is nine observations, that is `(10 - 2) + 1 = 9`. OBS= is valid only when an existing SAS data set is read.

Comparisons

- The OBS= data set option overrides the OBS= system option for the individual data set.
- While the OBS= data set option specifies an ending point for processing, the FIRSTOBS= data set option specifies a starting point. The two options are often used together to define a range of observations to be processed.
- The OBS= data set option enables you to select observations from SAS data sets. You can select observations to be read from external data files by using the OBS= option in the INFILE statement.

Examples

Example 1: Using OBS= to Specify When to Stop Processing Observations This example illustrates the result of using OBS= to tell SAS when to stop processing observations. This example creates a SAS data set and executes the PRINT procedure with FIRSTOBS=2 and OBS=12. The result is 11 observations, that is `(12 - 2) + 1 = 11`. The result of OBS= in this situation appears to be the observation number that SAS processes last, because the output starts with observation 2, and ends with observation 12. This situation is only a coincidence.

```
data Ages;
   input Name $ Age;
   datalines;
Miguel 53
Brad 27
Willie 69
Marc 50
Sylvia 40
Arun 25
Gary 40
Becky 51
Alma 39
Tom 62
```

```
Kris 66
Paul 60
Randy 43
Barbara 52
Virginia 72
;
proc print data=Ages (firstobs=2 obs=12);
run;
```

Output 2.1 PROC PRINT Output Using OBS= and FIRSTOBS=

```
                          The SAS System                          1

                    Obs    Name       Age

                      2    Brad        27
                      3    Willie      69
                      4    Marc        50
                      5    Sylvia      40
                      6    Arun        25
                      7    Gary        40
                      8    Becky       51
                      9    Alma        39
                     10    Tom         62
                     11    Kris        66
                     12    Paul        60
```

Example 2: Using OBS= with WHERE Processing This example illustrates the result of using OBS= along with WHERE processing. The example uses the data set that was created in Example 1, which contains 15 observations.

First, here is the PRINT procedure with a WHERE statement. The subset of the data results in 12 observations:

```
proc print data=Ages;
   where Age LT 65;
run;
```

Output 2.2 PROC PRINT Output Using a WHERE Statement

```
                          The SAS System                          1

                    Obs    Name       Age

                      1    Miguel      53
                      2    Brad        27
                      4    Marc        50
                      5    Sylvia      40
                      6    Arun        25
                      7    Gary        40
                      8    Becky       51
                      9    Alma        39
                     10    Tom         62
                     12    Paul        60
                     13    Randy       43
                     14    Barbara     52
```

Executing the PRINT procedure with the WHERE statement and OBS=10 results in 10 observations, that is **`(10 - 1) + 1 = 10`**. Note that with WHERE processing, SAS first subsets the data and applies OBS= to the subset.

```
proc print data=Ages (obs=10);
   where Age LT 65;
run;
```

Output 2.3 PROC PRINT Output Using a WHERE Statement and OBS=

```
                          The SAS System                          2

                    Obs    Name       Age

                      1    Miguel      53
                      2    Brad        27
                      4    Marc        50
                      5    Sylvia      40
                      6    Arun        25
                      7    Gary        40
                      8    Becky       51
                      9    Alma        39
                     10    Tom         62
                     12    Paul        60
```

The result of OBS= appears to be how many observations to process, because the output consists of 10 observations, ending with the observation number 12. However, the result is only a coincidence. If you apply FIRSTOBS=2 and OBS=10 to the subset, then the result is nine observations, that is **`(10 - 2) + 1 = 9`**. OBS= in this situation is neither the observation number to end with nor how many observations to process; the value is used in the formula to determine when to stop processing.

```
proc print data=Ages (firstobs=2 obs=10);
   where Age LT 65;
run;
```

Output 2.4 PROC PRINT Output Using WHERE Statement, OBS=, and FIRSTOBS=

```
                          The SAS System                          3

                    Obs    Name       Age

                      2    Brad        27
                      4    Marc        50
                      5    Sylvia      40
                      6    Arun        25
                      7    Gary        40
                      8    Becky       51
                      9    Alma        39
                     10    Tom         62
                     12    Paul        60
```

Example 3: Using OBS= When Observations Are Deleted This example illustrates the result of using OBS= for a data set that has deleted observations. The example uses the data set that was created in Example 1, with observation 6 deleted.

First, here is PROC PRINT output of the modified file:

```
proc print data=Ages;
run;
```

Output 2.5 PROC PRINT Output Showing Observation 6 Deleted

```
                              The SAS System                          1

                        Obs    Name        Age

                          1    Miguel       53
                          2    Brad         27
                          3    Willie       69
                          4    Marc         50
                          5    Sylvia       40
                          7    Gary         40
                          8    Becky        51
                          9    Alma         39
                         10    Tom          62
                         11    Kris         66
                         12    Paul         60
                         13    Randy        43
                         14    Barbara      52
                         15    Virginia     72
```

Executing the PRINT procedure with OBS=12 results in 12 observations, that is **(12 – 1) + 1 = 12**:

```
proc print data=Ages (obs=12);
run;
```

Output 2.6 PROC PRINT Output Using OBS=

```
                              The SAS System                          2

                        Obs    Name       Age

                          1    Miguel      53
                          2    Brad        27
                          3    Willie      69
                          4    Marc        50
                          5    Sylvia      40
                          7    Gary        40
                          8    Becky       51
                          9    Alma        39
                         10    Tom         62
                         11    Kris        66
                         12    Paul        60
                         13    Randy       43
```

The result of OBS= appears to be how many observations to process, because the output consists of 12 observations, ending with the observation number 13. However, if you apply FIRSTOBS=2 and OBS=12, the result is 11 observations, that is **(12 – 2) + 1 = 11**. OBS= in this situation is neither the observation number to end with nor how many observations to process; the value is used in the formula to determine when to stop processing.

```
proc print data=Ages (firstobs=2 obs=12);
run;
```

Output 2.7 PROC PRINT Output Using OBS= and FIRSTOBS=

```
                          The SAS System                          3

                    Obs    Name      Age

                      2    Brad       27
                      3    Willie     69
                      4    Marc       50
                      5    Sylvia     40
                      7    Gary       40
                      8    Becky      51
                      9    Alma       39
                     10    Tom        62
                     11    Kris       66
                     12    Paul       60
                     13    Randy      43
```

See Also

Data Set Options:

"FIRSTOBS= Data Set Option" on page 25

Statements:

"INFILE Statement" on page 1591

"WHERE Statement" on page 1792

System Options:

"OBS= System Option" on page 1948

For more information about using OBS= with WHERE processing, see "Processing a Segment of Data That Is Conditionally Selected" in *SAS Language Reference: Concepts*.

OBSBUF= Data Set Option

Determines the size of the view buffer for processing a DATA step view.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Valid only for a DATA step view

Syntax

OBSBUF=*n*

Syntax Description

n

specifies the number of observations that are read into the view buffer at a time.

Default: 32K bytes of memory are allocated for the default view buffer, which means that the default number of observations that can be read into the view buffer at one time depends on the observation length. Therefore, the default is the number of observations that can fit into 32K bytes. If the observation length is larger than 32K, then only one observation can be read into the buffer at a time.

Tip: To determine the observation length, which is its size in bytes, use PROC CONTENTS for the DATA step view.

CAUTION:
The maximum value for the OBSBUF= option depends on the amount of available memory. If you specify a value so large that the memory allocation of the view buffer fails, an out-of- memory error results. If you specify a value that is larger than the amount of available real memory and your operating environment allows SAS to perform the allocation using virtual memory, the result can be a decrease in performance due to excessive paging. △

Details

The OBSBUF= data set option specifies the number of observations that can be read into the view buffer at a time. The *view buffer* is a segment of memory that is allocated to hold output observations that are generated from a DATA step view. The size of the buffer determines how much data can be held in memory at one time. OBSBUF= enables you to tune the performance of reading data from a DATA step view.

The view buffer is shared between the request that opens the DATA step view (for example, a SAS procedure) and the DATA step view itself. Two computer tasks coordinate between requesting data and generating and returning the data as follows:

1. When a request task, such as a PRINT procedure, requests data, task switching occurs from the request task to the view task in order to execute the DATA step view and generate the observations. The DATA step view fills the view buffer with as many observations as possible.
2. When the view buffer is full, task switching occurs from the view task back to the request task in order to return the requested data. The observations are cleared from the view buffer.

The size of the view buffer determines how many generated observations can be held. The number of generated observations then determines how many times the computer must switch between the request task and the view task. For example, OBSBUF=1 results in task switching for each observation, while OBSBUF=10 results in 10 observations being read into the view buffer at a time. The larger the view buffer is, the less task switching is needed to process a DATA step view, which can speed up execution time.

To improve efficiency, first determine how many observations fits into the default buffer size, then set the view buffer so that it can hold more generated observations.

Note: Using OBSBUF= can improve processing efficiency by reducing task switching. However, the larger the view buffer size, the more time it takes to fill. This delays the task switching from the view task back to the request task in order to return the requested data. The delay is more apparent in interactive applications. For example, when you use the Viewtable window, the larger the view buffer, the longer it takes to display the requested observations, because the view buffer must be filled before even one observation is returned to the Viewtable. Therefore, before you set a very large view buffer size, consider the type of application that you are using to process the DATA step view as well as the amount of memory that you have available. △

Example

For this example, the observation length is 10K, which means that the default view buffer size, which is 32K, would result in three observations at a time to be read into the view buffer. The default view buffer size causes the execution time to be slower, because the computer must do task switching for every three observations that are generated.

To improve performance, the OBSBUF= data set option is set to 100, which causes one hundred observations at a time to be read into the view buffer and reduces task switching in order to process the DATA step view with the PRINT procedure:

```
data testview / view=testview;
   ... more SAS statements ...
run;

proc print data=testview (obsbuf=100);
run;
```

See Also

Data Set Options:

"SPILL= Data Set Option" on page 59

OUTREP= Data Set Option

Specifies the data representation for the output SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

See: OUTREP= Data Set Option in the documentation for your operating environment.

Syntax

OUTREP=*format*

Syntax Description

format

specifies the data representation, which is the form in which data is stored in a particular operating environment. Different operating environments use different standards or conventions for storing floating-point numbers (for example, IEEE or IBM Mainframe); for character encoding (ASCII or EBCDIC); for the ordering of bytes in memory (big Endian or little Endian); for word alignment (4-byte boundaries or 8-byte boundaries); for integer data-type length (16-bit, 32-bit, or 64-bit); and for doubles (byte-swapped or not).

Native data representation refers to an environment in which the data representation is comparable to the CPU that is accessing the file. For example, a

file that is in Windows data representation is native to the Windows operating environment.

By default, SAS creates a new SAS data set by using the native data representation of the CPU that is running SAS. Specifying the OUTREP= option enables you to create a SAS data set within the native environment that uses a foreign environment data representation. For example, in a UNIX environment, you can create a SAS data set that uses Windows data representation.

Values for OUTREP= are listed in the following table:

Table 2.2 Data Representation Values for OUTREP= Option

OUTREP= Value	Alias*	Environment
ALPHA_TRU64	ALPHA_OSF	Tru64 UNIX
ALPHA_VMS_32	ALPHA_VMS	OpenVMS on Alpha
ALPHA_VMS_64		OpenVMS on Alpha
HP_IA64	HP_ITANIUM	HP-UX on Itanium 64-bit platform
HP_UX_32	HP_UX	HP-UX on 32-bit platform
HP_UX_64		HP-UX on 64-bit platform
INTEL_ABI		ABI UNIX on Intel 32-bit platform
LINUX_32	LINUX	Linux for Intel Architecture on 32-bit platform
LINUX_IA64		Linux for Itanium-based system on 64-bit platform
LINUX_X86_64		LINUX on x64 64-bit platform
MIPS_ABI		ABI UNIX on 32-bit platform
MVS_32	MVS	z/OS on 32-bit platform
OS2		OS/2 on Intel 32-bit platform
RS_6000_AIX_32	RS_6000_AIX	AIX UNIX on 32-bit RS/6000
RS_6000_AIX_64		AIX UNIX on 64-bit RS/6000
SOLARIS_32	SOLARIS	Solaris on SPARC 32-bit platform
SOLARIS_64		Solaris on SPARC 64-bit platform
SOLARIS_X86_64		Solaris on x64 64-bit platform
VAX_VMS		OpenVMS VAX
VMS_IA64		OpenVMS for HP Integrity servers 64-bit platform
WINDOWS_32	WINDOWS	Microsoft Windows on 32-bit platform
WINDOWS_64		Microsoft Windows 64-bit Edition (for both Itanium-based systems and x64)

* It is recommended that you use the current values. The aliases are available for compatibility only.

Details

CAUTION:

Transcoding could result in character data loss when encodings are incompatible. For information about encoding and transcoding, see *SAS National Language Support (NLS): Reference Guide*. △

See Also

Statements:

OUTREP= option in "LIBNAME Statement" on page 1656

"Processing Data Using Cross-Environment Data Access (CEDA)" in *SAS Language Reference: Concepts*

POINTOBS= Data Set Option

Specifies whether SAS creates compressed data sets whose observations can be randomly accessed or sequentially accessed.

Valid in: DATA step and PROC steps

Category: Observation Control

Restriction: POINTOBS= is effective only when creating a compressed data set. Otherwise it is ignored.

Syntax

POINTOBS=YES | NO

Syntax Description

YES

causes SAS software to produce a compressed data set that might be randomly accessed by observation number. This is the default.

Examples of accessing data directly by observation number are:

- the POINT= option of the MODIFY and SET statements in the DATA step
- going directly to a specific observation number with PROC FSEDIT.

Tip: Specifying POINTOBS=YES does not affect the efficiency of retrieving information from a data set, but it does increase CPU usage by approximately 10% when creating a compressed data set and when updating or adding information to it.

NO

suppresses the ability to randomly access observations in a compressed data set by observation number.

Tip: Specifying POINTOBS=NO is desirable for applications where the ability to point directly to an observation by number within a compressed data set is not important.

If you do not need to access data by observation number, then you can improve performance by approximately 10% when creating a compressed data set and when updating or adding observations to it by specifying POINTOBS=NO.

Details

Note that REUSE=YES takes precedence over POINTOBS=YES. For example:

```
data test(compress=yes pointobs=yes reuse=yes);
```

results in a data set that has POINTOBS=NO. Because POINTOBS=YES is the default when you use compression, REUSE=YES causes POINTOBS= to change to NO.

See Also

Data Set Options:

"COMPRESS= Data Set Option" on page 19

"REUSE= Data Set Option" on page 56

System Options:

"COMPRESS= System Option" on page 1872

"REUSE= System Option" on page 1983

PW= Data Set Option

Assigns a READ, WRITE, and ALTER password to a SAS file, and enables access to a password-protected SAS file.

Valid in: DATA step and PROC steps

Category: Data Set Control

See: under UNIX in the documentation for your operating environment.

Syntax

PW=*password*

Syntax Description

password

must be a valid SAS name. See "Rules for Words and Names in the SAS Language" in *SAS Language Reference: Concepts*.

Details

The PW= option applies to all types of SAS files except catalogs. You can use this option to assign a password to a SAS file or to access a password-protected SAS file.

When replacing a SAS data set that is protected by an ALTER password, the new data set inherits the ALTER password. To change the ALTER password for the new data set, use the MODIFY statement in the DATASETS procedure.

Operating Environment Information: See the appropriate sections of the SAS documentation for your operating environment for more information about using passwords. △

Note: A SAS password does not control access to a SAS file beyond the SAS system. You should use the operating system-supplied utilities and file-system security controls in order to control access to SAS files outside of SAS. △

See Also

Data Set Options:

"ALTER= Data Set Option" on page 14

"ENCRYPT= Data Set Option" on page 23

"READ= Data Set Option" on page 51

"WRITE= Data Set Option" on page 71

"Manipulating Passwords" in "The DATASETS Procedure" in *Base SAS Procedures Guide*

"File Protection" in *SAS Language Reference: Concepts*

PWREQ= Data Set Option

Specifies whether to display a dialog box to enter a SAS data set password.

Valid in: DATA and PROC steps

Category: Data Set Control

Syntax

PWREQ=YES | NO

Syntax Description

YES
specifies to display a dialog box.

NO
prevents a dialog box from displaying. If a missing or invalid password is entered, the data set is not opened and an error message is written to the SAS log.

Details

In an interactive SAS session, the PWREQ= option controls whether a dialog box displays after a user enters an incorrect or a missing password for a SAS data set that is password protected. PWREQ= applies to data sets with read, write, or alter passwords. PWREQ= is most useful in SCL applications.

See Also

Data Set Options:

"ALTER= Data Set Option" on page 14

"ENCRYPT= Data Set Option" on page 23

"PW= Data Set Option" on page 49

"READ= Data Set Option" on page 51

"WRITE= Data Set Option" on page 71

READ= Data Set Option

Assigns a READ password to a SAS file that prevents users from reading the file, unless they enter the password.

Valid in: DATA step and PROC steps

Category: Data Set Control

Syntax

READ=*read-password*

Syntax Description

read-password
must be a valid SAS name. See "Rules for Words and Names in the SAS Language" in *SAS Language Reference: Concepts*.

Details

The READ= option applies to all types of SAS files except catalogs. You can use this option to assign a password to a SAS file or to access a read-protected SAS file.

Note: A SAS password does not control access to a SAS file beyond the SAS system. You should use the operating system-supplied utilities and file-system security controls in order to control access to SAS files outside of SAS. △

See Also

Data Set Options:

"ALTER= Data Set Option" on page 14
"ENCRYPT= Data Set Option" on page 23
"PW= Data Set Option" on page 49
"WRITE= Data Set Option" on page 71

"Manipulating Passwords" in "The DATASETS Procedure" in *Base SAS Procedures Guide*

"File Protection" in *SAS Language Reference: Concepts*

RENAME= Data Set Option

Changes the name of a variable.

Valid in: DATA step and PROC steps
Category: Variable Control

Syntax

RENAME=(*old-name-1=new-name-1* < *...old-name-n=new-name-n*>)

Syntax Description

old-name
the variable you want to rename.

new-name
the new name of the variable. It must be a valid SAS name.

Details

If you use the RENAME= data set option when you create a data set, the new variable name is included in the output data set. If you use RENAME= on an input data set, the new name is used in DATA step programming statements.

If you use RENAME= on an input data set that is used in a SAS procedure, SAS changes the name of the variable in that procedure. If you use RENAME= with WHERE processing such as a WHERE statement or a WHERE= data set option, the new name is applied before the data is processed. You must use the new name in the WHERE expression.

If you use RENAME= in the same DATA step with either the DROP= or the KEEP= data set option, the DROP= and the KEEP= data set options are applied before RENAME=. You must use the old name in the DROP= and KEEP= data set options. You cannot drop and rename the same variable in the same statement.

Note: The RENAME= data set option has an effect only on data sets that are opened in output mode. △

Comparisons

- The RENAME= data set option differs from the RENAME statement in the following ways:
  - The RENAME= data set option can be used in PROC steps and the RENAME statement cannot.
  - The RENAME statement applies to all output data sets. If you want to rename different variables in different data sets, you must use the RENAME= data set option.
  - To rename variables before processing begins, you must use a RENAME= data set option on the input data set or data sets.
- Use the RENAME statement or the RENAME= data set option when program logic requires that you rename variables such as two input data sets that have variables with the same name. To rename variables as a file management task, use the DATASETS procedure.

Examples

Example 1: Renaming a Variable at Time of Output This example uses RENAME= in the DATA statement to show that the variable is renamed at the time it is written to the output data set. The variable keeps its original name, X, during the DATA step processing:

```
data two(rename=(x=keys));
   set one;
   z=x+y;
run;
```

Example 2: Renaming a Variable at Time of Input This example renames variable X to a variable named KEYS in the SET statement, which is a rename before DATA step processing. KEYS, not X, is the name to use for the variable for DATA step processing.

```
data three;
   set one(rename=(x=keys));
   z=keys+y;
run;
```

Example 3: Renaming a Variable for a SAS Procedure with WHERE Processing This example renames variable Score1 to a variable named Score2 for the PRINT procedure. Because the new name is applied before the data is processed, the new name must be specified in the WHERE statement.

```
proc print data=test (rename=(score1=score2));
   where score2 gt 75;
run;
```

See Also

Data Set Options:

"DROP= Data Set Option" on page 22

"KEEP= Data Set Option" on page 36

Statements:

"RENAME Statement" on page 1743

"The DATASETS Procedure" in *Base SAS Procedures Guide*

REPEMPTY= Data Set Option

Specifies whether a new, empty data set can overwrite an existing SAS data set that has the same name.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with output data sets only.

Syntax

REPEMPTY=YES | NO

Syntax Description

YES
specifies that a new empty data set with a given name replaces an existing data set with the same name. This is the default.

Interaction: When REPEMPTY=YES and REPLACE=NO, then the data set is not replaced.

NO
specifies that a new empty data set with a given name does not replace an existing data set with the same name.

Tip: Use REPEMPTY=NO to prevent the following syntax error from replacing the existing data set B with the new empty data set B that is created by mistake:

```
data mylib.a set b;
```

Tip: For both the convenience of replacing existing data sets with new ones that contain data and the protection of not overwriting existing data sets with new empty ones that are created by accident, set REPLACE=YES and REPEMPTY=NO.

Comparisons

- For an individual data set, the REPEMPTY= data set option overrides the REPEMPTY= option in the LIBNAME statement.
- The REPEMPTY= and REPLACE= data set options apply to both permanent and temporary SAS data sets. The REPLACE system option, however, only applies to permanent SAS data sets.

See Also

Data Set Options:

"REPLACE= Data Set Option" on page 55

Statement Options:

REPEMPTY= in the LIBNAME statement on page 1660

System Options:

"REPLACE System Option" on page 1982

REPLACE= Data Set Option

Specifies whether a new SAS data set that contains data can overwrite an existing data set that has the same name.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with output data sets only.

Restriction: This option is valid only when creating a SAS data set.

Syntax

REPLACE=NO | YES

Syntax Description

NO
specifies that a new data set with a given name does not replace an existing data set with the same name.

YES
specifies that a new data set with a given name replaces an existing data set with the same name.

Comparisons

- The REPLACE= data set option overrides the REPLACE system option for the individual data set.
- The REPLACE system option only applies to permanent SAS data sets.

Example

Using the REPLACE= data set option in this DATA statement prevents SAS from replacing a permanent SAS data set named ONE in a library referenced by MYLIB:

```
data mylib.one(replace=no);
```

SAS writes a message in the log that tells you that the file has not been replaced.

See Also

System Options:

"REPLACE System Option" on page 1982

REUSE= Data Set Option

Specifies whether new observations can be written to freed space in compressed SAS data sets.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Use with output data sets only.

Syntax

REUSE=NO | YES

Syntax Description

NO
does not track and reuse space in compressed data sets. New observations are appended to the existing data set. Specifying the NO argument results in less efficient data storage if you delete or update many observations in the SAS data set.

YES
tracks and reuses space in compressed SAS data sets. New observations are inserted in the space that is freed when other observations are updated or deleted.

If you plan to use procedures that add observations to the end of SAS data sets (for example, the APPEND and FSEDIT procedures) with compressed data sets, use the REUSE=NO argument. REUSE=YES causes new observations to be added wherever there is space in the file, not necessarily at the end of the file.

Details

By default, new observations are appended to existing compressed data sets. If you want to track and reuse free space by deleting or updating other observations, use the REUSE= data set option when you create a compressed SAS data set.

REUSE= has meaning only when you are creating new data sets with the COMPRESS=YES data set option or system option. Using the REUSE= data set option when you are accessing an existing SAS data set has no effect.

Comparisons

The REUSE= data set option overrides the REUSE= system option.

REUSE=YES takes precedence over POINTOBS=YES. For example, the following statement results in a data set that has POINTOBS=NO:

```
data test(compress=yes pointobs=yes reuse=yes);
```

Because POINTOBS=YES is the default when you use compression, REUSE=YES causes POINTOBS= to change to NO.

See Also

Data Set Options:

"COMPRESS= Data Set Option" on page 19

System Options:

"REUSE= System Option" on page 1983

SORTEDBY= Data Set Option

Specifies how a data set is currently sorted.

Valid in: DATA step and PROC steps

Category: Data Set Control

Syntax

SORTEDBY=*by-clause* </ *collate-name*> | _NULL_

Syntax Description

***by-clause* < / *collate-name*>**
indicates how the data is currently sorted.

by-clause names the variables and options that you use in a BY statement in a PROC SORT step.

collate-name names the collating sequence that is used for the sort. By default, the collating sequence is that of your operating environment. A slash (/) must precede the collating sequence.

Operating Environment Information: For details about collating sequences, see the SAS documentation for your operating environment. △

NULL
removes any existing sort indicator.

Details

SAS determines whether a data set is already sorted by the key variable or variables in ascending order by checking the sort indicator. The sort indicator is stored in the data set descriptor information and is set from a previous sort. For detailed information about how the sort indicator is used and how it improves performance, see "The Sort Indicator" in *SAS Language Reference: Concepts* and the "SORTVALIDATE= System Option" in the *SAS Language Reference: Dictionary*.

The following example of the CONTENTS procedure **`Sort Information`** section containing the **`Validated`** attribute set to NO, indicates that the data set was sorted using the SORTEDBY= data set option.

```
Sort Information
 Sortedby var1
 Validated NO
 Character Set ANSI
```

Comparisons

- Use the CONTENTS statement in the DATASETS procedure to see how a data set is sorted.
- The SORTEDBY= option indicates how the data is sorted, but does not cause a data set to be sorted.

Examples

This example uses the SORTEDBY= data set option to specify how the data are currently sorted. The data set ORDERS is sorted by PRIORITY and by the descending values of INDATE. Once the data set is created, the sort indicator is stored with it. These statements create the data set ORDERS and record the sort indicator:

```
libname mylib 'SAS-library';
options yearcutoff=1920;

data mylib.orders(sortedby=priority
                  descending indate);
   input priority 1. +1 indate date7.
         +1 office $ code $;
   format indate date7.;
   datalines;
1 03may01 CH J8U
1 21mar01 LA M91
1 01dec00 FW L6R
1 27feb99 FW Q2A
2 15jan00 FW I9U
2 09jul99 CH P3Q
3 08apr99 CH H5T
3 31jan99 FW D2W
;
```

See Also

The CONTENTS statement in "The DATASETS Procedure" in *Base SAS Procedures Guide*

"The SORT Procedure" in *Base SAS Procedures Guide*

"The SQL Procedure" in *Base SAS Procedures Guide*

SPILL= Data Set Option

Specifies whether to create a spill file for non-sequential processing of a DATA step view.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: Valid only for a DATA step view

Syntax

SPILL=YES | NO

Syntax Description

YES
creates a spill file for non-sequential processing of a DATA step view. This is the default.

Interaction: A spill file is never created for sequential processing of a DATA step view.

Tip: A DATA step view that generates large amounts of observations can result in a very large spill file. You must have enough disk space to accommodate the spill file.

NO
does not create a spill file or reduces the size of a spill file.

Interaction: For direct (random) access, a spill file is always created even if you specify SPILL=NO.

Tip: If you do not have enough disk space to accommodate a resulting spill file from a DATA step view that generates a large amount of data, specify SPILL=NO.

Tip: For SAS procedures that process BY-group data, consider specifying SPILL=NO in order to write only the current BY group to the spill file.

Details

When a DATA step view is opened for non-sequential processing, a spill file is created by default. The *spill file* contains the observations that are generated by a DATA step view. Subsequent requests for data read the observations from the spill file rather than execute the DATA step view again. The spill file is a temporary file in the WORK library.

Non-sequential processing includes the following access methods, which are supported by several SAS statements and procedures. How the SPILL= data set option operates with each of the access methods is described below:

random access	retrieves observations directly either by an observation number or by the value of one or more variables through an index without reading all observations sequentially. Whether SPILL=YES or SPILL=NO, a spill file is always created, because the processing time to restart a DATA step view for each observation would be costly.
BY-group access	uses a BY statement to process observations that are ordered, grouped, or indexed according to the values of one or more variables. SPILL=YES creates a spill file the size of all the data that is

requested from the DATA step view. SPILL=NO writes only the current BY group to the spill file. The largest size of the spill file is a size to store the largest BY group.

two-pass access performs multiple sequential passes through the data. With SPILL=NO, no spill file is created. Instead, after the first pass through the data, the DATA step view is restarted for each subsequent pass through the data. If small amounts of data are returned by the DATA step view for each restart, the processing time to restart the view might become significant.

Note: With SPILL=NO, subsequent passes through the data could result in generating different data. Some processing might require using a spill file; for example, results from using random functions and computing values that are based on the current time of day could affect the data. △

Examples

Example 1: Using a Spill File for a Small Number of Large BY Groups This example creates a DATA step view that generates a large amount of random data and uses the UNIVARIATE procedure with a BY statement. The example illustrates the effects of SPILL= with a small number of large BY groups.

With SPILL=YES, all observations that are requested from the DATA step view are written to the spill file. With SPILL=NO, only the observations that are in the current BY group are written to the spill file. The information messages that are produced by this example show that the size of the spill file is reduced with SPILL=NO. However, the time to truncate the spill file for each BY group might add to the overall processing time for the DATA step view.

```
options msglevel=i;

data vw_few_large / view=vw_few_large;
   drop i;

   do byval = 'Group A', 'Group B', 'Group C';
      do i = 1 to 500000;
         r = ranuni(4);
         output;
      end;
   end;
run;

proc univariate data=vw_few_large (spill=yes) noprint;
   var r;
   by byval;
run;

proc univariate data=vw_few_large (spill=no) noprint;
   var r;
   by byval;
run;
```

Output 2.8 SAS Log Output

```
1    options msglevel=i;
2    data vw_few_large / view=vw_few_large;
3       drop i;
4
5       do byval = 'Group A', 'Group B', 'Group C';
6          do i = 1 to 500000;
7             r = ranuni(4);
8             output;
9          end;
10      end;
11   run;

NOTE: DATA STEP view saved on file WORK.VW_FEW_LARGE.
NOTE: A stored DATA STEP view cannot run under a different operating system.
NOTE: DATA statement used (Total process time):
      real time           21.57 seconds
      cpu time            1.31 seconds


12   proc univariate data=vw_few_large (spill=yes) noprint;
INFO: View WORK.VW_FEW_LARGE open mode: BY-group rewind.
13      var r;
14      by byval;
15   run;

INFO: View WORK.VW_FEW_LARGE opening spill file for output observations.
INFO: View WORK.VW_FEW_LARGE deleting spill file.  File size was 22506120 bytes.
NOTE: View WORK.VW_FEW_LARGE.VIEW used (Total process time):
      real time           40.68 seconds
      cpu time            12.71 seconds

NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           57.63 seconds
      cpu time            13.12 seconds


16
17   proc univariate data=vw_few_large (spill=no) noprint;
INFO: View WORK.VW_FEW_LARGE open mode: BY-group rewind.
18      var r;
19      by byval;
20   run;

INFO: View WORK.VW_FEW_LARGE opening spill file for output observations.
INFO: View WORK.VW_FEW_LARGE truncating spill file.  File size was 7502040 bytes.
NOTE: The above message was for the following by-group:
      byval=Group A
INFO: View WORK.VW_FEW_LARGE truncating spill file.  File size was 7534800 bytes.
NOTE: The above message was for the following by-group:
      byval=Group B
INFO: View WORK.VW_FEW_LARGE truncating spill file.  File size was 7534800 bytes.
NOTE: The above message was for the following by-group:
      byval=Group C
INFO: View WORK.VW_FEW_LARGE deleting spill file.  File size was 32760 bytes.
NOTE: View WORK.VW_FEW_LARGE.VIEW used (Total process time):
      real time           11.03 seconds
      cpu time            10.95 seconds

NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           11.04 seconds
      cpu time            10.96 seconds
```

Example 2: Using a Spill File for a Large Number of Small BY Groups This example creates a DATA step view that generates a large amount of random data and uses the UNIVARIATE procedure with a BY statement. This example illustrates the effects of SPILL= with a large number of small BY groups.

With SPILL=YES, all observations that are requested from the DATA step view are written to the spill file. With SPILL=NO, only the observations that are in the current BY group are written to the spill file. The information messages that are produced by this example show that the size of the spill file is reduced with SPILL=NO, and with small BY groups, this results in a large disk space savings.

```
options msglevel=i;
data vw_many_small / view=vw_many_small;
   drop i;

   do byval = 1 to 100000;
      do i = 1 to 5;
         r = ranuni(4);
         output;
      end;
   end;
run;

proc univariate data=vw_many_small (spill=yes) noprint;
   var r;
   by byval;
run;

proc univariate data=vw_many_small (spill=no) noprint;
   var r;
   by byval;
run;
```

Output 2.9 SAS Log Output

```
1    options msglevel=i;
2    data vw_many_small / view=vw_many_small;
3       drop i;
4
5       do byval = 1 to 100000;
6          do i = 1 to 5;
7             r = ranuni(4);
8             output;
9          end;
10      end;
11   run;

NOTE: DATA STEP view saved on file WORK.VW_MANY_SMALL.
NOTE: A stored DATA STEP view cannot run under a different operating system.
NOTE: DATA statement used (Total process time):
      real time           0.56 seconds
      cpu time            0.03 seconds


12   proc univariate data=vw_many_small (spill=yes) noprint;
INFO: View WORK.VW_MANY_SMALL open mode: BY-group rewind.
13      var r;
14      by byval;
15   run;

INFO: View WORK.VW_MANY_SMALL opening spill file for output observations.
INFO: View WORK.VW_MANY_SMALL deleting spill file.  File size was 8024240 bytes.
NOTE: View WORK.VW_MANY_SMALL.VIEW used (Total process time):
      real time           30.73 seconds
      cpu time            29.59 seconds

NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           30.96 seconds
      cpu time            29.68 seconds


16
17   proc univariate data=vw_many_small (spill=no) noprint;
INFO: View WORK.VW_MANY_SMALL open mode: BY-group rewind.
18      var r;
19      by byval;
20   run;

INFO: View WORK.VW_MANY_SMALL opening spill file for output observations.
INFO: View WORK.VW_MANY_SMALL truncating spill file.  File size was 65504 bytes.
NOTE: The above message was for the following by-group:
      byval=410
INFO: View WORK.VW_MANY_SMALL truncating spill file.  File size was 65504 bytes.
NOTE: The above message was for the following by-group:
      byval=819
INFO: View WORK.VW_MANY_SMALL truncating spill file.  File size was 65504 bytes.
NOTE: The above message was for the following by-group:
      byval=1229
   .
   . Deleted many INFO and NOTE messages for BY groups
   .
INFO: View WORK.VW_MANY_SMALL truncating spill file.  File size was 65504 bytes.
NOTE: The above message was for the following by-group:
      byval=99894
INFO: View WORK.VW_MANY_SMALL deleting spill file.  File size was 32752 bytes.
NOTE: View WORK.VW_MANY_SMALL.VIEW used (Total process time):
      real time           29.43 seconds
      cpu time            28.81 seconds

NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           29.43 seconds
      cpu time            28.81 seconds
```

Example 3: Using a Spill File with Two-Pass Access This example creates a DATA step view that generates a large amount of random data and uses the TRANSPOSE procedure. The example illustrates the effects of SPILL= with a procedure that requires two-pass access processing.

When PROC TRANSPOSE processes a DATA step view, the procedure must make two passes through the observations that the view generates. The first pass counts the number of observations and the second pass performs the transposition. With SPILL=YES, a spill file is created during the first pass, and the second pass reads the observations from the spill file. With SPILL=NO, a spill file is not created—after the first pass, the DATA step view is restarted.

Note that for the first TRANSPOSE procedure, which does not include the SPILL= data set option, even though a spill file is used by default, the informative message about the open mode is not displayed. This action occurs to reduce the amount of messages in the SAS log for users who are not using the SPILL= data set option.

```
options msglevel=i;
data vw_transpose/view=vw_transpose;
   drop i j;
   array x[10000];
   do i = 1 to 10;
      do j = 1 to dim(x);
         x[j] = ranuni(4);
      end;
      output;
   end;
run;
proc transpose data=vw_transpose out=transposed;
run;
proc transpose data=vw_transpose(spill=yes) out=transposed;
run;
proc transpose data=vw_transpose(spill=no) out=transposed;
run;
```

Output 2.10 SAS Log Output

```
1    options msglevel=i;
2    data vw_transpose/view=vw_transpose;
3       drop i j;
4       array x[10000];
5       do i = 1 to 10;
6          do j = 1 to dim(x);
7             x[j] = ranuni(4);
8          end;
9         output;
10      end;
11   run;

NOTE: DATA STEP view saved on file WORK.VW_TRANSPOSE.
NOTE: A stored DATA STEP view cannot run under a different operating system.
NOTE: DATA statement used (Total process time):
      real time           0.68 seconds
      cpu time            0.18 seconds


12   proc transpose data=vw_transpose out=transposed;
13   run;

INFO: View WORK.VW_TRANSPOSE opening spill file for output observations.
INFO: View WORK.VW_TRANSPOSE deleting spill file.  File size was 880000 bytes.
NOTE: View WORK.VW_TRANSPOSE.VIEW used (Total process time):
      real time           2.37 seconds
      cpu time            1.17 seconds

NOTE: There were 10 observations read from the data set WORK.VW_TRANSPOSE.
NOTE: The data set WORK.TRANSPOSED has 10000 observations and 11 variables.
NOTE: PROCEDURE TRANSPOSE used (Total process time):
      real time           4.17 seconds
      cpu time            1.51 seconds


14   proc transpose data=vw_transpose (spill=yes) out=transposed;
INFO: View WORK.VW_TRANSPOSE open mode: sequential.
15   run;

INFO: View WORK.VW_TRANSPOSE reopen mode: two-pass.
INFO: View WORK.VW_TRANSPOSE opening spill file for output observations.
INFO: View WORK.VW_TRANSPOSE deleting spill file.  File size was 880000 bytes.
NOTE: View WORK.VW_TRANSPOSE.VIEW used (Total process time):
      real time           0.95 seconds
      cpu time            0.92 seconds

NOTE: There were 10 observations read from the data set WORK.VW_TRANSPOSE.
NOTE: The data set WORK.TRANSPOSED has 10000 observations and 11 variables.
NOTE: PROCEDURE TRANSPOSE used (Total process time):
      real time           1.01 seconds
      cpu time            0.98 seconds


16   proc transpose data=vw_transpose (spill=no) out=transposed;
INFO: View WORK.VW_TRANSPOSE open mode: sequential.
17   run;

INFO: View WORK.VW_TRANSPOSE reopen mode: two-pass.
INFO: View WORK.VW_TRANSPOSE restarting for another pass through the data.
NOTE: View WORK.VW_TRANSPOSE.VIEW used (Total process time):
      real time           1.34 seconds
      cpu time            1.32 seconds

NOTE: The View WORK.VW_TRANSPOSE was restarted 1 times. The following view statistics
      only apply to the last view restart.
NOTE: There were 10 observations read from the data set WORK.VW_TRANSPOSE.
NOTE: The data set WORK.TRANSPOSED has 10000 observations and 11 variables.
NOTE: PROCEDURE TRANSPOSE used (Total process time):
      real time           1.42 seconds
      cpu time            1.40 seconds
```

See Also

Data Set Options:

"OBSBUF= Data Set Option" on page 44

TOBSNO= Data Set Option

Specifies the number of observations to send in a client/server transfer.

Valid in: DATA step and PROC steps

Category: Data Set Control

Restriction: The TOBSNO= option is valid only for data sets that are accessed through a SAS server via the REMOTE engine.

Syntax

TOBSNO=*n*

Syntax Description

n
specifies the number of observations to be transmitted.

Details

If the TOBSNO= option is not specified, its value is calculated based on the observation length and the size of the server's transmission buffers, as specified by the PROC SERVER statement TBUFSIZE= option.

The TOBSNO= option is valid only for data sets that are accessed through a SAS server via the REMOTE engine. If this option is specified for a data set opened for update or accessed via another engine, it is ignored.

See Also

"FOPEN Function" in *SAS Component Language: Reference*.

TYPE= Data Set Option

Specifies the data set type for a specially structured SAS data set.

Valid in: DATA step and PROC steps

Category: Data Set Control

Syntax

TYPE=*data-set-type*

Syntax Description

data-set-type
specifies the special type of the data set.

Details

Use the TYPE= data set option in a DATA step to create a special SAS data set in the proper format, or to identify the special type of the SAS data set in a procedure statement.

You can use the CONTENTS procedure to determine the type of a data set.

Most SAS data sets do not have a specified type. However, there are several specially structured SAS data sets that are used by some SAS/STAT procedures. These SAS data sets contain special variables and observations, and they are usually created by SAS statistical procedures. Because most of the special SAS data sets are used with SAS/STAT software, they are described in the *SAS/STAT User's Guide*. Some of the special data sets are CORR, COV, SSPC, EST, or FACTOR.

Other values are available in other SAS software products and are described in the appropriate documentation.

Note: If you use a DATA step with a SET statement to modify a special SAS data set, you must specify the TYPE= option in the DATA statement. The *data-set-type* is not automatically copied to the data set that is created. △

See Also

"Special SAS Data Sets" in the *SAS/STAT User's Guide*

"The CONTENTS Procedure" in the *Base SAS Procedures Guide*

WHERE= Data Set Option

Specifies specific conditions to use to select observations from a SAS data set.

Valid in: DATA step and PROC steps

Category: Observation Control

Restriction: Cannot be used with the POINT= option in the SET and MODIFY statements.

Syntax

WHERE=(*where-expression-1<logical-operator where-expression-n>*)

Syntax Description

where-expression
is an arithmetic or logical expression that consists of a sequence of operators, operands, and SAS functions. An operand is a variable, a SAS function, or a constant. An operator is a symbol that requests a comparison, logical operation, or arithmetic calculation. The expression must be enclosed in parentheses.

logical-operator
can be AND, AND NOT, OR, or OR NOT.

Details

- Use the WHERE= data set option with an input data set to select observations that meet the condition specified in the WHERE expression before SAS brings them into the DATA or PROC step for processing. Selecting observations that meet the conditions of the WHERE expression is the first operation SAS performs in each iteration of the DATA step.

  You can also select observations that are written to an output data set. In general, selecting observations at the point of input is more efficient than selecting them at the point of output. However, there are some cases when selecting observations at the point of input is not practical or not possible.
- You can apply OBS= and FIRSTOBS= processing to WHERE processing. For more information see "Processing a Segment of Data That is Conditionally Selected" in *SAS Language Reference: Concepts*.
- You cannot use the WHERE= data set option with the POINT= option in the SET and MODIFY statements.
- If you use both the WHERE= data set option and the WHERE statement in the same DATA step, SAS ignores the WHERE statement for data sets with the WHERE= data set option. However, you can use the WHERE= data set option with the WHERE command in SAS/FSP software.

Note: Using indexed SAS data sets can improve performance significantly when you are using WHERE expressions to access a subset of the observations in a SAS data set. See "Understanding SAS Indexes" in *SAS Language Reference: Concepts* for a complete discussion of WHERE expression processing with indexed data sets and a list of guidelines to consider before indexing your SAS data sets. △

Comparisons

- The WHERE statement applies to all input data sets, whereas the WHERE= data set option selects observations only from the data set for which it is specified.
- Do not confuse the purpose of the WHERE= data set option. The DROP= and KEEP= data set options select variables for processing, while the WHERE= data set option selects observations.

Examples

Example 1: Selecting Observations from an Input Data Set This example uses the WHERE= data set option to subset the SALES data set as it is read into another data set:

```
data whizmo;
   set sales(where=(product='whizmo'));
run;
```

Example 2: Selecting Observations from an Output Data Set This example uses the WHERE= data set option to subset the SALES output data set:

```
data whizmo(where=(product='whizmo'));
   set sales;
run;
```

See Also

Statements:

"WHERE Statement" on page 1792

"WHERE-Expression Processing" in *SAS Language Reference: Concepts*

WHEREUP= Data Set Option

Specifies whether to evaluate new observations and modified observations against a WHERE expression.

Valid in: DATA step and PROC steps

Category: Observation Control

Syntax

WHEREUP=NO | YES

Syntax Description

NO
does not evaluate added observations and modified observations against a WHERE expression.

YES
evaluates added observations and modified observations against a WHERE expression.

Details

Specify WHEREUP=YES when you want any added observations or modified observations to match a specified WHERE expression.

Examples

Example 1: Accepting Updates That Do Not Match the WHERE Expression This example shows how WHEREUP= permits observations to be updated and added even though the modified observation does not match the WHERE expression:

```
data a;
   x=1;
   output;
   x=2;
   output;
run;

data a;
   modify a(where=(x=1) whereup=no);
   x=3;
   replace; /* Update does not match WHERE expression */
   output; /* Add does not match WHERE expression */
run;
```

In this example, SAS updates the observation and adds the new observation to the data set.

Example 2: Rejecting Updates That Do Not Match the WHERE Expression In this example, WHEREUP= does not permit observations to be updated or added when the update and the add do not match the WHERE expression:

```
data a;
   x=1;
   output;
   x=2;
   output;
run;

data a;
   modify a(where=(x=1) whereup=yes);
   x=3;
   replace; /* Update does not match WHERE expression */
   output; /* Add does not match WHERE expression */
run;
```

In this example, SAS does not update the observation nor does it add the new observation to the data set.

See Also

Data Set Option:

"WHERE= Data Set Option" on page 67

WRITE= Data Set Option

Assigns a WRITE password to a SAS file that prevents users from writing to a file, unless they enter the password.

Valid in: DATA step and PROC steps

Category: Data Set Control

Syntax

WRITE=*write-password*

Syntax Description

write-password
must be a valid SAS name.

See: "Rules for Words and Names in the SAS Language" in *SAS Language Reference: Concepts*

Details

The WRITE= option applies to all types of SAS files except catalogs. You can use this option to assign a password to a SAS file or to access a write-protected SAS file.

Note: A SAS password does not control access to a SAS file beyond the SAS system. You should use the operating system-supplied utilities and file-system security controls in order to control access to SAS files outside of SAS. △

See Also

Data Set Options:

"ALTER= Data Set Option" on page 14
"ENCRYPT= Data Set Option" on page 23
"PW= Data Set Option" on page 49
"READ= Data Set Option" on page 51

"Manipulating Passwords" in "The DATASETS Procedure" in *Base SAS Procedures Guide*

Data Set Options Documented in Other SAS Publications

In addition to data set options documented in *SAS Language Reference: Dictionary*, data set options are also documented in the following publications:

"*SAS Companion for Windows*" on page 72
"*SAS Companion for OpenVMS on HP Integrity Servers*" on page 72
"*SAS Companion for UNIX Environments*" on page 72

SAS Companion for Windows

Data Set Option	Description
SGIO=	Activates the Scatter/Gather I/O feature for a dataset.

SAS Companion for OpenVMS on HP Integrity Servers

The data set options listed here are documented only in *SAS Companion for OpenVMS on HP Integrity Servers*. Other data set options in *SAS Companion for OpenVMS on HP Integrity Servers* contain information specific to the OpenVMS operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter data set options are not listed here.

Data Set Option	**Description**
ALQ=	Specifies how many disk blocks to initially allocate to a new SAS data set.
ALQMULT=	Specifies the number of pages that are preallocated to a file.
BKS=	Specifies the bucket size for a new data set.
CACHENUM=	Specifies the number of I/O data caches used per SAS file.
CACHESIZE=	Controls the size of the I/O data cache that is allocated for a SAS file.
DEQ=	Specifies how many disk blocks to add when OpenVMS automatically extends a SAS data set during a write operation.
DEQMULT=	Specifies the number of pages to extend a SAS file.
LOCKREAD	Specifies whether to read a record if a lock cannot be obtained for the record.
LOCKWAIT	Indicates whether SAS should wait for a locked record.
MBF	Specifies the multibuffer count for a data set.

SAS Companion for UNIX Environments

The data set options listed here are documented only in *SAS Companion for UNIX Environments*. Other data set options in *SAS Companion for UNIX Environments* contain information specific to the UNIX operating environment, where the main

documentation is in *SAS Language Reference: Dictionary*. These latter data set options are not listed here.

Data Set Option	Description
ALTER=	Specifies a password for a SAS file that prevents users from replacing or deleting the file, but permits read and write access.
BUFNO=	Specifies the number of buffers to be allocated for processing a SAS data set.
BUFSIZE=	Specifies the size of a permanent buffer page for an output SAS data set.
FILECLOSE=	Specifies how a tape is positioned when a SAS data set is closed.
PW=	Assigns a READ, WRITE, or ALTER password to a SAS file, and enables access to a password-protected SAS file.
USEDIRECTIO	Turns on direct I/O for a library that contains the file to which the ENABLEDIRECTIO option has been applied.

SAS Companion for z/OS

The data set options listed here are documented only in *SAS Companion for z/OS*. Other data set options in *SAS Companion for z/OS* contain information specific to the z/OS operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter data set options are not listed here.

Data Set Option	Description
ALTER=	Assigns an alter password to a SAS file and enables access to a password-protected SAS file.
BUFSIZE=	Specifies the permanent buffer page size for an output SAS data set.
FILEDISP=	Specifies the initial disposition for a sequential access bound SAS data library.

SAS National Language Support: Reference Guide

Data Set Option	Description
ENCODING=	Overrides the encoding to use for reading or writing a SAS data set.

SAS Scalable Performance Data Engine: Reference

Data Set Option	Description
ASYNCINDEX=	Specifies to create the indexes in parallel when creating multiple indexes on an SPD Engine data set.
BYNOEQUALS=	Specifies whether the output order of data set observations with identical values for the BY variable are guaranteed to be in data set order.
BYSORT=	Specifies for the SPD Engine to perform an automatic sort when it encounters a BY statement.
COMPRESS=	Specifies to compress SPD Engine data sets on disk as they are being created.
ENCRYPT=	Specifies whether to encrypt an output SPD Engine data set.
ENDOBS=	Specifies the end observation number in a user-defined range of observations to be processed.
IDXWHERE=	Specifies to use indexes when processing WHERE expressions in the SPD Engine.
IOBLOCKSIZE=	Specifies the number of observations in a block to be stored in or read from an SPD Engine data component file that is compressed.
LISTFILES=	Specifies whether the CONTENTS procedure lists the complete pathnames of all the component files.
PADCOMPRESS=	Specifies a number of bytes to add to compression blocks in a data set opened for UPDATE.
PARTSIZE=	When an SPD Engine data set is created, specifies the largest size (in megabytes) that the data component partitions can be. This is a fixed size. This specification applies only to the data component files.
STARTOBS=	Specifies the starting observation number in a user-defined range of observations to be processed.
SYNCADD=	Specifies to process one observation at a time or multiple observations at a time.
THREADNUM=	Specifies the number of I/O threads the SPD Engine can spawn for processing an SPD Engine data set.
UNIQUESAVE=	Specifies to save observations with non-unique key values (the rejected observations) to a separate data set when appending or inserting observations to data sets with unique indexes.
WHERENOINDEX=	Specifies, when making WHERE expression evaluations, a list of indexes to exclude.

SAS/ACCESS for Relational Databases: References

Data Set Option	Description
AUTHID=	Enables you to qualify the specified table with an authorization ID, user ID, or group ID.
AUTOCOMMIT=	Specifies whether to enable the DBMS autocommit capability.
BL_ALLOW_READ_ACCESS=	Specifies that the original table data is still visible to readers during bulk load.
BL_ALLOWWRITE_ACCESS=	Specifies that table data is still accessible to readers and writers while import is in progress.
BL_BADDATA_FILE=	Specifies where to put records that failed to process internally.
BL_BADFILE=	Identifies a file that contains records that were rejected during a bulk load.
BL_CODEPAGE=	Identifies the codepage that the DBMS engine uses to convert SAS character data to the current database codepage during a bulk load.
BL_CONTROL=	Identifies a file containing SQLLDR control statements that describe the data to be included in a bulk load.
BL_COPY_LOCATION=	Specifies the directory to which DB2 saves a copy of the loaded data. This option is valid only when used in conjunction with BL_RECOVERABLE=YES.
BL_CPU_PARALLELISM=	Specifies the number of processes or threads that are used when building table objects.
BL_DATA_BUFFER_SIZE=	Specifies the total amount of memory that is allocated for the bulk load utility to use as a buffer for transferring data.
BL_DATAFILE=	Identifies the file that contains the data that is loaded or appended into a DBMS table during a bulk load.
BL_DB2CURSOR=	Specifies a string that contains a valid DB2 SELECT statement that points to either local or remote objects (tables or views).
BL_DB2DEVT_PERM=	Specifies the unit address or generic device type that is used for the permanent data sets created by the LOAD utility, as well as SYSIN, SYSREC, and SYSPRINT when they are allocated by SAS.
BL_DB2DEVT_TEMP=	Specifies the unit address or generic device type that is used for the temporary data sets created by the LOAD utility (PNCH, COPY1, COPY2, RCPY1, RCPY2, WORK1, WORK2).
BL_DB2DISC=	Specifies the SYSDISC data set name for the LOAD utility.
BL_DB2ERR=	Specifies the SYSERR data set name for the LOAD utility.
BL_DB2IN=	Specifies the SYSIN data set name for the LOAD utility
BL_DB2LDCT1=	Specifies a string in the LOAD utility control statement, between LOAD DATA and INTO TABLE.
BL_DB2LDCT2=	Specifies a string in the LOAD utility control statement, between INTO TABLE table-name and (field-specification).
BL_DB2LDCT3=	Specifies a string in the LOAD utility control statement, after (*field-specification*)

Data Set Option	Description
BL_DB2LDEXT=	Specifies the mode of execution for the DB2 LOAD utility.
BL_DB2MAP=	Specifies the SYSMAP data set name for the LOAD utility.
BL_DB2PRINT=	Specifies the SYSPRINT data set name for the LOAD utility.
BL_DB2PRNLOG=	Determines whether the SYSPRINT output is written to the SAS log.
BL_DB2REC=	Specifies the SYSREC data set name for the LOAD utility
BL_DB2RECSP=	Determines the number of cylinders to specify as the primary allocation for the SYSREC data set when it is created.
BL_DB2RSTRT=	Tells the LOAD utility whether the current load is a restart and, for a restart, indicates where to begin.
BL_DB2SPC_PERM=	Determines the number of cylinders to specify as the primary allocation for the permanent data sets that are created by the LOAD utility.
BL_DB2SPC_TEMP=	Determines the number of cylinders to specify as the primary allocation for the temporary data sets that are created by the LOAD utility.
BL_DB2TBLXST=	Indicates whether the LOAD utility runs against an existing table
BL_DB2UTID=	Specifies a unique identifier for a given run of the DB2 LOAD utility.
BL_DELETE_DATAFILE=	Deletes the data file that is created for the DBMS bulk load facility.
BL_DELIMITER=	Specifies override of the default delimiter character for separating columns of data during data transfer or retrieval during bulk load or bulk unload.
BL_DIRECT_PATH=	Sets the Oracle SQL*Loader DIRECT option.
BL_DISCARDFILE=	Identifies the file that contains the records that were filtered out of a bulk load because they did not match the criteria specified in the CONTROL file.
BL_DISCARDS=	"Specifies whether and when to stop processing a job, based on the number of discarded records.
BL_DISK_PARALLELISM=	Specifies the number of processes or threads that are used when writing data to disk.
BL_ERRORS=	Specifies whether and when to stop processing a job based on the number of failed records.
BL_EXCEPTION=	Specifies the exception table into which rows in error are copied.
BL_FAILEDDATA=	Specifies where to put records that could not be written to the database.
BL_INDEX_OPTIONS=	Enables you to specify SQL*Loader Index options with bulk loading.
BL_INDEXING_MODE=	Used to indicate which scheme the DB2 load utility should use with respect to index maintenance.
BL_KEEPIDENTITY=	Determines whether the identity column that is created during a bulk load is populated with values generated by Microsoft SQL Server or with values provided by the user.
BL_KEEPNULLS=	Indicates how NULL values in Microsoft SQL Server columns that accept NULL are handled during a bulk load.

Data Set Option	Description
BL_LOAD_METHOD=	Specifies the method by which data is loaded into an Oracle table during bulk loading.
BL_LOAD_REPLACE=	Specifies whether DB2 appends or replaces rows during bulk loading
BL_LOG=	Identifies a log file that contains information such as statistics and error information for a bulk load.
BL_METHOD=	Specifies which bulk loading method to use for DB2.
BL_OPTIONS=	Passes options to the DBMS bulk load facility, affecting how it loads and processes data.
BL_PARFILE=	Creates a file that contains the SQL*Loader command line options.
BL_PORT_MAX=	Sets the highest available port number for concurrent uploads.
BL_PORT_MIN=	Sets the lowest available port number for concurrent uploads.
BL_PRESERVE_BLANKS=	Determines how the SQL*Loader handles requests to insert blank spaces into CHAR/VARCHAR2 columns with the NOT NULL constraint.
BL_RECOVERABLE=	Determines whether the LOAD process is recoverable.
BL_REMOTE_FILE=	Specifies the base filename and location of DB2 LOAD temporary files.
BL_RETRIES=	Specifies the number of attempts to make for a job.
BL_RETURN_ WARNINGS_AS_ ERRORS=	Specifies whether SQL*Loader (bulkload) warnings should surface in SAS through the SYSERR macro warnings or as errors.
BL_SERVER_DATAFILE=	Specifies the name and location of the data file as seen by the DB2 server instance.
BL_SQLLDR_PATH=	Specifies the location of the SQLLDR executable file.
BL_SUPPRESS_NULLIF=	Indicates whether to suppress the NULLIF clause for the specified columns when a table is created in order to increase performance.
BL_USE_PIPE=	Specifies a named pipe for data transfer.
BL_WARNING_COUNT=	Specifies the maximum number of row warnings to allow before you abort the load operation.
BUFFERS=	Specifies the number of shared memory buffers to be used for transferring data from SAS to Teradata.
BULK_BUFFER=	Specifies the number of bulk rows that the SAS/ACCESS engine can buffer for output.
BULKLOAD=	Loads rows of data as one unit.
BULKUNLOAD	Rapidly retrieves (fetches) large number of rows from a data set.
CAST=	Specifies whether data conversions should be performed on the Teradata DBMS server or by SAS.
CAST_OVERHEAD_ MAXPERCENT=	Specifies the overhead limit for data conversions to be performed in Teradata instead of SAS.
COMMAND_TIMEOUT=	Specifies the number of seconds to wait before a command times out.
CURSOR_TYPE=	Specifies the cursor type for read only and updatable cursors.

Data Set Option	Description
DBCOMMIT=	Causes an automatic COMMIT (a permanent writing of data to the DBMS) after a specified number of rows have been processed.
DBCONDITION=	Specifies criteria for subsetting and ordering DBMS data.
DBCREATE_TABLE_OPTS=	Specifies DBMS-specific syntax to be added to the CREATE TABLE statement.
DBFORCE=	Specifies whether to force the truncation of data during insert processing.
DBGEN_NAME=	Specifies how SAS renames columns automatically when they contain characters that SAS does not allow.
DBINDEX=	Detects and verifies that indexes exist on a DBMS table. If they do exist and are of the correct type, a join query that is passed to the DBMS might improve performance.
DBKEY=	Specifies a key column to optimize DBMS retrieval. Can improve performance when you are processing a join that involves a large DBMS table and a small SAS data set or DBMS table.
DBLABEL=	Specifies whether to use SAS variable labels or SAS variable names as the DBMS column names during output processing.
DBLINK=	Specifies a link from your default database to another database on the server to which you are connected in the Sybase interface; and specifies a link from your local database to database objects on another server in the Oracle interface.
DBMASTER=	Designates which table is the larger table when you are processing a join that involves tables from two different types of databases.
DBMAX_TEXT=	Determines the length of any very long DBMS character data type that is read into SAS or written from SAS when you are using a SAS/ACCESS engine.
DBNULL=	Indicates whether NULL is a valid value for the specified columns when a table is created.
DBNULLKEYS=	Controls the format of the WHERE clause with regard to NULL values when you use the DBKEY= data set option.
DBPROMPT=	Specifies whether SAS displays a window that prompts you to enter DBMS connection information.
DBSASLABEL=	Specifies how the engine returns column labels.
DBSASTYPE=	Specifies data types to override the default SAS data types during input processing.
DBSLICE=	Specifies user-supplied WHERE clauses to partition a DBMS query for threaded reads.
DBSLICEPARM=	Controls the scope of DBMS threaded reads and the number of DBMS connections.
DBTYPE=	Specifies a data type to use instead of the default DBMS data type when SAS creates a DBMS table.
DEGREE=	Determines whether DB2 uses parallelism.
DISTRIBUTE_ON	Specifies a column name to use in the DISTRIBUTE ON clause of the CREATE TABLE statement.

Data Set Option	Description
ERRLIMIT=	Specifies the number of errors that are allowed before SAS stops processing and issues a rollback.
ESCAPE_BACKSLASH=	Specifies whether backslashes in literals are preserved during data copy from a SAS data set to a table.
IGNORE_ READ_ONLY_COLUMNS=	Specifies whether to ignore or include columns whose data types are read-only when generating an SQL statement for inserts or updates.
IN=	Enables you to specify the database or tablespace in which you want to create a new table.
INSERT_SQL=	Determines the method that is used to insert rows into a data source.
INSERTBUFF=	Specifies the number of rows in a single DBMS insert.
KEYSET_SIZE=	Specifies the number of rows in the cursor that are key set driven.
LOCATION=	Enables you to further specify exactly where a table resides.
LOCKTABLE=	Places exclusive or shared locks on tables.
MBUFFSIZE=	Specifies the size of the shared memory buffers to be used for transferring data from SAS to Teradata.
ML_CHECKPOINT=	Specifies the interval between checkpoint operation, in minutes.
ML_ERROR1=	Specifies the name of a temporary table that MultiLoad uses to track errors that were generated during the acquisition phase of a bulk-load operation.
ML_ERROR2=	Specifies the name of a temporary table that MultiLoad uses to track errors that were generated during the application phase of a bulk-load operation.
ML_LOG=	Specifies a prefix for the names of the temporary tables that MultiLoad uses during a bulk-load operation.
ML_RESTART=	Specifies the name of a temporary table that is used by MultiLoad to track checkpoint information.
ML_WORK=	Specifies the name of a temporary table that MultiLoad uses to store intermediate data.
MULTILOAD=	Specifies whether Teradata insert and append operations should use the Teradata MultiLoad utility.
MULTISTMT=	Specifies whether insert statements are to be sent to Teradata one at a time or in a group.
NULLCHAR=	Indicates how missing SAS character values are handled during insert, update, DBINDEX=, and DBKEY= processing.
NULLCHARVAL=	Defines the character string that replaces missing SAS character values during insert, update, DBINDEX=, and DBKEY= processing.
OR_PARTITION=	Allows reading, updating, and deleting from a particular partition in a partitioned table, also inserting and bulk-loading into a particular partition in a partitioned table.
OR_UPD_NOWHERE=	Specifies whether SAS uses an extra WHERE clause when updating rows with no lockingSpecifies whether SAS uses an extra WHERE clause when updating rows with no locking.

Data Set Option	Description
ORHINTS=	Specifies Oracle hints to pass to Oracle from a SAS statement or SQL procedure.
PRESERVE_COL_NAMES=	Preserves spaces, special characters, and case-sensitivity in DBMS column names when you create DBMS tables.
QUALIFIER=	Specifies the qualifier to use when you are reading database objects, such as DBMS tables and views.
QUERY_TIMEOUT=	Specifies the number of seconds of inactivity to wait before canceling a query.
READ_ISOLATION_LEVEL=	Specifies which level of read isolation locking to use when you are reading data.
READ_LOCK_TYPE=	Specifies how data in a DBMS table is locked during a read transaction.
READ_MODE_WAIT=	Specifies during SAS/ACCESS read operations whether Teradata waits to acquire a lock or fails your request when the DBMS resource is locked by a different user.
READBUFF=	Specifies the number of rows of DBMS data to read into the buffer.
SASDATEFMT=	Changes the SAS date format of a DBMS column.
SCHEMA=	Enables you to read a data source, such as a DBMS table and view, in the specified schema.
SEGMENT_NAME=	Enables you to control the segment in which you create a table.
SET=	Specifies whether duplicate rows are allowed when creating a table.
SLEEP=	Specifies the number of minutes that MultiLoad waits before it retries logging in to Teradata.
TENACITY=	Specifies how many hours MultiLoad continues to retry logging on to Teradata if the maximum number of Teradata utilities are already running.
TRAP151=	Enables columns that cannot be updated to be removed from a FOR UPDATE OF clause so updating of columns can proceed as normal.
UPDATE_ISOLATION_ LEVEL=	Defines the degree of isolation of the current application process from other concurrently running application processes.
UPDATE_LOCK_TYPE=	Specifies how data in a DBMS table is locked during an update transaction.
UPDATE_MODE_WAIT=	Specifies during SAS/ACCESS update operations whether the DBMS waits to acquire a lock or fails your request when the DBMS resource is locked by a different user.
UPDATE_SQL=	Determines the method that is used to update and delete rows in a data source.
UPDATEBUFF=	Specifies the number of rows that are processed in a single DBMS update or delete operation.

CHAPTER

3

Formats

Definition of Formats

A *format* is an instruction that SAS uses to write data values. You use formats to control the written appearance of data values, or, in some cases, to group data values together for analysis. For example, the WORDS22. format, which converts numeric values to their equivalent in words, writes the numeric value 692 as **`six hundred ninety-two`**.

Syntax

SAS formats have the following form:

<$>*format*<*w*>.<*d*>

where

$
indicates a character format; its absence indicates a numeric format.

format
names the format. The format is a SAS format or a user-defined format that was previously defined with the VALUE statement in PROC FORMAT. For more information about user-defined formats, see “The FORMAT Procedure” in *Base SAS Procedures Guide*.

w
specifies the format width, which for most formats is the number of columns in the output data.

d
specifies an optional decimal scaling factor in the numeric formats.

Formats always contain a period (.) as a part of the name. If you omit the *w* and the *d* values from the format, SAS uses default values. The *d* value that you specify with a format tells SAS to display that many decimal places. Formats never change or truncate the internally stored data values.

For example, in DOLLAR10.2, the *w* value of 10 specifies a maximum of 10 columns for the value. The *d* value of 2 specifies that two of these columns are for the decimal part of the value, which leaves eight columns for all the remaining characters in the value. The remaining columns include the decimal point, the remaining numeric value, a minus sign if the value is negative, the dollar sign, and commas, if any.

If the format width is too narrow to represent a value, SAS tries to squeeze the value into the space available. Character formats truncate values on the right. Numeric formats sometimes revert to the BEST*w.d* format. SAS prints asterisks if you do not specify an adequate width. In the following example, the result is x=**.

```
x=123;
put x= 2.;
```

If you use an incompatible format, such as using a numeric format to write character values, SAS first attempts to use an analogous format of the other type. If this attempt fails, an error message that describes the problem appears in the SAS log.

When the value of *d* is greater than fifteen, the precision of the decimal value after the 15th decimal place might not be accurate.

Using Formats

Ways to Specify Formats

You can use formats in the following ways:

- in a PUT statement
- with the PUT, PUTC, or PUTN functions
- with the %SYSFUNC macro function
- in a FORMAT statement in a DATA step or a PROC step
- in an ATTRIB statement in a DATA step or a PROC step.

PUT Statement

The PUT statement with a format after the variable name uses a format to write data values in a DATA step. For example, this PUT statement uses the DOLLAR*w.d* format to write the numeric value for AMOUNT as a dollar amount:

```
amount=1145.32;
put amount dollar10.2;
```

The DOLLAR*w.d* format in the PUT statement produces this result:

```
$1,145.32
```

See “PUT Statement” on page 1708 for more information.

PUT Function

The PUT function converts a numeric variable, a character variable, or a constant using any valid format and returns the resulting character value. For example, the following statement converts the value of a numeric variable into a two-character hexadecimal representation:

```
num=15;
char=put(num,hex2.);
```

The PUT function returns a value of 0F, which is assigned to the variable CHAR.

The PUT function is useful for converting a numeric value to a character value. See “PUT Function” on page 1056 for more information.

%SYSFUNC

The %SYSFUNC (or %QSYSFUNC) macro function executes SAS functions or user-defined functions and applies an optional format to the result of the function outside a DATA step. For example, the following program writes a numeric value in a macro variable as a dollar amount.

```
%macro tst(amount);
  %put %sysfunc(putn(&amount,dollar10.2));
%mend tst;

%tst (1154.23);
```

For more information, see *SAS Macro Language: Reference*.

FORMAT Statement

The FORMAT statement permanently associates a format with a variable. SAS uses the format to write the values of the variable that you specify. For example, the following statement in a DATA step associates the COMMA*w.d* numeric format with the variables SALES1 through SALES3:

```
format sales1-sales3 comma10.2;
```

Because the FORMAT statement permanently associates a format with a variable, any subsequent DATA step or PROC step uses COMMA10.2 to write the values of SALES1, SALES2, and SALES3. See “FORMAT Statement” on page 1576 for more information.

Note: If you assign formats with a FORMAT statement before a PUT statement, all leading blanks are trimmed. Formats that are associated with variables by using a FORMAT statement behave like formats that are used with a colon (:) modifier in a subsequent PUT statement. For details about using the colon format modifier, see “PUT Statement, List” on page 1731. △

ATTRIB Statement

The ATTRIB statement can also associate a format, as well as other attributes, with one or more variables. For example, in the following statement the ATTRIB statement permanently associates the COMMA*w.d* format with the variables SALES1 through SALES3:

```
attrib sales1-sales3 format=comma10.2;
```

Because the ATTRIB statement permanently associates a format with a variable, any subsequent DATA step or PROC step uses COMMA10.2 to write the values of SALES1, SALES2, and SALES3. See "ATTRIB Statement" on page 1448 for more information.

Permanent versus Temporary Association

When you specify a format in a PUT statement, SAS uses the format to write data values during the DATA step but does not permanently associate the format with a variable. To permanently associate a format with a variable, use a FORMAT statement or an ATTRIB statement in a DATA step. SAS permanently associates a format with the variable by modifying the descriptor information in the SAS data set.

Using a FORMAT statement or an ATTRIB statement in a PROC step associates a format with a variable for that PROC step, as well as for any output data sets that the procedure creates that contain formatted variables. For more information about using formats in SAS procedures, see *Base SAS Procedures Guide*.

User-Defined Formats

In addition to the formats that are supplied with Base SAS software, you can create your own formats. In Base SAS software, PROC FORMAT allows you to create your own formats for both character and numeric variables. For more information, see "The FORMAT Procedure" in *Base SAS Procedures Guide*.

When you execute a SAS program that uses user-defined formats, these formats should be available. The two ways to make these formats available are

- to create permanent, not temporary, formats with PROC FORMAT
- to store the source code that creates the formats (the PROC FORMAT step) with the SAS program that uses them.

To create permanent SAS formats, see "The FORMAT Procedure" in *Base SAS Procedures Guide*.

If you execute a program that cannot locate a user-defined format, the result depends on the setting of the FMTERR system option. If the user-defined format is not found, then these system options produce these results:

System Options	Results
FMTERR	SAS produces an error that causes the current DATA or PROC step to stop.
NOFMTERR	SAS continues processing and substitutes a default format, usually the BEST*w*. or $*w*. format.

Although using NOFMTERR enables SAS to process a variable, you lose the information that the user-defined format supplies.

To avoid problems, make sure that your program has access to all user-defined formats that are used.

Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms

Definitions

Integer values for binary integer data are typically stored in one of three sizes: one-byte, two-byte, or four-byte. The ordering of the bytes for the integer varies depending on the platform (operating environment) on which the integers were produced.

The ordering of bytes differs between the "big endian" and "little endian" platforms. These colloquial terms are used to describe byte ordering for IBM mainframes (big endian) and for Intel-based platforms (little endian). In the SAS System, the following platforms are considered big endian: AIX, HP-UX, IBM mainframe, Macintosh, and Solaris on SPARC. The following platforms are considered little endian: Intel ABI, Linux, OpenVMS Alpha, OpenVMS Integrity, Solaris on x64, Tru64 UNIX, and Windows.

How Bytes are Ordered Differently

On big endian platforms, the value 1 is stored in binary and is represented here in hexadecimal notation. One byte is stored as 01, two bytes as 00 01, and four bytes as 00 00 00 01. On little endian platforms, the value 1 is stored in one byte as 01 (the same as big endian), in two bytes as 01 00, and in four bytes as 01 00 00 00.

If an integer is negative, the "two's complement" representation is used. The high-order bit of the most significant byte of the integer will be set on. For example, –2 would be represented in one, two, and four bytes on big endian platforms as FE, FF FE, and FF FF FF FE respectively. On little endian platforms, the representation would be FE, FE FF, and FE FF FF FF. These representations result from the output of the integer binary value –2 expressed in hexadecimal representation.

Writing Data Generated on Big Endian or Little Endian Platforms

SAS can read signed and unsigned integers regardless of whether they were generated on a big endian or a little endian system. Likewise, SAS can write signed and unsigned integers in both big endian and little endian format. The length of these integers can be up to eight bytes.

The following table shows which format to use for various combinations of platforms. In the Signed Integer column, "no" indicates that the number is unsigned and cannot be negative. "Yes" indicates that the number can be either negative or positive.

Table 3.1 SAS Formats and Byte Ordering

Platform For Which the Data Was Created	Platform That Writes the Data	Signed Integer	Format
big endian	big endian	yes	IB or S370FIB
big endian	big endian	no	PIB, S370FPIB, S370FIBU
big endian	little endian	yes	S370FIB

Platform For Which the Data Was Created	Platform That Writes the Data	Signed Integer	Format
big endian	little endian	no	S370FPIB
little endian	big endian	yes	IBR
little endian	big endian	no	PIBR
little endian	little endian	yes	IB or IBR
little endian	little endian	no	PIB or PIBR
big endian	either	yes	S370FIB
big endian	either	no	S370FPIB
little endian	either	yes	IBR
little endian	either	no	PIBR

Integer Binary Notation and Different Programming Languages

The following table compares integer binary notation according to programming language.

Table 3.2 Integer Binary Notation and Programming Languages

Language	2 Bytes	4 Bytes
SAS	IB2. , IBR2., PIB2., PIBR2., S370FIB2., S370FIBU2., S370FPIB2.	IB4., IBR4., PIB4., PIBR4., S370FIB4., S370FIBU4., S370FPIB4.
PL/I	FIXED BIN(15)	FIXED BIN(31)
Fortran	INTEGER*2	INTEGER*4
COBOL	COMP PIC 9(4)	COMP PIC 9(8)
IBM assembler	H	F
C	short	long

Data Conversions and Encodings

An encoding maps each character in a character set to a unique numeric representation, resulting in a table of all code points. A single character can have different numeric representations in different encodings. For example, the ASCII encoding for the dollar symbol $ is 24 hexadecimal. The Danish EBCDIC encoding for the dollar symbol $ is 67 hexadecimal. In order for a version of SAS that normally uses ASCII to properly interpret a data set that is encoded in Danish EBCDIC, the data must be transcoded.

Transcoding is the process of moving data from one encoding to another. When transcoding the ASCII dollar sign to the Danish EBCDIC dollar sign, the hexadecimal representation for the character is converted from the value 24 to a 67.

If you want to know the encoding of a particular SAS data set, for SAS 9 and above follow these steps:

1. Locate the data set with SAS Explorer.
2. Right-click the data set.
3. Select Properties from the menu.
4. Click the Details tab.
5. The encoding of the data set is listed, along with other information.

Some situations where data might commonly be transcoded are:

- when you share data between two different SAS sessions that are running in different locales or in different operating environments,
- when you perform text-string operations, such as converting to uppercase or lowercase,
- when you display or print characters from another language,
- when you copy and paste data between SAS sessions running in different locales.

For more information about SAS features designed to handle data conversions from different encodings or operating environments, see *SAS National Language Support (NLS): Reference Guide*.

Working with Packed Decimal and Zoned Decimal Data

Definitions

Packed decimal	specifies a method of encoding decimal numbers by using each byte to represent two decimal digits. Packed decimal representation stores decimal data with exact precision. The fractional part of the number is determined by the informat or format because there is no separate mantissa and exponent. An advantage of using packed decimal data is that exact precision can be maintained. However, computations involving decimal data might become inexact due to the lack of native instructions.
Zoned decimal	specifies a method of encoding decimal numbers in which each digit requires one byte of storage. The last byte contains the number's sign as well as the last digit. Zoned decimal data produces a printable representation.
Nibble	specifies 1/2 of a byte.

Types of Data

Packed Decimal Data

A packed decimal representation stores decimal digits in each "nibble" of a byte. Each byte has two nibbles, and each nibble is indicated by a hexadecimal character. For example, the value 15 is stored in two nibbles, using the hexadecimal characters 1 and 5.

The sign indication is dependent on your operating environment. On IBM mainframes, the sign is indicated by the last nibble. With formats, C indicates a positive value, and D indicates a negative value. With informats, A, C, E, and F indicate positive values, and B and D indicate negative values. Any other nibble is invalid for signed packed decimal data. In all other operating environments, the sign is indicated in its own byte. If the high-order bit is 1, then the number is negative. Otherwise, it is positive.

The following applies to packed decimal data representation:

- You can use the S370FPD format on all platforms to obtain the IBM mainframe configuration.
- You can have unsigned packed data with no sign indicator. The packed decimal format and informat handles the representation. It is consistent between ASCII and EBCDIC platforms.
- Note that the S370FPDU format and informat expects to have an F in the last nibble, while packed decimal expects no sign nibble.

Zoned Decimal Data

The following applies to zoned decimal data representation:

- A zoned decimal representation stores a decimal digit in the low order nibble of each byte. For all but the byte containing the sign, the high-order nibble is the numeric zone nibble (F on EBCDIC and 3 on ASCII).
- The sign can be merged into a byte with a digit, or it can be separate, depending on the representation. But the standard zoned decimal format and informat expects the sign to be merged into the last byte.
- The EBCDIC and ASCII zoned decimal formats produce the same printable representation of numbers. There are two nibbles per byte, each indicated by a hexadecimal character. For example, the value 15 is stored in two bytes. The first byte contains the hexadecimal value F1 and the second byte contains the hexadecimal value C5.

Packed Julian Dates

The following applies to packed Julian dates:

- The two formats and informats that handle Julian dates in packed decimal representation are PDJULI and PDJULG. PDJULI uses the IBM mainframe year computation, while PDJULG uses the Gregorian computation.
- The IBM mainframe computation considers 1900 to be the base year, and the year values in the data indicate the offset from 1900. For example, 98 means 1998, 100 means 2000, and 102 means 2002. 1998 would mean 3898.
- The Gregorian computation allows for 2-digit or 4-digit years. If you use 2-digit years, SAS uses the setting of the YEARCUTOFF= system option to determine the true year.

Platforms Supporting Packed Decimal and Zoned Decimal Data

Some platforms have native instructions to support packed and zoned decimal data, while others must use software to emulate the computations. For example, the IBM mainframe has an Add Pack instruction to add packed decimal data, but the Intel-based platforms have no such instruction and must convert the decimal data into some other format.

Languages Supporting Packed Decimal and Zoned Decimal Data

Several languages support packed decimal and zoned decimal data. The following table shows how COBOL picture clauses correspond to SAS formats and informats.

IBM VS COBOL II clauses	Corresponding S370Fxxx formats/informats
PIC S9(X) PACKED-DECIMAL	S370FPDw.
PIC 9(X) PACKED-DECIMAL	S370FPDUw.
PIC S9(W) DISPLAY	S370FZDw.
PIC 9(W) DISPLAY	S370FZDUw.
PIC S9(W) DISPLAY SIGN LEADING	S370FZDLw.
PIC S9(W) DISPLAY SIGN LEADING SEPARATE	S370FZDSw.
PIC S9(W) DISPLAY SIGN TRAILING SEPARATE	S370FZDTw.

For the packed decimal representation listed above, X indicates the number of digits represented, and W is the number of bytes. For PIC S9(X) PACKED-DECIMAL, W is **`ceil((x+1)/2)`**. For PIC 9(X) PACKED-DECIMAL, W is **`ceil(x/2)`**. For example, PIC S9(5) PACKED-DECIMAL represents five digits. If a sign is included, six nibbles are needed. **`ceil((5+1)/2)`** has a length of three bytes, and the value of W is 3.

Note that you can substitute COMP-3 for PACKED-DECIMAL.

In IBM assembly language, the P directive indicates packed decimal, and the Z directive indicates zoned decimal. The following shows an excerpt from an assembly language listing, showing the offset, the value, and the DC statement:

```
offset  value (in hex)      inst label     directive

+000000 00001C                2 PEX1       DC PL3'1'
+000003 00001D                3 PEX2       DC PL3'-1'
+000006 F0F0C1                4 ZEX1       DC ZL3'1'
+000009 F0F0D1                5 ZEX2       DC ZL3'1'
```

In PL/I, the FIXED DECIMAL attribute is used in conjunction with packed decimal data. You must use the PICTURE specification to represent zoned decimal data. There is no standardized representation of decimal data for the Fortran or the C languages.

Summary of Packed Decimal and Zoned Decimal Formats and Informats

SAS uses a group of formats and informats to handle packed and zoned decimal data. The following table lists the type of data representation for these formats and informats. Note that the formats and informats that begin with S370 refer to IBM mainframe representation.

Format	Type of data representation	Corresponding informat	Comments
PD	Packed decimal	PD	Local signed packed decimal
PK	Packed decimal	PK	Unsigned packed decimal; not specific to your operating environment
ZD	Zoned decimal	ZD	Local zoned decimal
none	Zoned decimal	ZDB	Translates EBCDIC blank (hexadecimal 40) to EBCDIC zero (hexadecimal F0); corresponds to the informat as zoned decimal
none	Zoned decimal	ZDV	Non-IBM zoned decimal representation
S370FPD	Packed decimal	S370FPD	Last nibble C (positive) or D (negative)
S370FPDU	Packed decimal	S370FPDU	Last nibble always F (positive)
S370FZD	Zoned decimal	S370FZD	Last byte contains sign in upper nibble: C (positive) or D (negative)
S370FZDU	Zoned decimal	S370FZDU	Unsigned; sign nibble always F
S370FZDL	Zoned decimal	S370FZDL	Sign nibble in first byte in informat; separate leading sign byte of hexadecimal C0 (positive) or D0 (negative) in format
S370FZDS	Zoned decimal	S370FZDS	Leading sign of - (hexadecimal 60) or + (hexadecimal 4E)
S370FZDT	Zoned decimal	S370FZDT	Trailing sign of - (hexadecimal 60) or + (hexadecimal 4E)
PDJULI	Packed decimal	PDJULI	Julian date in packed representation - IBM computation
PDJULG	Packed decimal	PDJULG	Julian date in packed representation - Gregorian computation
none	Packed decimal	RMFDUR	Input layout is: *mmssttt*F
none	Packed decimal	SHRSTAMP	Input layout is: *yyyyddd*F*hhmmssth*, where *yyyyddd*F is the packed Julian date; *yyyy* is a 0-based year from 1900

Format	Type of data representation	Corresponding informat	Comments
none	Packed decimal	SMFSTAMP	Input layout is: *xxxxxxxxyyyydddF*, where *yyyydddF* is the packed Julian date; *yyyy* is a 0-based year from 1900
none	Packed decimal	PDTIME	Input layout is: 0*hhmmss*F
none	Packed decimal	RMFSTAMP	Input layout is: 0*hhmmss*F*yyyydddF*, where *yyyydddF* is the packed Julian date; yyyy is a 0-based year from 1900

Working with Dates and Times Using the ISO 8601 Basic and Extended Notations

ISO 8601 Formatting Symbols

The following list explains the formatting symbols that are used to notate the ISO 8601 dates, time, datetime, durations, and interval values:

n specifies a number that represents the number of years, months, or days

P indicates that the duration that follows is specified by the number of years, months, days, hours, minutes, and seconds

T indicates that a time value follows. Any value with a time must begin with T.

Requirement: Time values that are read by the extended notation informats that begin with the characters E8601 must use an uppercase T.

W indicates that the duration is specified in weeks.

Z indicates that the time value is the time in Greenwich, England, or UTC time.

+|- the + indicates the time zone offset to the east of Greenwich, England. The - indicates the time zone offset to the west of Greenwich, England.

yyyy specifies a four-digit year

mm as part of a date, specifies a two-digit month, 01 - 12

dd specifies a two-digit day, 01 - 31

hh specifies a two-digit hour, 00 - 24

mm as part of a time, specifies a two-digit minute, 00 - 59

ss	specifies a two-digit second, 00 - 59
fff | *ffffff*	specifies an optional fraction of a second using the digits 0 - 9:
	fff use 1 - 3 digits for values read by the $N8601B informat and the $N8601E informat
	ffffff use 1 - 6 digits for informat other than the $N8601B informat and the $N8601E informat
Y	indicates that a year value proceeds this character in a duration
M	as part of a date, indicates that a month value proceeds this character in a duration
D	indicates that a day value proceeds this character in a duration
H	indicates that an hour value proceeds this character in a duration
M	as part of a time, indicates that a minute value proceeds this character in a duration
S	indicates that a seconds value proceeds this character in a duration

Writing ISO 8601 Date, Time, and Datetime Values

SAS uses the formats in the following table to write date, time, and datetime values in the ISO 8601 basic and extended notations from SAS date, time, and datetime values.

Table 3.3 Formats for Writing ISO 8601 Dates, Times, and Datetimes

Date, Time, or Datetime	ISO 8601 Notation	Example	Format
Basic Notations			
Date	*yyyymmdd*	20080915	B8601DA*w*.
Time	*hhmmssffffff*	155300322348	B8601TM*w.d*
Time with time zone	*hhmmss*+|-*hhmm*	155300+0500	B8601TZ*w.d*
	*hhmmss*Z	155300Z	B8601TZ*w.d*
Convert to local time with time zone	hhmmss+|-hhmm	155300+0500	B8601LZ*w.d*
Datetime	*yyyymmdd*T*hhmmssffffff*	20080915T155300	B8601DT*w.d*
Datetime with timezone	*yyyymmdd*T*hhmmss*+|-*hhmm*	20080915T155300+0500	B8601DZ*w.d*
	*yyyymmdd*T*hhmmss*Z	20080915T155300Z	B8601DZ*w.d*
Write the date from a datetime	*yyyymmdd*	20080915	B8601DN*w*.
Extended Notations			
Date	*yyyy-mm-dd*	2008-09-15	E8601DA*w*.
Time	*hh:mm:ss.ffffff*	15:53:00.322348	E8601TM*w.d*
Time with time zone	*hh:mm:ss.ffffff*+|-*hh:mm*	15:53:00+05:00	E8601TZ*w.d*
Convert to local time with time zone	*hh:mm:ss.ffffff*+|-*hh:mm*	15:53:00+05:00	E8601LZ*w.d*

Date, Time, or Datetime	ISO 8601 Notation	Example	Format
Datetime	*yyyy-mm-dd*T*hh:mm:ss.ffffff*	2008-09-15T15:53:00	E8601DT*w.d*
Datetime with time zone	*yyyy-mm-dd*T*hh:mm:ss.nnnnnn*+|-*hh:mm*	2008-09-15T15:53:00+05:00	E8601DZ*w.d*
Write the date from a datetime	*yyyy-mm-dd*	2008-09-15	E8601DN*w*.

An asterisk (*)used in place of a date or time formatted value that is out-of-range.

Writing ISO 8601 Duration, Datetime, and Interval Values

Duration, Datetime, and Interval Formats

SAS writes duration, datetime, and interval values from character data using these formats:

Time Component	ISO 8601 Notation	Example	Format
Duration - Basic Notation	P*yyyymmdd*T*hhmmssfff*	P20080915T155300	$N8601BA
	-P*yyyymmdd*T*hhmmssfff*	-P20080915T155300	$N8601BA
Duration - Extended Notation	P*yyyy-mm-dd*T*hh:mm:ss.fff*	P2008-09-15T15:53:00	$N8601EA
	-P*yyyy-mm-dd*T*hh:mm:ss.fff*	-P2008-09-15T15:53:00	$N8601EA
Duration - Basic and Extended Notation	P*n*Y*n*M*n*DT*n*H*n*M*n*S	P2y10m14dT20h13m45s	$N8601B $N8601E
	-P*n*Y*n*M*n*DT*n*H*n*M*n*S	-P2y10m14dT20h13m45s	$N8601B $N8601E
	P*n*W (weeks)	P6w	$N8601B $N8601E
Interval - Basic Notation	*yyyymmdd*T*hhmmssfff*/ yyyymmddT*hhmmssfff*	20080915T155300/ 20101113T000000	$N8601BA
	P*n*Y*n*M*n*DT*n*H*n*M*n*S/ *yyyymmdd*T*hhmmssfff*	P2y10M14dT20h13m45s/ 20080915T155300	$N8601B
	*yyyymmdd*T*hhmmssfff*/ P*n*Y*n*M*n*DT*n*H*n*M*n*S	20080915T155300/ P2y10M14dT20h13m45s	$N8601BA
Interval- Extended Notation	*yyyy-mm-dd*T*hh:mm:ss.fff*/ *yyyy-mm-dd*T*hh:mm:ss.fff*	2008-09-15T15:53:00/ 2010-11-13T00:00:00	$N8601EA
	P*n*Y*n*M*n*DT*n*H*n*M*n*S/ *yyyy-mm-dd*T*hh:mm:ss.fff*	P2y10M14dT20h13m45s/ 2008-09-15T15:53:00	$N8601E

Time Component	ISO 8601 Notation	Example	Format
	*yyyy-mm-dd*T*hh:mm:ss.fff*/ P*n*Y*n*M*n*DT*n*H*n*M*n*S	2008-09-15T15:53:00/ P2y10M14dT20h13m45s	$N8601EA
Datetime-Basic Notation	*yyyymmdd*T*hhmmss.fff*+|-*hhmm*	20080915T155300	$N8601BA
	(all blank)		$N8601B
			$N8601BA
			$N8601E
			$N8601EA
Datetime-Extended Notation	*yyyy-mm-dd*T*hh:mm:ss.fff*+|-*hhmm*	2008-09-15T15:53:00 +04:30	$N8601EA
	(all blank)		$N8601B
			$N8601BA
			$N8601E
			$N8601EA

Writing Omitted Components

An omitted component can be represented by a hyphen (-) or an x in the extended datetime form *yyyy-mm-dd*T*hh:mm:ss* and in the extended duration form P*yyyy-mm-dd*T*hh:mm:ss*.

Omitted components in the durations form P*n*Y*n*M*n*DT*n*H*n*M*n*S are dropped, they do not contain a hyphen or x. For example, P2mT4H.

The following formats write omitted components that use the hyphen and the x:

Format	Datetime Form	Duration Form	Examples
$N8601H	*yyyy-mm-dd*T*hh:mm:ss*	P*n*Y*n*M*n*DT*n*H*n*M*n*S	–09-15T15:-:53 P2Y2DT4H5M6S/ –09-15T15:-:00
$N8601EH	*yyyy-mm-dd*T*hh:mm:ss*	P*yyyy-mm-dd*T*hh:mm:ss*	P000—02T02:55:20/ 2008—15T-:-:45
$N8601X	*yyyy-mm-dd*T*hh:mm:ss*	P*n*Y*n*M*n*DT*n*H*n*M*n*S	P2Y2DT4H5M6S/ x-09-15T15:x:00
$N8601EX	*yyyy-mm-dd*T*hh:mm:ss*	P*yyyy-mm-dd*T*hh:mm:ss*	P0003-x-02T02:55:20/ 2008-x-15Tx:x:45

Datetime values with omitted components that are formatted with either the $N8601B format or the $N8601BA format are formatted in the extended notation using the hyphen for omitted components to ensure accurate data. For example, when the month is an omitted component, the value 2008—15 is written and not 2008-15.

The extended notation with hyphens is also used in place of the basic notation if a duration is formatted by using the $N8601BA format. Using the same date, P2008—15 is written and not P2008-15.

Writing Truncated Duration, Datetime, and Interval Values

Duration, datetime, or interval values can be truncated when one or more lower order values is 0 or is not significant. When SAS writes a truncated value using the formats $N8601B, $N8601BA, $N8601E, and $N8601EA, the display of the value stops at the last non-missing component.

When you format a truncated value by using either the $N8601H format or the $N8601EH format, the lower order components are written with a hyphen. When you format a truncated value by using the $N8601X format or the $N8601EX format, the lower order components are written with an x.

The following examples show truncated values:

```
p00030202T1031
2008-09-15T15/2010-09-15T15:53
-p0003-03-03T-:-:-
P2y3m4dT5h6m
2008-09-xTx:x:x
2008
```

Normalizing Duration Components

When a value for a duration component is greater than the largest standard value for a component, SAS normalizes the component except when the duration component is a single component. The following table shows examples of normalized duration components:

Duration	Extended Normalized Duration
p3y13m	p0004-01
pt24h24m65s	P----01T-:25:05
p3y13mT24h61m	P0004-01-01T01:01
p0004-13	p0005-01
p0003-02-61T15:61:61	P0003-04-01T16:02:01
p13m	P13M

If a component contains the largest value, such as 60 for minutes or seconds, SAS normalizes the value and replaces the value with a hyphen. For example, **pT12:60:13** becomes **PT13:-:13**.

Thirty days is used to normalize a month.

Dates and times in a datetime value that are greater than the standard value for the component are not normalized. They produce an error.

Fractions in Durations, Datetime, and Interval Values

Ending components can contain a fraction that consists of a period or a comma, followed by one to three digits. The following examples show the use of fractions in duration, datetime, and interval values:

```
200809.5
P2008-09-15T10.33
2008-09-15/P0003-03-03,333
```

Formats by Category

There are four categories of formats in this list:

Category	Description
Character	instructs SAS to write character data values from character variables.
Date and Time	instructs SAS to write data values from variables that represent dates, times, and datetimes.
ISO 8601	instructs SAS to write date, time, and datetime values using the ISO 8601 standard.
Numeric	instructs SAS to write numeric data values from numeric variables.

Formats that support national languages can be found in *SAS National Language Support (NLS): Reference Guide*. A listing of national language formats is provided in "Formats Documented in Other SAS Publications" on page 285.

Storing user-defined formats is an important consideration if you associate these formats with variables in permanent SAS data sets, especially those data sets shared with other users. For information about creating and storing user-defined formats, see "The FORMAT Procedure" in *Base SAS Procedures Guide*.

The following table provides brief descriptions of the SAS formats. For more detailed descriptions, see the dictionary entry for each format.

Table 3.4 Categories and Descriptions of Formats

Category	Formats	Description
Character	"$ASCII*w*. Format" on page 108	Converts native format character data to ASCII representation.
	"$BASE64X*w*. Format" on page 109	Converts character data into ASCII text by using Base 64 encoding.
	"$BINARY*w*. Format" on page 110	Converts character data to binary representation.
	"$CHAR*w*. Format" on page 111	Writes standard character data.
	"$EBCDIC*w*. Format" on page 112	Converts native format character data to EBCDIC representation.
	"$HEX*w*. Format" on page 113	Converts character data to hexadecimal representation.
	"$MSGCASE*w*. Format" on page 114	Writes character data in uppercase when the MSGCASE system option is in effect.
	"$OCTAL*w*. Format" on page 126	Converts character data to octal representation.
	"$QUOTE*w*. Format" on page 128	Writes data values that are enclosed in double quotation marks.
	"$REVERJ*w*. Format" on page 129	Writes character data in reverse order and preserves blanks.

Category	Formats	Description
	"$REVERS*w*. Format" on page 130	Writes character data in reverse order and left aligns
	"$UPCASE*w*. Format" on page 131	Converts character data to uppercase.
	"$VARYING*w*. Format" on page 132	Writes character data of varying length.
	"$*w*. Format" on page 134	Writes standard character data.
Date and Time	"$N8601B*w.d* Format" on page 115	Writes ISO 8601 duration, datetime, and interval forms using the basic notations P*n*Y*n*M*n*DT*n*H *n*M*n*S and *yyyymmdd*T*hhmmss*.
	"$N8601BA*w.d* Format" on page 117	Writes ISO 8601 duration, datetime, and interval forms using the basic notations P*yyyymmdd*T*hhmmss* and *yyyymmdd* T*hhmmss*.
	"$N8601Ew.d Format" on page 118	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*n*Y*n*M*n*DT*n* H*n*M*n*S and *yyyy-mm-dd*T *hh:mm:ss*.
	"$N8601EA*w.d* Format" on page 119	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*.
	"$N8601EH*w.d* Format" on page 121	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using a hyphen (-)for omitted components.
	"$N8601EX*w.d* Format" on page 122	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using an x for digit of an omitted component.
	"$N8601H*w.d* Format" on page 123	Writes ISO 8601 duration, datetime, and interval forms P *n*Y*n*M*n*DT*n*H*n*M *n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using a hyphen (-)for omitted components in datetime values.
	"$N8601X*w.d* Format" on page 125	Writes ISO 8601 duration, datetime, and interval forms P *n*Y*n*M*n*DT*n*H*n*M *n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using an x for each digit of an omitted component in datetime values.
	"B8601DA*w*. Format" on page 138	Writes date values using the IOS 8601 base notation *yyyymmdd* .
	"B8601DN*w*. Format" on page 139	Writes the date from a datetime value using the ISO 8601 basic notation *yyyymmdd*.
	"B8601DT*w.d* Format" on page 140	Writes datetime values in the ISO 8601 basic notation *yyyymmdd* T*hhmmssffffff*.
	"B8601DZ*w*. Format" on page 141	Writes datetime values in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime and time zone basic notation *yyyymmdd* T*hhmmss*+|-*hhmm*.

Category	Formats	Description
	"B8601LZ*w.* Format" on page 143	Writes time values as local time by appending a time zone offset difference between the local time and UTC, using the ISO 8601 basic time notation *hhmmss+|-hhmm*.
	"B8601TM*w.d* Format" on page 144	Writes time values using the ISO 8601 basic notation *hhmmssffff* .
	"B8601TZ*w.* Format" on page 145	Adjusts time values to the Coordinated Universal Time (UTC) and writes them using the ISO 8601 basic time notation *hhmmss+|- hhmm*.
	"DATE*w*. Format" on page 151	Writes date values in the form *ddmmmyy*, *ddmmmyyyy* , or *dd-mmm-yyyy*.
	"DATEAMPM*w.d* Format" on page 153	Writes datetime values in the form *ddmmmyy:hh:mm:ss.ss* with AM or PM.
	"DATETIME*w.d* Format" on page 154	Writes datetime values in the form *ddmmmyy:hh:mm:ss.ss*.
	"DAY*w*. Format" on page 156	Writes date values as the day of the month.
	"DDMMYY*w*. Format" on page 157	Writes date values in the form *ddmm<yy> yy* or *dd/mm/<yy>yy* , where a forward slash is the separator and the year appears as either 2 or 4 digits.
	"DDMMYY*xw*. Format" on page 158	Writes date values in the form *ddmm<yy> yy* or *dd-mm-yy<yy>*, where the *x* in the format name is a character that represents the special character that separates the day, month, and year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.
	"DOWNAME*w*. Format" on page 163	Writes date values as the name of the day of the week.
	"DTDATE*w*. Format" on page 164	Expects a datetime value as input and writes date values in the form *ddmmmyy* or *ddmmmyyyy*.
	"DTMONYY*w*. Format" on page 165	Writes the date part of a datetime value as the month and year in the form *mmmyy* or *mmmyyyy*.
	"DTWKDATX*w*. Format" on page 166	Writes the date part of a datetime value as the day of the week and the date in the form *day-of-week*, *dd month-name yy* (or *yyyy*).
	"DTYEAR*w*. Format" on page 168	Writes the date part of a datetime value as the year in the form *yy* or *yyyy*.
	"DTYYQC*w*. Format" on page 169	Writes the date part of a datetime value as the year and the quarter and separates them with a colon (:).
	"E8601DA*w.* Format" on page 171	Writes date values using the ISO 8601 extended notation *yyyy-mm-dd*.
	"E8601DN*w*. Format" on page 172	Writes the date from a SAS datetime value using the ISO 8601 extended notation *yyyy-mm-dd*.
	"E8601DT*w.d* Format" on page 173	Writes datetime values in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.

Category	Formats	Description
	"E8601DZ*w.* Format" on page 174	Writes datetime values in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime and time zone extended notations *yyyy-mm-dd*T*hh:mm:ss+|-hh:mm*.
	"E8601LZ*w.* Format" on page 175	Writes time values as local time, appending the Coordinated Universal Time (UTC) offset for the local SAS session, using the ISO 8601 extended time notation *hh:mm:ss+|-hh:mm*.
	"E8601TM*w.d* Format" on page 177	Writes time values using the ISO 8601 extended notation *hh:mm:ss.ffffff*.
	"E8601TZ*w.d* Format" on page 179	Adjusts time values to the Coordinated Universal Time (UTC) and writes them using the ISO 8601 extended notation *hh*:*mm* :*ss+|-hh*:*mm*.
	"HHMM*w.d* Format" on page 185	Writes time values as hours and minutes in the form *hh:mm* .
	"HOUR*w.d* Format" on page 188	Writes time values as hours and decimal fractions of hours.
	"JULDAY*w.* Format" on page 193	Writes date values as the Julian day of the year.
	"JULIAN*w.* Format" on page 194	Writes date values as Julian dates in the form *yyddd* or *yyyyddd*.
	"MMDDYY*w.* Format" on page 196	Writes date values in the form *mmdd*<*yy*> *yy* or *mm/dd/*<*yy*>*yy* , where a forward slash is the separator and the year appears as either 2 or 4 digits.
	"MMDDYY*xw.* Format" on page 198	Writes date values in the form *mmdd*<*yy*> *yy* or *mm-dd-*<*yy*>*yy* , where the *x* in the format name is a character that represents the special character which separates the month, day, and year. The special character can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.
	"MMSS*w.d* Format" on page 200	Writes time values as the number of minutes and seconds since midnight.
	"MMYY*w.* Format" on page 201	Writes date values in the form *mm*M<*yy*> *yy*, where M is the separator and the year appears as either 2 or 4 digits.
	"MMYY*xw.* Format" on page 203	Writes date values in the form *mm*<*yy*> *yy* or *mm-*<*yy*>*yy*, where the *x* in the format name is a character that represents the special character that separates the month and the year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.
	"MONNAME*w.* Format" on page 204	Writes date values as the name of the month.
	"MONTH*w.* Format" on page 205	Writes date values as the month of the year.

Category	Formats	Description
	"MONYY*w*. Format" on page 206	Writes date values as the month and the year in the form *mmmyy* or *mmmyyyy*.
	"PDJULG*w*. Format" on page 213	Writes packed Julian date values in the hexadecimal format *yyyyddd*F for IBM.
	"PDJULI*w*. Format" on page 214	Writes packed Julian date values in the hexadecimal format *ccyyddd*F for IBM.
	"QTR*w*. Format" on page 224	Writes date values as the quarter of the year.
	"QTRR*w*. Format" on page 225	Writes date values as the quarter of the year in Roman numerals.
	"TIME*w.d* Format" on page 245	Writes time values as hours, minutes, and seconds in the form *hh:mm:ss.ss*.
	"TIMEAMPM*w.d* Format" on page 247	Writes time values as hours, minutes, and seconds in the form *hh:mm:ss.ss* with AM or PM.
	"TOD*w.d* Format" on page 249	Writes SAS time values and the time portion of SAS datetime values in the form *hh:mm:ss.ss*.
	"WEEKDATE*w*. Format" on page 255	Writes date values as the day of the week and the date in the form *day-of-week*, *month-name dd*, *yy* (or *yyyy*).
	"WEEKDATX*w*. Format" on page 257	Writes date values as the day of the week and date in the form *day-of-week*, *dd month-name yy* (or *yyyy*).
	"WEEKDAY*w*. Format" on page 258	Writes date values as the day of the week.
	"WEEKU*w*. Format" on page 259	Writes a week number in decimal format by using the U algorithm.
	"WEEKV*w*. Format" on page 261	Writes a week number in decimal format by using the V algorithm.
	"WEEKW*w*. Format" on page 263	Writes a week number in decimal format by using the W algorithm.
	"WORDDATE*w*. Format" on page 265	Writes date values as the name of the month, the day, and the year in the form *month-name dd*, *yyyy*.
	"WORDDATX*w*. Format" on page 266	Writes date values as the day, the name of the month, and the year in the form *dd month-name yyyy*.
	"YEAR*w*. Format" on page 269	Writes date values as the year.
	"YYMM*w*. Format" on page 270	Writes date values in the form <*yy*>*yy*M *mm*, where M is a character separator to indicate that the month number follows the M and the year appears as either 2 or 4 digits.
	"YYMM*x*w. Format" on page 271	Writes date values in the form <*yy*>*yymm* or <*yy*>*yy-mm*, where the *x* in the format name is a character that represents the special character that separates the year and the month, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.

Category	Formats	Description
	"YYMMDD*w*. Format" on page 273	Writes date values in the form *yymmdd* or < *yy*>*yy-mm-dd*, where a dash is the separator and the year appears as either 2 or 4 digits.
	"YYMMDD*xw*. Format" on page 274	Writes date values in the form *yymmdd* or < *yy*>*yy-mm-dd*, where the *x* in the format name is a character that represents the special character which separates the year, month, and day. The special character can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.
	"YYMON*w*. Format" on page 276	Writes date values in the form *yymmm* or *yyyymmm* .
	"YYQ*w*. Format" on page 277	Writes date values in the form <*yy*>*yy*Q *q*, where Q is the separator, the year appears as either 2 or 4 digits, and *q* is the quarter of the year.
	"YYQ*xw*. Format" on page 278	Writes date values in the form <*yy*>*yyq* or <*yy*>*yy-q*, where the *x* in the format name is a character that represents the special character that separates the year and the quarter or the year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.
	"YYQR*w*. Format" on page 280	Writes date values in the form <*yy*>*yy*Q *qr*, where Q is the separator, the year appears as either 2 or 4 digits, and *qr* is the quarter of the year expressed in roman numerals.
	"YYQR*xw*. Format" on page 281	Writes date values in the form <*yy*>*yy qr* or <*yy*>*yy-qr*, where the *x* in the format name is a character that represents the special character that separates the year and the quarter or the year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits and *qr* is the quarter of the year expressed in roman numerals.
ISO 8601	"$N8601B*w.d* Format" on page 115	Writes ISO 8601 duration, datetime, and interval forms using the basic notations P*n*Y*n*M*n*DT*n*H *n*M*n*S and *yyyymmdd*T*hhmmss*.
	"$N8601BA*w.d* Format" on page 117	Writes ISO 8601 duration, datetime, and interval forms using the basic notations P*yyyymmdd*T*hhmmss* and *yyyymmdd* T*hhmmss*.
	"$N8601Ew.d Format" on page 118	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*n*Y*n*M*n*DT*n* H*n*M*n*S and *yyyy-mm-dd*T *hh:mm:ss*.
	"$N8601EA*w.d* Format" on page 119	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*.
	"$N8601EH*w.d* Format" on page 121	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using a hyphen (-)for omitted components.

Category	Formats	Description
	"$N8601EX*w.d* Format" on page 122	Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using an x for digit of an omitted component.
	"$N8601H*w.d* Format" on page 123	Writes ISO 8601 duration, datetime, and interval forms P *n*Y*n*M*n*DT*n*H*n*M *n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using a hyphen (-)for omitted components in datetime values.
	"$N8601X*w.d* Format" on page 125	Writes ISO 8601 duration, datetime, and interval forms P *n*Y*n*M*n*DT*n*H*n*M *n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using an x for each digit of an omitted component in datetime values.
	"B8601DA*w*. Format" on page 138	Writes date values using the IOS 8601 base notation *yyyymmdd* .
	"B8601DN*w*. Format" on page 139	Writes the date from a datetime value using the ISO 8601 basic notation *yyyymmdd*.
	"B8601DT*w.d* Format" on page 140	Writes datetime values in the ISO 8601 basic notation *yyyymmdd* T*hhmmssffffff*.
	"B8601DZ*w*. Format" on page 141	Writes datetime values in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime and time zone basic notation *yyyymmdd* T*hhmmss+|-hhmm*.
	"B8601LZ*w*. Format" on page 143	Writes time values as local time by appending a time zone offset difference between the local time and UTC, using the ISO 8601 basic time notation *hhmmss+|-hhmm*.
	"B8601TM*w.d* Format" on page 144	Writes time values using the ISO 8601 basic notation *hhmmssffff* .
	"B8601TZ*w*. Format" on page 145	Adjusts time values to the Coordinated Universal Time (UTC) and writes them using the ISO 8601 basic time notation *hhmmss+|- hhmm*.
	"E8601DA*w*. Format" on page 171	Writes date values using the ISO 8601 extended notation *yyyy-mm-dd*.
	"E8601DN*w*. Format" on page 172	Writes the date from a SAS datetime value using the ISO 8601 extended notation *yyyy-mm-dd*.
	"E8601DT*w.d* Format" on page 173	Writes datetime values in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.
	"E8601DZ*w*. Format" on page 174	Writes datetime values in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime and time zone extended notations *yyyy-mm-dd*T*hh:mm:ss+|-hh*:*mm*.
	"E8601LZ*w*. Format" on page 175	Writes time values as local time, appending the Coordinated Universal Time (UTC) offset for the local SAS session, using the ISO 8601 extended time notation *hh:mm:ss+|-hh:mm*.

Category	Formats	Description
	"E8601TM*w.d* Format" on page 177	Writes time values using the ISO 8601 extended notation *hh:mm:ss.ffffff*.
	"E8601TZ*w.d* Format" on page 179	Adjusts time values to the Coordinated Universal Time (UTC) and writes them using the ISO 8601 extended notation *hh*:*mm* :*ss*+|-*hh*:*mm*.
Numeric	"BEST*w*. Format" on page 134	SAS chooses the best notation.
	"BESTD*w.p* Format" on page 136	Prints numeric values, lining up decimal places for values of similar magnitude, and prints integers without decimals.
	"BINARY*w*. Format" on page 137	Converts numeric values to binary representation.
	"COMMA*w.d* Format" on page 147	Writes numeric values with a comma that separates every three digits and a period that separates the decimal fraction.
	"COMMAX*w.d* Format" on page 148	Writes numeric values with a period that separates every three digits and a comma that separates the decimal fraction.
	"D*w.p* Format" on page 149	Prints numeric values, possibly with a great range of values, lining up decimal places for values of similar magnitude.
	"DOLLAR*w.d* Format" on page 160	Writes numeric values with a leading dollar sign, a comma that separates every three digits, and a period that separates the decimal fraction.
	"DOLLARX*w.d* Format" on page 161	Writes numeric values with a leading dollar sign, a period that separates every three digits, and a comma that separates the decimal fraction.
	"E*w*. Format" on page 170	Writes numeric values in scientific notation.
	"FLOAT*w.d* Format" on page 181	Generates a native single-precision, floating-point value by multiplying a number by 10 raised to the *d*th power.
	"FRACT*w*. Format" on page 183	Converts numeric values to fractions.
	"HEX*w*. Format" on page 184	Converts real binary (floating-point) values to hexadecimal representation.
	"IB*w.d* Format" on page 189	Writes native integer binary (fixed-point) values, including negative values.
	"IBR*w.d* Format" on page 190	Writes integer binary (fixed-point) values in Intel and DEC formats.
	"IEEE*w.d* Format" on page 192	Generates an IEEE floating-point value by multiplying a number by 10 raised to the *d*th power.
	"NEGPAREN*w.d* Format" on page 208	Writes negative numeric values in parentheses.
	"NUMX*w.d* Format" on page 209	Writes numeric values with a comma in place of the decimal point.

Category	Formats	Description
	"OCTAL*w*. Format" on page 210	Converts numeric values to octal representation.
	"PD*w*.*d* Format" on page 211	Writes data in packed decimal format.
	"PERCENT*w*.*d* Format" on page 216	Writes numeric values as percentages.
	"PERCENTN*w*.*d* Format" on page 217	Produces percentages, using a minus sign for negative values.
	"PIB*w*.*d* Format" on page 219	Writes positive integer binary (fixed-point) values.
	"PIBR*w*.*d* Format" on page 221	Writes positive integer binary (fixed-point) values in Intel and DEC formats.
	"PK*w*.*d* Format" on page 222	Writes data in unsigned packed decimal format.
	"PVALUE*w*.*d* Format" on page 223	Writes *p*-values.
	"RB*w*.*d* Format" on page 226	Writes real binary data (floating-point) in real binary format.
	"ROMAN*w*. Format" on page 228	Writes numeric values as roman numerals.
	"S370FF*w*.*d* Format" on page 229	Writes native standard numeric data in IBM mainframe format.
	"S370FIB*w*.*d* Format" on page 230	Writes integer binary (fixed-point) values, including negative values, in IBM mainframe format.
	"S370FIBU*w*.*d* Format" on page 231	Writes unsigned integer binary (fixed-point) values in IBM mainframe format.
	"S370FPD*w*.*d* Format" on page 233	Writes packed decimal data in IBM mainframe format.
	"S370FPDU*w*.*d* Format" on page 235	Writes unsigned packed decimal data in IBM mainframe format.
	"S370FPIB*w*.*d* Format" on page 236	Writes positive integer binary (fixed-point) values in IBM mainframe format.
	"S370FRB*w*.*d* Format" on page 237	Writes real binary (floating-point) data in IBM mainframe format.
	"S370FZD*w*.*d* Format" on page 239	Writes zoned decimal data in IBM mainframe format.
	"S370FZDL*w*.*d* Format" on page 240	Writes zoned decimal leading–sign data in IBM mainframe format.
	"S370FZDS*w*.*d* Format" on page 241	Writes zoned decimal separate leading-sign data in IBM mainframe format.
	"S370FZDT*w*.*d* Format" on page 242	Writes zoned decimal separate trailing-sign data in IBM mainframe format.
	"S370FZDU*w*.*d* Format" on page 243	Writes unsigned zoned decimal data in IBM mainframe format.

Category	Formats	Description
	"SSN*w*. Format" on page 244	Writes Social Security numbers.
	"VAXRB*w.d* Format" on page 252	Writes real binary (floating-point) data in VMS format.
	"VMSZN*w.d* Format" on page 253	Generates VMS and MicroFocus COBOL zoned numeric data.
	"*w.d* Format" on page 254	Writes standard numeric data one digit per byte.
	"WORDF*w*. Format" on page 267	Writes numeric values as words with fractions that are shown numerically.
	"WORDS*w*. Format" on page 268	Writes numeric values as words.
	"Z*w.d* Format" on page 283	Writes standard numeric data with leading 0s.
	"ZD*w.d* Format" on page 284	Writes numeric data in zoned decimal format .

Dictionary

$ASCII*w*. Format

Converts native format character data to ASCII representation.

Category: Character
Alignment: left

Syntax

$ASCII*w*.

Syntax Description

w
specifies the width of the output field.
Default: 1
Range: 1–32767

Details

If ASCII is the native format, no conversion occurs.

Comparisons

- On EBCDIC systems, $ASCII*w*. converts EBCDIC character data to ASCII*w*.

- On all other systems, $ASCII*w*. behaves like the $CHAR*w*. format.

Examples

```
put x $ascii3.;
```

Value of **x**	Results
abc	**616263**
ABC	**414243**
();	**28293B**

* The results are hexadecimal representations of ASCII codes for characters. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one character.

$BASE64X*w*. Format

Converts character data into ASCII text by using Base 64 encoding.

Category: Character

Alignment: left

Syntax

$BAS64X*w*.

Syntax Description

w

specifies the width of the output field.

Default: 1

Range: 1-32767

Details

Base 64 is an industry encoding method whose encoded characters are determined by using a positional scheme that uses only ASCII characters. Several Base 64 encoding schemes have been defined by the industry for specific uses, such as e-mail or content masking. SAS maps positions 0 - 61 to the characters A - Z, a - z, and 0 - 9. Position 62 maps to the character +, and position 63 maps to the character /.

The following are some uses of Base 64 encoding:

- embed binary data in an XML file
- encode passwords
- encode URLs

The '=' character in the encoded results indicates that the results have been padded with zero bits. In order for the encoded characters to be decoded, the '=' must be included in the value to be decoded.

Examples

put x $base64x64.;

Value of x	Results
"FCA01A7993BC"	RkNBMDFBNzk5M0JD
"MyPassword"	TXlQYXNzd29yZA==
"www.mydomain.com/myhiddenURL"	d3d3Lm15ZG9tYWluLmNvbi9teWhpZGRlblVSTA==

See Also

Informat:

"$BASE64X*w*. Informat" on page 1281

The XMLDOUBLE option of the LIBNAME Statement for the XML engine, in *SAS XML LIBNAME Engine: User's Guide*

$BINARY*w*. Format

Converts character data to binary representation.

Category: Character
Alignment: left

Syntax

$BINARY*w*.

Syntax Description

w

specifies the width of the output field.

Default: The default width is calculated based on the length of the variable to be printed.

Range: 1–32767

Comparisons

The $BINARY*w*. format converts character values to binary representation. The BINARY*w*. format converts numeric values to binary representation.

Examples

```
put @1 name $binary16.;
```

Value of **name**	Results	
	ASCII	EBCDIC
	----+----1----+----2	**----+----1----+----2**
AB	**0100000101000010**	**1100000111000010**

$CHAR*w*. Format

Writes standard character data.

Category: Character
Alignment: left

Syntax

$CHAR*w*.

Syntax Description

w
specifies the width of the output field.
Default: 8 if the length of variable is undefined; otherwise, the length of the variable
Range: 1–32767

Comparisons

- The $CHAR*w*. format is identical to the $*w*. format.
- The $CHAR*w*. and $*w*. formats do not trim leading blanks. To trim leading blanks, use the LEFT function to left align character data, or use the PUT statement with the colon (:) format modifier and the format of your choice to produce list output.
- Use the following table to compare the SAS format $CHAR8. with notation in other programming languages:

Language	Notation
SAS	$CHAR8.
C	char [8]
COBOL	PIC x(8)
Fortran	A8
PL/I	A(8)

Examples

```
put @7 name $char4.;
```

Value of **name**	Results
	`----+----1`
XYZ	**XYZ**

$EBCDIC*w*. Format

Converts native format character data to EBCDIC representation.

Category: Character
Alignment: left

Syntax

$EBCDIC*w*.

Syntax Description

w
specifies the width of the output field.
Default: 1
Range: 1–32767

Details

If EBCDIC is the native format, no conversion occurs.

Comparisons

- On ASCII systems, $EBCDIC*w*. converts ASCII character data to EBCDIC.
- On all other systems, $EBCDIC*w*. behaves like the $CHAR*w*. format.

Examples

```
put name $ebcdic3.;
```

Value of **name**	Results*
qrs	**9899A2**
QRS	**D8D9E2**
+;>	**4E5E6E**

* The results are shown as hexadecimal representations of EBCDIC codes for characters. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one character.

$HEX*w*. Format

Converts character data to hexadecimal representation.

Category: Character

Alignment: left

See: $HEX*w*. Format in the documentation for your operating environment.

Syntax

$HEX*w*.

Syntax Description

w

specifies the width of the output field.

Default: The default width is calculated based on the length of the variable to be printed.

Range: 1–32767

Tip: To ensure that SAS writes the full hexadecimal equivalent of your data, make *w* twice the length of the variable or field that you want to represent.

Tip: If *w* is greater than twice the length of the variable that you want to represent, $HEX*w*. pads it with blanks.

Details

The $HEX*w*. format converts each character into two hexadecimal characters. Each blank counts as one character, including trailing blanks.

Comparisons

The HEX*w*. format converts real binary numbers to their hexadecimal equivalent.

Examples

```
put @5 name $hex4.;
```

Value of **name**	Results	
	EBCDIC	ASCII
	`----+----1`	`----+----1`
AB	**C1C2**	**4142**

$MSGCASE*w*. Format

Writes character data in uppercase when the MSGCASE system option is in effect.

Category: Character
Alignment: left

Syntax

$MSGCASE*w*.

Syntax Description

w
specifies the width of the output field.

Default: 1 if the length of the variable is undefined. Otherwise, the default is the length of the variable

Range: 1–32767

Details

When the MSGCASE= system option is in effect, all notes, warnings, and error messages that SAS generates appear in uppercase. Otherwise, all notes, warnings, and error messages appear in mixed case. You specify the MSGCASE= system option in the configuration file or during the SAS invocation.

Operating Environment Information: For more information about the MSGCASE= system option, see the SAS documentation for your operating environment. △

Examples

```
put name $msgcase.;
```

Value of **name**	Results
sas	**SAS**

$N8601B*w.d* Format

Writes ISO 8601 duration, datetime, and interval forms using the basic notations P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyymmdd*T*hhmmss*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: No

ISO 8601 Element: 5.4.4 Complete representation

Syntax

$N8601B*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d

specifies the number of digits to the right of the lowest order component. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601B format writes ISO 8601 duration, datetime, and interval values as character data for the following basic notations:

P*n*Y*n*M*n*DT*n*H*n*M*n*S

*yyyymmdd*T*hhmmss*

P*n*Y*n*M*n*DT*n*H*n*M*n*S/*yyyymmdd*T*hhmmss*

*yyyymmdd*T*hhmmss*T/P*n*Y*n*M*n*DT*n*H*n*M*n*S

The lowest order component can contain fractions, as in these examples:

```
p2y3.5m
p00020304T05.335
```

Examples

```
put nb $n8601b.;
```

Value of nb	Results
0002405050112FFC	P2Y4M5DT5H1M12S
2008915155300FFD	20080915T155300
2008915000000FFD2010915000000FFD	20080915T000000/20100915T000000
0033104030255FFC2008915155300FFD	P33Y1M4DT3H2M55S/20080915T155300

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$N8601BA*w.d* Format

Writes ISO 8601 duration, datetime, and interval forms using the basic notations P*yyyymmdd*T*hhmmss* and *yyyymmdd*T*hhmmss*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: No

ISO 8601 Element: 5.5.4.2 Alternative format

Syntax

$N8601BA*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d
specifies the number of digits to the right of the lowest order component. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601BA format writes ISO 8601 duration, datetime, and interval values as character data for the following basic notations:

P*yyyymmdd*T*hhmmss*

*yyyymmdd*T*hhmmss*

P*yyyymmdd*T*hhmmss*/*yyyymmdd*T*hhmmss*

*yyyymmdd*T*hhmmss*/P*yyyymmdd*T*hhmmss*

The lowest order component can contain fractions, as in these examples:

```
p00023.5
00020304T05.335
```

Examples

```
put @1 nba $N8601ba.;
```

Value of nba	Results
00024050501127D0	P00020405T050112.5
2008915155300FFD	20080915T155300
000230405060752820089151553 00FFD	P00020304T050607.33/20080915T155300

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$N8601Ew.d Format

Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyy-mm-dd*T*hh:mm:ss*.

Category: Date and Time
ISO 8601
Alignment: left
Time Zone Format: No
ISO 8601 Element: 5.4.4 Complete representation

Syntax

$N8601E*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 50
Range: 1 - 200
Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d
specifies the number of digits to the right of the lowest order component. This argument is optional.
Default: 0
Range: 0 - 3

Details

The $N8601B format writes ISO 8601 duration, datetime, and interval values as character data for the following basic notations:

P*n*Y*n*M*n*DT*n*H*n*M*n*S

*yyyy-mm-dd*T*hh:mm:ss*

P*n*Y*n*M*n*DT*n*H*n*M*n*S/*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*T/P*n*Y*n*M*n*DT*n*H*n*M*n*S

The lowest order component can contain fractions, as in these examples:

```
p2y3.5m
p0002--03--04T05.335
```

Examples

```
put @1 ne $n8601e.;
```

Value of ne	Results
`00024050501127D0`	`P2Y4M5DT5H1M12.5S`
`2008915155300FFD`	`2008-09-15T15:53:00`
`2008915000000FFD2010915000000FFD`	`2008-09-15T00:00:00/2010-09-15T00:00:00`
`0033104030255FFC2008915155300FFD`	`P33Y1M4DT3H2M55S/2008-09-15T15:53:00`

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$N8601EA*w.d* Format

Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: No

ISO 8601 Element: 5.4.4 Complete representation

Syntax

$N8601EA*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d
specifies the number of digits to the right of the lowest order component. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601EA format writes ISO 8601 duration, datetime, and interval values as character data for the following basic notations:

P*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*

P*yyyy-mm-dd*T*hh:mm:ss*/*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*/P*yyyy-mm-dd*T*hh:mm:ss*

The lowest order component can contain fractions, as in these examples:

```
p00023.5
0002--03--04T05.335
```

Examples

```
put @1 nea $N8601ea.;
```

Value of nea	Results
00024050501127D0	P0002-04-05T05:01:12.500
2008915155300FFD	2008-09-15T15:53:00
000230405060752820089151553OOFFD	P0002-03-04T05:06:07.330/2008-09-15T15:53:00

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$N8601EH*w.d* Format

Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using a hyphen (-)for omitted components.

Category: Date and Time
ISO 8601

Time Zone Format: No

ISO 8601 Element: 5.4.4 Complete representation

Syntax

$N8601EH*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d
specifies the number of digits to the right of the lowest order component. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601H format writes ISO 8601 duration, datetime, and interval values as character data, using a hyphen (-)to represent omitted components, for the following extended notations:

P*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*

P*yyyy-mm-dd*T*hh:mm:ss*/*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*/P*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*/*yyyy-mm-dd*T*hh:mm:ss*

Omitted datetime components are always displayed, they are never truncated.

Examples

```
put a $n8601eh.;
```

Value of a	Results
`00023FFFFFFFFFC2008FFF15FFFFFFD`	`P0002-03---T-:-:-/2008------T15:-:-`
`2008FFF15FFFFFFdFFFF3FF1553FFFFC`	`2008------T15:-:-/P---03---T15:53:-`

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$N8601EX*w.d* Format

Writes ISO 8601 duration, datetime, and interval forms using the extended notations P*yyyy-mm-dd*T*hh:mm:ss* and *yyyy-mm-dd*T*hh:mm:ss*, using an x for digit of an omitted component.

Category: Date and Time
ISO 8601
Alignment: left
Time Zone Format: No
ISO 8601 Element: 5.5.3, 5.5.4.1, 5.5.4.2

Syntax

$N8601X*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d
specifies the number of digits to the right of the lowest order component. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601H format writes ISO 8601 duration, datetime, and interval values as character data, using a hyphen (-) to represent omitted components, for the following extended notations:

P*yyyy-mm-dd*T*hh:mm:ss*
*yyyy-mm-dd*T*hh:mm:ss*
P*yyyy-mm-dd*T*hh:mm:ss*/*yyyy-mm-dd*T*hh:mm:ss*
*yyyy-mm-dd*T*hh:mm:ss*/P*yyyy-mm-dd*T*hh:mm:ss*
*yyyy-mm-dd*T*hh:mm:ss*/*yyyy-mm-dd*T*hh:mm:ss*

Omitted datetime components are always displayed, they are never truncated.

Examples

```
put nex $n8601ex.;
```

Value of nex	Results
`00023FFFFFFFFFC2008FFF15FFFFFFD`	`P0002-03xxTxx:xx:xx/2008--xx-xxT15:xx:xx`
`2008FFF15FFFFFFdFFFF3FF1553FFFFC`	`2008-xx-xxT15:xx:xx/Pxxxx-03-xxT15:53:xx`

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$N8601H*w.d* Format

Writes ISO 8601 duration, datetime, and interval forms P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using a hyphen (-)for omitted components in datetime values.

Category: Date and Time
ISO 8601
Alignment: left
Time Zone Format: No
ISO 8601 Element: 5.5.3, 5.5.4.1, 5.5.4.2

Syntax

$N8601H*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d

specifies the number of digits to the right of the lowest order component. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601H format writes ISO 8601 durations, intervals, and datetimes in the following forms, omitting components in the P*n*Y*n*M*n*DT*n*H*n*M*n*S form and using a hyphen (-)to represent omitted components in the datetime form:

P*n*Y*n*M*n*DT*n*H*n*M*n*S

*yyyy-mm-dd*T*hh:mm:ss*

P*n*Y*n*M*n*DT*n*H*n*M*n*S/*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*T/P*n*Y*n*M*n*DT*n*H*n*M*n*S

*yyyy-mm-dd*T*hh:mm:ss*/*yyyy-mm-dd*T*hh:mm:ss*

Omitted datetime components are always displayed, they are never truncated.

Examples

```
put nh $n8601h.;
```

Value of nh	Results
`0002304FFFFFFFFC2008FFF15FFFFFFD`	`P2Y3M4D/2008------T15:-:-`
`FFFF102FFFFFFFFD2008FFF15FFFFFFD`	`---01-02T-:-:-0/2008------T15:-:-`

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$N8601X*w.d* Format

Writes ISO 8601 duration, datetime, and interval forms P*n*Y*n*M*n*DT*n*H*n*M*n*S and *yyyy-mm-dd*T*hh:mm:ss*, dropping omitted components in duration values and using an x for each digit of an omitted component in datetime values.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: No

ISO 8601 Element: 5.5.3, 5.5.4.1, 5.5.4.2

Syntax

$N8601X*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d

specifies the number of digits to the right of the lowest order component. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601X format writes ISO 8601 durations, intervals, and datetimes in the following forms, omitting components in the P*n*Y*n*M*n*DT*n*H*n*M*n*S form and using an x to represent omitted components in the datetime form:

P*n*Y*n*M*n*DT*n*H*n*M*n*S

*yyyy-mm-dd*T*hh:mm:ss*

P*n*Y*n*M*n*DT*n*H*n*M*n*S/*yyyy-mm-dd*T*hh:mm:ss*

*yyyy-mm-dd*T*hh:mm:ss*T/P*n*Y*n*M*n*DT*n*H*n*M*n*S

*yyyy-mm-dd*T*hh:mm:ss*/*yyyy-mm-dd*T*hh:mm:ss*

Omitted datetime components are always displayed, they are never truncated.

Examples

```
put nx $n8601x.;
```

Value of nx	Results
0002304FFFFFFFFC2008FFF15FFFFFFD	P2Y3M4D/2008-xx-xxT15:xx:xx
FFFF102FFFFFFFFD2008FFF15FFFFFd	xxxx-01-02Txx:xx:xx/2008-xx-xxT15:xx:xx

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

$OCTALw. Format

Converts character data to octal representation.

Category: Character
Alignment: left

Syntax

$OCTAL*w.*

Syntax Description

w

specifies the width of the output field.

Default: The default width is calculated based on the length of the variable to be printed.

Range: 1–32767

Tip: Because each character value generates three octal characters, increase the value of *w* by three times the length of the character value.

Comparisons

The $OCTAL*w*. format converts character values to the octal representation of their character codes. The OCTAL*w*. format converts numeric values to octal representation.

Example

The following example shows ASCII output when you use the $OCTAL*w*. format.

```
data _null_;
   infile datalines truncover;
   input item $5.;
   put item $octal15.;
   datalines;
art
rice
bank
;
run;
```

SAS writes the following results to the log.

```
141162164040040
162151143145040
142141156153040
```

$QUOTE*w*. Format

Writes data values that are enclosed in double quotation marks.

Category: Character
Alignment: left

Syntax

$QUOTE*w*.

Syntax Description

w
specifies the width of the output field.

Default: 2 if the length of the variable is undefined. Otherwise, the default is the length of the variable + 2

Range: 2–32767

Tip: Make *w* wide enough to include the left and right quotation marks.

Details

The following list describes the output that SAS produces when you use the $QUOTE*w*. format. For examples of these items, see "Examples" on page 129.

- If your data value is not enclosed in quotation marks, SAS encloses the output in double quotation marks.
- If your data value is not enclosed in quotation marks, but the value contains a single quotation mark, SAS
  - encloses the data value in double quotation marks
  - does not change the single quotation mark.
- If your data value begins and ends with single quotation marks, and the value contains double quotation marks, SAS
  - encloses the data value in double quotation marks
  - duplicates the double quotation marks that are found in the data value
  - does not change the single quotation marks.
- If your data value begins and ends with single quotation marks, and the value contains two single contiguous quotation marks, SAS
  - encloses the value in double quotation marks
  - does not change the single quotation marks.
- If your data value begins and ends with single quotation marks, and contains both double quotation marks and single, contiguous quotation marks, SAS
  - encloses the value in double quotation marks
  - duplicates the double quotation marks that are found in the data value
  - does not change the single quotation marks.

- If the length of the target field is not large enough to contain the string and its quotation marks, SAS returns all blanks.

Examples

```
put name $quote20.;
```

Value of **name**	Results
	`----+----1`
`SAS`	`"SAS"`
`SAS's`	`"SAS's"`
`'ad"verb"'`	`"'ad""verb""'"`
`'ad''verb'`	`"'ad''verb'"`
`'"ad"''"verb"'`	`"'""ad""''""verb""'"`
`deoxyribonucleotide`	

$REVERJ*w*. Format

Writes character data in reverse order and preserves blanks.

Category: Character
Alignment: right

Syntax

$REVERJ*w*.

Syntax Description

w
specifies the width of the output field.
Default: 1 if *w* is not specified
Range: 1–32767

Comparisons

The $REVERJ*w*. format is similar to the $REVERS*w*. format except that $REVERS*w*. left aligns the result by trimming all leading blanks.

Examples

```
put @1 name $reverj7.;
```

Name	Results
	----+----1
ABCD###	**DCBA**
###ABCD	**DCBA**

* The character # represents a blank space.

$REVERS*w*. Format

Writes character data in reverse order and left aligns

Category: Character
Alignment: left

Syntax

$REVERS*w*.

Syntax Description

w
specifies the width of the output field.

Default: 1 if *w* is not specified

Range: 1–32767

Comparisons

The $REVERS*w*. format is similar to the $REVERJ*w*. format except that $REVERJ*w*. does not left align the result.

Examples

```
put @1 name $revers7.;
```

Name	Results
	`----+----1`
ABCD###	**DCBA**
###ABCD	**DCBA**

* The character # represents a blank space.

$UPCASE*w*. Format

Converts character data to uppercase.

Category: Character

Alignment: left

Syntax

$UPCASE*w*.

Syntax Description

w
specifies the width of the output field.

Default: 8 if the length of the variable is undefined. Otherwise, the default is the length of the variable

Range: 1–32767

Details

Special characters, such as hyphens and other symbols, are not altered.

Examples

```
put @1 name $upcase9.;
```

Value of **name**	Results
	`----+----1`
`coxe-ryan`	`COXE-RYAN`

$VARYING*w*. Format

Writes character data of varying length.

Valid: in DATA step

Category: Character

Alignment: left

Syntax

$VARYING*w*. *length-variable*

Syntax Description

w
specifies the maximum width of the output field for any output line or output file record.

Default: 8 if the length of the variable is undefined. Otherwise, the default is the length of the variable

Range: 1–32767

length-variable
specifies a numeric variable that contains the length of the current value of the character variable. SAS obtains the value of the *length-variable* by reading it directly from a field that is described in an INPUT statement, reading the value of a variable in an existing SAS data set, or calculating its value.

Requirement: You must specify *length-variable* immediately after $VARYING*w*. in a SAS statement.

Restriction: *Length-variable* cannot be an array reference.

Tip: If the value of *length-variable* is 0, negative, or missing, SAS writes nothing to the output field. If the value of *length-variable* is greater than 0 but less than *w*, SAS writes the number of characters that are specified by *length-variable*. If *length-variable* is greater than or equal to *w*, SAS writes *w* columns.

Details

Use $VARYING*w*. when the length of a character value differs from record to record. After writing a data value with $VARYING*w*., the pointer's position is the first column after the value.

Examples

Example 1: Obtaining a Variable Length Directly An existing data set variable contains the length of a variable. The data values and the results follow the explanation of this SAS statement:

```
put @10 name $varying12. varlen;
```

NAME is a character variable of length 12 that contains values that vary from 1 to 12 characters in length. VARLEN is a numeric variable in the same data set that contains the actual length of NAME for the current observation.

Value of name*	Results
	`----+----1----+----2----+`
New York 8	**New York**
Toronto 7	**Toronto**
Buenos Aires 12	**Buenos Aires**
Tokyo 5	**Tokyo**

* The value of NAME appears before the value of VARLEN.

Example 2: Obtaining a Variable Length Indirectly Use the LENGTH function to determine the length of a variable. The data values and the results follow the explanation of these SAS statements:

```
varlen=length(name);
put @10 name $varying12. varlen;
```

The assignment statement determines the length of the varying-length variable. The variable VARLEN contains this length and becomes the *length-variable* argument to the $VARYING12. format.

Values*	Results
	`----+----1----+----2----+`
New York	**New York**
Toronto	**Toronto**
Buenos Aires	**Buenos Aires**
Tokyo	**Tokyo**

* The value of NAME appears before the value of VARLEN.

$w. Format

Writes standard character data.

Category: Character
Alignment: left
Alias: $F*w*.

Syntax

$*w*.

Syntax Description

w
specifies the width of the output field. You can specify a number or a column range.
Default: 1 if the length of the variable is undefined. Otherwise, the default is the length of the variable.
Range: 1–32767

Comparisons

The $*w*. format and the $CHAR*w*. format are identical, and they do not trim leading blanks. To trim leading blanks, use the LEFT function to left align character data, or use list output with the colon (:) format modifier and the format of your choice.

Examples

```
put @10 name $5.;
put name $ 10-15;
```

Value of **name***	Results
	----+----1----+----2
#Cary	**Cary**
Tokyo	**Tokyo**

* The character # represents a blank space.

BEST*w*. Format

SAS chooses the best notation.

Category: Numeric
Alignment: right
See: BEST*w*. Format in the documentation for your operating environment.

Syntax

BEST*w*.

Syntax Description

w
specifies the width of the output field.

Default: 12

Tip: If you print numbers between 0 and .01 exclusively, then use a field width of at least 7 to avoid excessive rounding. If you print numbers between 0 and -.01 exclusively, then use a field width of at least 8.

Range: 1–32

Details

When a format is not specified for writing a numeric value, SAS uses the BEST*w*. format as the default format. The BEST*w*. format writes numbers as follows:

- Values are written with the maximum precision, as determined by the width.
- Integers are written without decimals.
- Numbers with decimals are written with as many digits to the left and right of the decimal point as needed or as allowed by the width.
- Values that can be written within the given width are written without trailing zeros.
- Values that cannot be written within the given width are written with the maximum allowable number of decimal places as determined by the width.
- Extreme values might be written in scientific notation.

SAS stores the complete value regardless of the format that is used.

Comparisons

- The BEST*w*. format writes as many significant digits as possible in the output field, but if the numbers vary in magnitude, the decimal points do not line up. Integers print without a decimal.
- The D*w.p* format writes numbers with the desired precision and more alignment than the BEST*w* format.
- The BESTD*w.p* format is a combination of the BEST*w*. format and the D*w.p* format in that it formats all numeric data, and it does a better job of aligning decimals than the BEST*w*. format.
- The *w.d* format aligns decimal points, if possible, but does not necessarily show the same precision for all numbers.

Examples

The following statements produce these results.

SAS Statements	Results
	`----+----1----+----2`
`x=1257000;` `put x best6.;`	`1.26E6`
`x=1257000;` `put x best3.;`	`1E6`

See Also

Format:
"BESTD*w.p* Format" on page 136

BESTD*w.p* Format

Prints numeric values, lining up decimal places for values of similar magnitude, and prints integers without decimals.

Category: Numeric
Alignment: right

Syntax

BESTD*w.p*

Syntax Description

w
specifies the width of the output field.
Default: 12
Range: 1–32

p
specifies the precision. This argument is optional.
Default: 3
Range: 0 to *w*–1
Requirement: must be less than *w*
Tip: If *p* is omitted or is specified as 0, then *p* is set to 3.

Details

The BESTD*w.p* format writes numbers so that the decimal point aligns in groups of values with similar magnitude. Integers are printed without a decimal point. Larger values of *p* print the data values with more precision and potentially more shifts in the decimal point alignment. Smaller values of *p* print the data values with less precision and a greater chance of decimal point alignment.

The format chooses the number of decimal places to print for ranges of values, even when the underlying values can be represented with fewer decimal places.

Comparisons

- The BEST*w*. format writes as many significant digits as possible in the output field, but if the numbers vary in magnitude, the decimal points do not line up. Integers print without a decimal.
- The D*w.p* format writes numbers with the desired precision and more alignment than the BEST*w* format.
- The BESTD*w.p* format is a combination of the BEST*w*. format and the D*w.p* format in that it formats all numeric data, and it does a better job of aligning decimals than the BEST*w*. format.
- The *w.d* format aligns decimal points, if possible, but it does not necessarily show the same precision for all numbers.

Examples

```
put x bestd14.;
```

Data Line	Results
	—-+—-1—-+
12345	12345
123.45	123.4500000
1.2345	1.2345000
.12345	0.1234500
1.23456789	1.23456789

See Also

Formats:
"BEST*w*. Format" on page 134
"D*w.p* Format" on page 149

BINARY*w*. Format

Converts numeric values to binary representation.

Category: Numeric
Alignment: left

Syntax

BINARY*w*.

Syntax Description

w
specifies the width of the output field.
Default: 8
Range: 1–64

Comparisons

BINARY*w*. converts numeric values to binary representation. The $BINARY*w*. format converts character values to binary representation.

Examples

```
put @1 x binary8.;
```

Value of **x**	Results
	`----+----1`
`123.45`	`01111011`
`123`	`01111011`
`-123`	`10000101`

B8601DA*w*. Format

Writes date values using the IOS 8601 base notation *yyyymmdd*.

Category: Date and Time
ISO 8601
Alignment: left
Time Zone Format: No
ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

B8601DA*w*.

Syntax Description

w
specifies the width of the output field.
Default: 10
Requirement: The width of the output field must be 10.

Details

The B8601DA format writes the ISO 8601 basic date notation *yyyymmdd*:

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 31

Examples

```
put bda $b8601da.;
```

Value of bda	Results
17790	20080915

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

B8601DN*w*. Format

Writes the date from a datetime value using the ISO 8601 basic notation *yyyymmdd*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: No

ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

B8601DN*w.*

Syntax Description

w
specifies the width of the output field.

Default: 10

Requirement: The width of the input field must be 10.

Details

The B8601DN format writes the date from a datetime value using the ISO 8601 basic date notation *yyyymmdd*:

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 31

Examples

```
put bdn b8601dn.;
```

Value of bdn	Results
1537113180	20080915

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

B8601DT*w.d* Format

Writes datetime values in the ISO 8601 basic notation *yyyymmdd*T*hhmmssffffff*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: No

ISO 8601 Element: 5.4.1 Complete representation

Syntax

B8601DT*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 19

Range: 19 - 26

d

specifies the number of digits to the right of the seconds value that represents a fraction of a second. This argument is optional.

Default: 0

Range: 0 - 6

Details

The B8601DT format writes ISO 8601 basic datetime notation *yyyymmdd*T*hhmmss*.*ffffff*:

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 31
hh	is a two-digit hour (zero padded), between 00 - 23
mm	is a two-digit minute (zero padded), between 00 - 59
ss	is a two-digit second (zero padded), between 00 - 59
.ffffff	are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 - .

Examples

```
put bdt b8601dt.;
```

Value of bdt	Results
——+——1	
1537113180	**20080915T155300**

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

B8601DZ*w.* Format

Writes datetime values in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime and time zone basic notation *yyyymmdd*T*hhmmss*+|-*hhmm*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: Yes

ISO 8601 Element: 5.4.1 Complete representation

Syntax

B8601DZ*w*.

Syntax Description

w
specifies the width of the output field.
Default: 26
Range: 20 - 35

Details

UTC values specify a time and a time zone based on the zero meridian in Greenwich, England. The B8602DZ format writes SAS datetime values for the zero meridian date and time using one of the following ISO 8601 basic datetime notations:

*yyyymmdd*T*hhmmss*+|–*hhmm*
is the form used when*w* is large enough to support this time zone notation.

*yyyymmdd*T*hhmmss*Z
is the form used when *w* is not large enough to support the +|-*hhmm* time zone notation.

where

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 31
hh	is a two-digit hour (zero padded), between 00 - 23
mm	is a two-digit minute (zero padded), between 00 - 59
ss	is a two-digit second (zero padded), between 00 - 59
Z	indicates that the time is for zero meridian (Greenwich, England) or UTC time
+|-*hhmm*	is an hour and minute signed offset from zero meridian time. Note that the offset must be +|-*hhmm* (that is, + or - and four characters). Use + for time zones east of the zero meridian and use - for time zones west of the zero meridian. For example, +0200 indicates a two-hour time difference to the east of the zero meridian, and -0600 indicates a six-hour time differences to the west of the zero meridian. **Restriction:** The shorter form +|-*hh* is not supported.

Examples

SAS Statement	Value of bdz	Results
`put bdz b8601dz20.;`	`1537113180`	`20080915T155300Z`
`put bdz b8601dz26.;`	`1537113180`	`20080915T155300+0000`

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

B8601LZ*w*. Format

Writes time values as local time by appending a time zone offset difference between the local time and UTC, using the ISO 8601 basic time notation *hhmmss*+|-*hhmm*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: Yes. The format appends the UTC offset to the value as determined by the local SAS session.

ISO 8601 Element: 5.3.3, 5.3.4.2

Syntax

B8601LZ*w*.

Syntax Description

w
specifies the width of the output field.

Default: 14

Range: 9 - 20

Details

The B8602LZ format writes time values without making any adjustments and appends the UTC time zone offset for the local SAS session, using the following ISO 8601 basic notation:

hhmmss+|–*hhmm*

where

hh	is a two-digit hour (zero padded), between 00 - 23
mm	is a two-digit minute (zero padded), between 00 - 59
ss	is a two-digit second (zero padded), between 00 - 59
+|-*hhmm*	is an hour and minute signed offset from zero meridian time. Note that the offset must be +|-*hhmm* (that is, + or - and five characters). Use + for time zones east of the zero meridian and use - for time zones west of the zero meridian. For example, +0200 indicates a two hour time difference to the east of the zero meridian, and -0600 indicates a six hour time differences to the west of the zero meridian.

Restriction: The shorter form +|-*hh* is not supported.

When SAS reads a UTC time by using the B8601TZ informat, and the adjusted time is greater than 24 hours or less than 00 hours, SAS adjusts the value so that the time is between 000000 and 235959. If the B8601LZ format attempts to format a time outside of this time range, the time is formatted with stars to indicate that the value is out of range.

Examples

The following PUT statement writes the time for the Eastern Standard time zone:

```
put blz b8601lz.;
```

Value of blz	Results
46380	125300-0500

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

B8601TM*w.d* Format

Writes time values using the ISO 8601 basic notation *hhmmssffff*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: No

ISO 8601 Element: 5.3.1.1 Complete representation

Syntax

B8601TM*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 8

Range: 8 - 15

d

specifies the number of digits to the right of the seconds value that represent a fraction of a second. This argument is optional.

Default: 0

Range: 0 - 6

Details

The B8601TM format writes SAS time values using the following ISO 8601 basic time notation *hhmmssffffff*:

hh	is a two-digit hour (zero padded), between 00 - 23.
mm	is a two-digit minute (zero padded), between 00 - 59.
ss	is a two-digit second (zero padded), between 00 - 59.
ffffff	are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 - 9.

Examples

```
put btm b8601tm.;.
```

Value of btm	Results
57180	155300

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

B8601TZ*w.* Format

Adjusts time values to the Coordinated Universal Time (UTC) and writes them using the ISO 8601 basic time notation *hhmmss*+|-*hhmm*.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Format: Yes

ISO 8601 Element: 5.3.3, 5.3.4

Syntax

B8601TZ*w.*

Syntax Description

w
specifies the width of the output field.

Default: 14

Range: 9–20

Details

UTC time values specify a time and a time zone based on the zero meridian in Greenwich, England. The B8602TZ format adjusts the time value to be the time at the zero meridian and writes it in one of the following ISO 8601 basic time notations:

hhmmss+|–*hhmm*	is the form used when *w* is large enough to support this time notation.
*hhmmss*Z	is the form used when *w* is not large enough to support the +|-*hhmm* time zone notation.

where

hh	is a two-digit hour (zero padded), between 00 and 23.
mm	is a two-digit minute (zero padded), between 00 and 59.
ss	is a two-digit second (zero padded), between 00 and 59.
Z	indicates that the time is for zero meridian (Greenwich, England) or UTC time.
+|–*hh:mm*	is an hour and minute signed offset from zero meridian time. Note that the offset must be +|–*hhmm* (that is, + or – and four characters). Use + for time zones east of the zero meridian and use – for time zones west of the zero meridian. For example, +0200 indicates a two hour time difference to the east of the zero meridian, and –0600 indicates a six hour time differences to the west of the zero meridian. **Restriction:** The shorter form +|–*hh* is not supported.

When SAS reads a UTC time by using the B8601TZ informat, and the adjusted time is greater than 24 hours or less than 00 hours, SAS adjusts the value so that the time is between 000000 and 240000. If the B8601TZ format attempts to format a time outside of this time range, the time is formatted with stars to indicate that the value is out of range.

Comparisons

For time values between 000000 and 240000, the B8601TZ format adjusts the time value to be the time at the zero meridian and writes it in the international standard extended time notation. The B8601LZ format makes no adjustment to the time and writes time values in the international standard extended time notation, using a UTC time zone offset for the local SAS session.

Examples

```
put btz b8601tz.;
```

Values for btz	Results
73441	**202401+0000**

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

COMMA*w.d* Format

Writes numeric values with a comma that separates every three digits and a period that separates the decimal fraction.

Category: Numeric
Alignment: right

Syntax

COMMA*w*.*d*

Syntax Description

w
specifies the width of the output field.
Default: 6
Range: 1–32
Tip: Make *w* wide enough to write the numeric values, the commas, and the optional decimal point.

d
specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.
Range: 0–31
Requirement: must be less than *w*

Details

The COMMA*w*.*d* format writes numeric values with a comma that separates every three digits and a period that separates the decimal fraction.

Comparisons

- The COMMA*w*.*d* format is similar to the COMMAX*w*.*d* format, but the COMMAX*w*.*d* format reverses the roles of the decimal point and the comma. This convention is common in European countries.
- The COMMA*w*.*d* format is similar to the DOLLAR*w*.*d* format except that the COMMA*w*.*d* format does not print a leading dollar sign.

Examples

```
put @10 sales comma10.2;
```

Value of **sales**	Results
	----+----1----+----2
23451.23	23,451.23
123451.234	123,451.23

See Also

Formats:

"COMMAX*w.d* Format" on page 148
"DOLLAR*w.d* Format" on page 160

COMMAX*w.d* Format

Writes numeric values with a period that separates every three digits and a comma that separates the decimal fraction.

Category: Numeric

Alignment: right

Syntax

COMMAX*w.d*

Syntax Description

w

specifies the width of the output field. This argument is optional.

Default: 6

Range: 1–32

Tip: Make *w* wide enough to write the numeric values, the commas, and the optional decimal point.

d

specifies the number of digits to the right of the decimal point in the numeric value.

Range: 0–31

Requirement: must be less than *w*

Details

The COMMAX*w.d* format writes numeric values with a period that separates every three digits and with a comma that separates the decimal fraction.

Comparisons

The COMMA*w.d* format is similar to the COMMAX*w.d* format, but the COMMAX*w.d* format reverses the roles of the decimal point and the comma. This convention is common in European countries.

Examples

```
put @10 sales commax10.2;
```

Value of **sales**	Results
	`----+----1----+----2`
23451.23	**23.451,23**
123451.234	**123.451,23**

D*w.p* Format

Prints numeric values, possibly with a great range of values, lining up decimal places for values of similar magnitude.

Category: Numeric

Alignment: right

Syntax

D*w.p*

Syntax Description

w

specifies the width of the output field. This argument is optional.

Default: 12

Range: 1–32

p
specifies the precision. This argument is optional.

Default: 3

Range: 0–9

Requirement: *p* must be less than *w*

Tip: If *p* is omitted or is specified as 0, then *p* is set to 3.

Tip: If zero is the desired precision, use the *w.d* format in place of the D*w.p* format.

Details

The D*w.p* format writes numbers so that the decimal point aligns in groups of values with similar magnitude. Larger values of *p* print the data values with more precision and potentially more shifts in the decimal point alignment. Smaller values of *p* print the data values with less precision and a greater chance of decimal point alignment.

Comparisons

- The BEST*w*. format writes as many significant digits as possible in the output field, but if the numbers vary in magnitude, the decimal points do not line up.
- D*w.p* writes numbers with the desired precision and more alignment than the BEST*w* format.
- The BESTD*w.p* format is a combination of the BEST*w*. format and the D*w.p* format in that it formats all numeric data, and it does a better job of aligning decimals than the BEST*w*. format.
- The *w.d* format aligns decimal points, if possible, but it does not necessarily show the same precision for all numbers.

Examples

```
put @1 x d10.4;
```

Value of **x**	Results
	----+----1----+----2
12345	**12345.0**
1234.5	**1234.5**
123.45	**123.45000**
12.345	**12.34500**
1.2345	**1.23450**
.12345	**0.12345**

See Also

Format:

"BESTD*w.p* Format" on page 136

DATE*w*. Format

Writes date values in the form *ddmmmyy*, *ddmmmyyyy*, or *dd-mmm-yyyy*.

Category: Date and Time

Alignment: right

Syntax

DATE*w*.

Syntax Description

w

specifies the width of the output field.

Default: 7

Range: 5–11

Tip: Use a width of 9 to print a 4-digit year without a separator between the day, month, and year. Use a width of 11 to print a 4-digit year using a hyphen as a separator between the day, month, and year

Details

The DATE*w*. format writes SAS date values in the form *ddmmmyy*, *ddmmmyyyy*, or *dd-mmm-yyyy*, where

dd
is an integer that represents the day of the month.

mmm
is the first three letters of the month name.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

Examples

The example table uses the input value of 17607, which is the SAS date value that corresponds to March 16, 2008.

SAS Statement	Results
	`----+----1----+`
`put day date5.;`	`16MAR`
`put day date6.;`	` 16MAR`
`put day date7.;`	`16MAR08`
`put day date8.;`	` 16MAR08`
`put day date9.;`	`16MAR2008`
`put day date11.;`	`16-MAR-2008`

See Also

Function:
"DATE Function" on page 634
Informat:
"DATE*w*. Informat" on page 1322

DATEAMPM*w.d* Format

Writes datetime values in the form *ddmmmyy:hh:mm:ss.ss* with AM or PM.

Category: Date and Time

Alignment: right

Syntax

DATEAMPM*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 19

Range: 7–40

Tip: SAS requires a minimum *w* value of 13 to write AM or PM. For widths between 10 and 12, SAS writes a 24-hour clock time.

d

specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Requirement: must be less than *w*

Range: 0–39

Note: If *w*–*d*< 17, SAS truncates the decimal values. △

Details

The DATEAMPM*w.d* format writes SAS datetime values in the form *ddmmmyy:hh:mm:ss.ss*, where

dd
is an integer that represents the day of the month.

mmm
is the first three letters of the month name.

yy
is a two-digit integer that represents the year.

hh
is an integer that represents the hour.

mm
is an integer that represents the minutes.

ss.ss
is the number of seconds to two decimal places.

Comparisons

The DATEAMPM*w.d* format is similar to the DATETIME*w.d* format except that DATEAMPM*w.d* prints AM or PM at the end of the time.

Examples

The example table uses the input value of 1347455694, which is the SAS datetime value that corresponds to 11:01:34 a.m. on April 20, 2003.

SAS Statement	Results
	`----+----1----+----2----+`
`put event dateampm.;`	`20APR03:11:01:34 AM`
`put event dateampm7.;`	`20APR03`
`put event dateampm10.;`	`20APR:11`
`put event dateampm13.;`	`20APR03:11 AM`
`put event dateampm22.2;`	`20APR03:11:01:34.00 AM`

See Also

Format:

"DATETIME*w.d* Format" on page 154

DATETIME*w.d* Format

Writes datetime values in the form *ddmmmyy:hh:mm:ss.ss*.

Category: Date and Time
Alignment: right

Syntax

DATETIME*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 16

Range: 7–40

Tip: SAS requires a minimum *w* value of 16 to write a SAS datetime value with the date, hour, and seconds. Add an additional two places to *w* and a value to *d* to return values with optional decimal fractions of seconds.

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Requirement: must be less than *w*

Range: 0–39

Note: If *w*–*d*< 17, SAS truncates the decimal values. △

Details

The DATETIME*w.d* format writes SAS datetime values in the form *ddmmmyy:hh:mm:ss.ss*, where

dd
is an integer that represents the day of the month.

mmm
is the first three letters of the month name.

yy
is a two-digit integer that represents the year.

hh
is an integer that represents the hour in 24–hour clock time.

mm
is an integer that represents the minutes.

ss.ss
is the number of seconds to two decimal places.

Examples

The example table uses the input value of 1447213759, which is the SAS datetime value that corresponds to 3:49:19 a.m. on November 10, 2005.

SAS Statement	Results
	`----+----1----+----2`
`put event datetime.;`	`10NOV05:03:49:19`
`put event datetime7.;`	`10NOV05`
`put event datetime12.;`	`  10NOV05:03`
`put event datetime18.;`	`  10NOV05:03:49:19`
`put event datetime18.1;`	`10NOV05:03:49:19.0`
`put event datetime19.;`	` 10NOV2005:03:49:19`
`put event datetime20.1;`	`10NOV2005:03:49:19.0`
`put event datetime21.2;`	`10NOV2005:03:49:19.00`

See Also

Formats:

"DATE*w*. Format" on page 151
"TIME*w*.*d* Format" on page 245

Function:

"DATETIME Function" on page 637

Informats:

"DATE*w*. Informat" on page 1322
"DATETIME*w*. Informat" on page 1324
"TIME*w*. Informat" on page 1393

DAY*w*. Format

Writes date values as the day of the month.

Category: Date and Time
Alignment: right

Syntax

DAY*w*.

Syntax Description

w
specifies the width of the output field.

Default: 2
Range: 2–32

Examples

The example table uses the input value of 16601, which is the SAS date value that corresponds to June 14, 2005.

SAS Statement	Results
	`----+----1`
`put date day2.;`	`14`

DDMMYY*w*. Format

Writes date values in the form *ddmm<yy>yy* or *dd/mm/<yy>yy*, where a forward slash is the separator and the year appears as either 2 or 4 digits.

Category: Date and Time

Alignment: right

Syntax

DDMMYY*w*.

Syntax Description

w

specifies the width of the output field.

Default: 8

Range: 2–10

Interaction: When *w* has a value of from 2 to 5, the date appears with as much of the day and the month as possible. When *w* is 7, the date appears as a two-digit year without slashes.

Details

The DDMMYY*w*. format writes SAS date values in the form *ddmm<yy>yy* or *dd/mm/<yy>yy*, where

dd

is an integer that represents the day of the month.

/

is the separator.

mm

is an integer that represents the month.

<yy>yy

is a two-digit or four-digit integer that represents the year.

Examples

The following examples use the input value of 16794, which is the SAS date value that corresponds to December 24, 2005.

SAS Statement	Results
	`----+----1----+`
`put date ddmmyy5.;`	`24/12`
`put date ddmmyy6.;`	`241205`

SAS Statement	Results
`put date ddmmyy7.;`	`241205`
`put date ddmmyy8.;`	`24/12/05`
`put date ddmmyy10.;`	`24/12/2005`

See Also

Formats:

"DATE*w*. Format" on page 151
"DDMMYY*xw*. Format" on page 158
"MMDDYY*w*. Format" on page 196
"YYMMDD*w*. Format" on page 273

Function:

"MDY Function" on page 924

Informats:

"DATE*w*. Informat" on page 1322
"DDMMYY*w*. Informat" on page 1326
"MMDDYY*w*. Informat" on page 1349
"YYMMDD*w*. Informat" on page 1410

DDMMYY*xw*. Format

Writes date values in the form *ddmm<yy>yy* or *dd-mm-yy<yy>*, where the *x* in the format name is a character that represents the special character that separates the day, month, and year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.

Category: Date and Time
Alignment: right

Syntax

DDMMYY*xw.*

Syntax Description

x
identifies a separator or specifies that no separator appear between the day, the month, and the year. Valid values for *x* are:

B
separates with a blank

C
separates with a colon

D
separates with a dash

N
indicates no separator

P
separates with a period

S
separates with a slash.

w
specifies the width of the output field.

Default: 8

Range: 2–10

Interaction: When *w* has a value of from 2 to 5, the date appears with as much of the day and the month as possible. When *w* is 7, the date appears as a two-digit year without separators.

Interaction: When *x* has a value of N, the width range changes to 2–8.

Details

The DDMMYY*xw*. format writes SAS date values in the form *ddmm<yy>yy* or *ddxmmx<yy>yy*, where

dd
is an integer that represents the day of the month.

x
is a specified separator.

mm
is an integer that represents the month.

<yy>yy
is a two-digit or four-digit integer that represents the year.

Examples

The following examples use the input value of 18031, which is the SAS date value that corresponds to May 14, 2009.

SAS Statement	Results
	`----+----1----+`
`put date ddmmyyc5.;`	`14:05`
`put date ddmmyyd8.;`	`14-05-09`
`put date ddmmyyp10.;`	`14.05.2009`
`put date ddmmyyn8.;`	`14052009`

See Also

Formats:

"DATE*w*. Format" on page 151
"DDMMYY*w*. Format" on page 157
"MMDDYY*xw*. Format" on page 198
"YYMMDD*xw*. Format" on page 274

Functions:

"DAY Function" on page 637
"MDY Function" on page 924
"MONTH Function" on page 936
"YEAR Function" on page 1233

Informat:

"DDMMYY*w*. Informat" on page 1326

DOLLAR*w.d* Format

Writes numeric values with a leading dollar sign, a comma that separates every three digits, and a period that separates the decimal fraction.

Category: Numeric
Alignment: right

Syntax

DOLLAR*w*.*d*

Syntax Description

w
specifies the width of the output field.

Default: 6

Range: 2–32

d
specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Range: 0–31

Requirement: must be less than *w*

Details

The DOLLAR*w*.*d* format writes numeric values with a leading dollar sign, a comma that separates every three digits, and a period that separates the decimal fraction.

The hexadecimal representation of the code for the dollar sign character ($) is 5B on EBCDIC systems and 24 on ASCII systems. The monetary character that these codes represent might be different in other countries, but DOLLAR*w.d* always produces one of these codes. If you need another monetary character, define your own format with the FORMAT procedure. See "The FORMAT Procedure" in *Base SAS Procedures Guide* for more details.

Comparisons

- The DOLLAR*w.d* format is similar to the DOLLARX*w.d* format, but the DOLLARX*w.d* format reverses the roles of the decimal point and the comma. This convention is common in European countries.
- The DOLLAR*w.d* format is the same as the COMMA*w.d* format except that the COMMA*w.d* format does not write a leading dollar sign.

Examples

```
put @3 netpay dollar10.2;
```

Value of **netpay**	Results
	`----+----1----+`
`1254.71`	`$1,254.71`

See Also

Formats:

"COMMA*w.d* Format" on page 147
"DOLLARX*w.d* Format" on page 161

DOLLARX*w.d* Format

Writes numeric values with a leading dollar sign, a period that separates every three digits, and a comma that separates the decimal fraction.

Category: Numeric
Alignment: right

Syntax

DOLLARX*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 6

Range: 2–32

d
specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Default: 0

Range: 0–31

Requirement: must be less than *w*

Details

The DOLLARX*w.d* format writes numeric values with a leading dollar sign, with a period that separates every three digits, and with a comma that separates the decimal fraction.

The hexadecimal representation of the code for the dollar sign character ($) is 5B on EBCDIC systems and 24 on ASCII systems. The monetary character that these codes represent might be different in other countries, but DOLLARX*w.d* always produces one of these codes. If you need another monetary character, define your own format with the FORMAT procedure. See "The FORMAT Procedure" in *Base SAS Procedures Guide* for more details.

Comparisons

- The DOLLARX*w.d* format is similar to the DOLLAR*w.d* format, but the DOLLARX*w.d* format reverses the roles of the decimal point and the comma. This convention is common in European countries.
- The DOLLARX*w.d* format is the same as the COMMAX*w.d* format except that the COMMA*w.d* format does not write a leading dollar sign.

Examples

```
put @3 netpay dollarx10.2;
```

Value of **netpay**	Results
	`----+----1----+`
`1254.71`	`$1.254,71`

See Also

Formats:

"COMMAX*w.d* Format" on page 148

"DOLLAR*w.d* Format" on page 160

DOWNAME*w*. Format

Writes date values as the name of the day of the week.

Category: Date and Time
Alignment: right

Syntax

DOWNAME*w*.

Syntax Description

w
specifies the width of the output field.
Default: 9
Range: 1–32
Tip: If you omit *w*, SAS prints the entire name of the day.

Details

If necessary, SAS truncates the name of the day to fit the format width. For example, the DOWNAME2. prints the first two letters of the day name.

Examples

The example table uses the input value of 13589, which is the SAS date value that corresponds to March 16, 1997.

SAS Statement	Results
	`----+----1`
`put date downame.;`	`Sunday`

See Also

Format:
"WEEKDAY*w*. Format" on page 258

DTDATE*w*. Format

Expects a datetime value as input and writes date values in the form *ddmmmyy* or *ddmmmyyyy*.

Category: Date and Time
Alignment: right

Syntax

DTDATE*w*.

Syntax Description

w
specifies the width of the output field.
Default: 7
Range: 5–9
Tip: Use a width of 9 to print a 4–digit year.

Details

The DTDATE*w*. format writes SAS date values in the form *ddmmmyy* or *ddmmmyyyy*, where

dd
is an integer that represents the day of the month.

mmm
are the first three letters of the month name.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

Comparisons

The DTDATE*w*. format produces the same type of output that the DATE*w*. format produces. The difference is that the DTDATE*w*. format requires a datetime value.

Examples

The example table uses a datetime value of 16APR2000:10:00:00 as input, and prints both a two-digit and a four-digit year for the DTDATE*w*. format.

SAS Statement	Results
	`----+----+`
`put trip_date=dtdate.;`	`16APR00`
`put trip_date=dtdate9.;`	`16APR2000`

See Also

Formats:

"DATE*w*. Format" on page 151

DTMONYY*w*. Format

Writes the date part of a datetime value as the month and year in the form *mmmyy* or *mmmyyyy*.

Category: Date and Time

Alignment: right

Syntax

DTMONYY*w*.

Syntax Description

w

specifies the width of the output field.

Default: 5

Range: 5–7

Details

The DTMONYY*w*. format writes SAS datetime values in the form *mmmyy* or *mmmyyyy*, where

mmm

is the first three letters of the month name.

yy or *yyyy*

is a two–digit or four-digit integer that represents the year.

Comparisons

The DTMONYY*w*. format and the MONYY*w*. format are similar in that they both write date values. The difference is that DTMONYY*w*. expects a datetime value as input, and MONYY*w*. expects a SAS date value.

Examples

The example table uses as input the value 1476598132, which is the SAS datetime value that corresponds to October 16, 2006, at 06:08:52 a.m.

SAS Statement	Results
	`----+----1`
`put date dtmonyy.;`	`OCT06`
`put date dtmonyy5.;`	`OCT06`
`put date dtmonyy6.;`	` OCT06`
`put date dtmonyy7.;`	`OCT2006`

See Also

Formats:

"DATETIME*w*.*d* Format" on page 154
"MONYY*w*. Format" on page 206

DTWKDATX*w*. Format

Writes the date part of a datetime value as the day of the week and the date in the form *day-of-week*, *dd month-name yy* (or *yyyy*).

Category: Date and Time
Alignment: right

Syntax

DTWKDATX*w*.

Syntax Description

w
specifies the width of the output field.
Default: 29
Range: 3–37

Details

The DTWKDATX*w*. format writes SAS date values in the form *day-of-week*, *dd month-name*, *yy* or *yyyy*, where

day-of-week
is either the first three letters of the day name or the entire day name.

dd
is an integer that represents the day of the month.

month-name
is either the first three letters of the month name or the entire month name.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

Comparisons

The DTWKDATX*w*. format is similar to the WEEKDATX*w*. format in that they both write date values. The difference is that DTWKDATX*w*. expects a datetime value as input, and WEEKDATX*w*. expects a SAS date value.

Examples

The example table uses as input the value 1476598132, which is the SAS datetime value that corresponds to October 16, 2002, at 06:08:52 a.m.

SAS Statement	Results
	`----+----1----+----2----+----3`
`put date dtwkdatx.;`	`Monday, 16 October 2006`
`put date dtwkdatx3.;`	`Mon`
`put date dtwkdatx8.;`	`Mon`
`put date dtwkdatx25.;`	`Monday, 16 Oct 2006`

See Also

Formats:

"DATETIME*w*.*d* Format" on page 154
"WEEKDATX*w*. Format" on page 257

DTYEAR*w*. Format

Writes the date part of a datetime value as the year in the form *yy* or *yyyy*.

Category: Date and Time

Alignment: right

Syntax

DTYEAR*w*.

Syntax Description

w

specifies the width of the output field.

Default: 4

Range: 2–4

Comparisons

The DTYEAR*w*. format is similar to the YEAR*w*. format in that they both write date values. The difference is that DTYEAR*w*. expects a datetime value as input, and YEAR*w*. expects a SAS date value.

Examples

The example table uses as input the value 1476598132, which is the SAS datetime value that corresponds to October 16, 2006, at 06:08:52 a.m.

SAS Statement	Results
	`----+----1`
`put date dtyear.;`	`2006`
`put date dtyear2.;`	`06`
`put date dtyear3.;`	` 06`
`put date year4.;`	`2006`

See Also

Formats:

"DATETIME*w*.*d* Format" on page 154

"YEAR*w*. Format" on page 269

DTYYQC*w*. Format

Writes the date part of a datetime value as the year and the quarter and separates them with a colon (:).

Category: Date and Time

Alignment: right

Syntax

DTYYQC*w*.

Syntax Description

w
specifies the width of the output field.

Default: 4

Range: 4–6

Details

The DTYYQC*w*. format writes SAS datetime values in the form *yy* or *yyyy*, followed by a colon (:) and the numeric value for the quarter of the year.

Examples

The example table uses as input the value 1476598132, which is the SAS datetime value that corresponds to October 16, 2006, at 06:08:52 p.m..

SAS Statement	Results
	`----+----1`
`put date dtyyqc.;`	`06:4`
`put date dtyyqc4.;`	`06:4`
`put date dtyyqc5.;`	` 06:4`
`put date dtyyqc6.;`	`2006:4`

See Also

Formats:

"DATETIME*w*.*d* Format" on page 154

E*w*. Format

Writes numeric values in scientific notation.

Category: Numeric

Alignment: right

See: E*w*. Format in the documentation for your operating environment.

Syntax

E*w*.

Syntax Description

w

specifies the width of the output field. The output field can dispaly up to fourteen significant digits.

Default: 12

Range: 7–32

Details

When formatting values in scientific notation, the E format reserves the first column of the result for a minus sign and formats up to fourteen significant digits.

Examples

```
put @1 x e10.;
```

Value of **x**	Results
	----+----1----+
1257	**1.257E+03**
-1257	**-1.257E+03**

E8601DA*w*. Format

Writes date values using the ISO 8601 extended notation *yyyy-mm-dd*.

Category: Date and Time
ISO 8601
Alignment: left
Alias: IS8601DA
Time Zone Format: No
ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

E8601DA*w*.

Syntax Description

w
specifies the width of the output field.
Default: 10
Requirement: The width of the output field must be 10.

Details

The E8601DA format writes a date in the ISO 8601 extended notation *yyyy-mm-dd*:

yyyy	is a four-digit year, such as 2008.
mm	is a two-digit month (zero padded) between 01 and 12.
dd	is a two-digit day of the month (zero padded) between 01 and 31.

Examples

```
put eda e8601da.;
```

Value for eda	Results
17790	2008-09-15

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

E8601DN*w*. Format

Writes the date from a SAS datetime value using the ISO 8601 extended notation *yyyy-mm-dd*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601DN

Time Zone Format: No

ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

E8601DN*w*.

Syntax Description

w
specifies the width of the input field.

Default: 10

Requirement: The width of the input field must be 10.

Details

The E8601DN formats writes the date in the ISO 8601 extended date notation *yyyy-mm-dd*:

yyyy	is a four-digit year, such as 2008.
mm	is a two-digit month (zero padded) between 01 and 12.
dd	is a two-digit day of the month (zero padded) between 01 and 31.

Examples

```
put edn e8601dn.;
```

Value for edn	Results
1537113180	2008-09-15

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

E8601DT*w.d* Format

Writes datetime values in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601DT

Time Zone Format: No

ISO 8601 Element: 5.4.1 Complete representation

Syntax

E8601DT*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 19

Range: 19 - 26

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0 - 6

Details

The E8602DT format writes datetime values using the ISO 8601 extended datetime notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*:

yyyy	is a four-digit year, such as 2008.
mm	is a two-digit month (zero padded) between 01 and 12.
dd	is a two-digit day of the month (zero padded) between 01 and 31.
hh	is a two-digit hour (zero padded), between 00 - 23.
mm	is a two-digit minute (zero padded), between 00 - 59.
ss	is a two-digit second (zero padded), between 00 - 59.
.ffffff	are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 - 9.

Examples

```
put edt e8601dt.;
```

Value of edt	Results
1537113180	2008-09-15T15:53:00

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

E8601DZ*w.* Format

Writes datetime values in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime and time zone extended notations *yyyy-mm-dd*T*hh:mm:ss*+|-*hh*:*mm*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601DZ

Time Zone Format: Yes

ISO 8601 Element: 5.4.1 Complete representation

Syntax

E8601DZ*w.*

Syntax Description

w
specifies the width of the output field.

Default: 26

Range: 20 - 35

Details

UTC values specify a time and a time zone based on the zero meridian in Greenwich, England. The E8602DZ format writes SAS datetime values using one of the following ISO 8601 extended datetime notations:

*yyyy-mm-dd*T*hh:mm:ss*+|–*hh:mm*	is the form used when *w* is large enough to support this time zone notation.
*yyyy-mm-dd*T*hh:mm:ss*Z	is the form used when *w* is not large enough to support the +|-*hhmm* time zone notation.

where

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 31
hh	is a two-digit hour (zero padded), between 00 - 24
mm	is a two-digit minute (zero padded), between 00 - 59
ss	is a two-digit second (zero padded), between 00 - 59
Z	indicates that the time is for zero meridian (Greenwich, England) or UTC time.
+|-*hh:mm*	is an hour and minute signed offset from zero meridian time. Note that the offset must be +|-*hh:mm* (that is, + or - and five characters). Use + for time zones east of the zero meridian and use - for time zones west of the zero meridian. For example, +02:00 indicates a two hour time difference to the east of the zero meridian, and -06:00 indicates a six hour time differences to the west of the zero meridian. **Restriction:** The shorter form +|-*hh* is not supported.

Examples

```
put edz e8601dz.;
```

Value of edz	Results
1537113180	2008-09-15T15:53:00+00:00
1537102380	2008-09-15T12:53:00+00:00

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

E8601LZ*w.* Format

Writes time values as local time, appending the Coordinated Universal Time (UTC) offset for the local SAS session, using the ISO 8601 extended time notation *hh:mm:ss*+|-*hh:mm*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601LZ

Time Zone Format: Yes. The format appends the UTC offset to the value as determined by the local SAS session.

ISO 8601 Element: 5.3.1.1 Complete representation

Syntax

E8601LZ*w.*

Syntax Description

w
specifies the width of the output field.

Default: 14

Range: 9 - 20

Details

The E8602LZ format writes time values without making any adjustments and appends the UTC time zone offset for the local SAS session, using one of the following ISO 8601 extended time notations:

hh:mm:ss+|–*hh:mm*	is the form used when *w* is large enough to support this time notation.
*hh:mm:ss*Z	is the form used when *w* is not large enough to support the +|-*hh:mm* time zone notation.

where

hh	is a two-digit hour (zero padded), between 00 - 23.
mm	is a two-digit minute (zero padded), between 00 - 59.
ss	is a two-digit second (zero padded), between 00 - 59.
Z	indicate zero meridian (Greenwich, England) or UTC time.
+|-*hh:mm*	is an hour and minute signed offset from zero meridian time. Note that the offset must be +|-*hh:mm* (that is, + or - and five characters). Use + for time zones east of the zero meridian and use - for time zones west of the zero meridian. For example, +02:00 indicates a two hour time difference to the east of the zero meridian, and -06:00 indicates a six hour time differences to the west of the zero meridian. **Restriction:** The shorter form +|-*hh* is not supported.

SAS writes the time value using the form *hh:mm.ffffff* and appends the time zone indicator +|-*hh:mm* based on the time zone offset from the zero meridian for the local SAS session, or Z. The Z time zone indicator is used for format lengths that are less than 14.

If the same time is written using both zone indicators, they indicate two different times based on the UTC. For example, if the local SAS session uses Eastern Standard

Time in the US, and the time value is 45824, SAS would write 12:43:44-04:00 or 12:43:44Z. The time 12:43:44–04:00 is the time 16:43:44+00:00 at the zero meridian. The Z indicates that the time is the time at the zero meridian, or 12:43:44+00:00.

When SAS reads a UTC time by using the E8601TZ informat, and the adjusted time is greater than 24 hours or less than 00 hours, SAS adjusts the value so that the time is between 00:00:00 and 24:00:00. If the E8601TZ format attempts to format a time outside of this time range, the time is formatted with stars to indicate that the value is out of range.

Examples

The following PUT statement write the time for the Eastern Standard time zone.

```
put elz e8601lz.;
```

Value of elz	Results
`46380`	`12:53:00-5:00`

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

E8601TM*w.d* Format

Writes time values using the ISO 8601 extended notation *hh:mm:ss.ffffff*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601TM

Time Zone Format: No

ISO 8601 Element: 5.3.1.1 Complete representation and 5.3.1.3 Representation of decimal fractions

Syntax

E8601TM*w*.*d*

Syntax Description

w
specifies the width of the output field.

Default: 8

Range: 8 - 15

d

specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0 - 6

Details

The E8601TM format writes SAS time values using the following ISO 8601 extended time notation:

hh:mm:ss.ffffff

hh	is a two-digit hour (zero padded), between 00 - 23.
mm	is a two-digit minute (zero padded), between 00 - 59.
ss	is a two-digit second (zero padded), between 00 - 59.
.ffffff	are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 - 9.

Examples

```
put etm e8601tm.;
```

Value of etm	Results
57180	15:53:00

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

E8601TZ*w.d* Format

Adjusts time values to the Coordinated Universal Time (UTC) and writes them using the ISO 8601 extended notation *hh*:*mm*:*ss*+|-*hh*:*mm*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601TZ

Time Zone Format: Yes

ISO 8601 Element: 5.3.1.1 Complete representation

Syntax

E8601TZ*w*.*d*

Syntax Description

w
specifies the width of the output field.

Default: 14

Range: 9 - 20

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0 - 6

Details

UTC time values specify a time and a time zone based on the zero meridian in Greenwich, England. The E8602TZ format writes time values in one of the following ISO 8601 extended time notations:

hh:mm:ss+|–*hh:mm*	is the form used when *w* is large enough to support this time zone notation.
*hh:mm:ss*Z	is the form used when *w* is not large enough to support the +|-*hh:mm* time zone notation.

where

hh	is a two-digit hour (zero padded), between 00 - 23
mm	is a two-digit minute (zero padded), between 00 - 59
ss	is a two-digit second (zero padded), between 00 - 59
Z	indicate zero meridian (Greenwich, England) or UTC time
+|-*hh:mm*	is an hour and minute signed offset from zero meridian time. Note that the offset must be +|-*hh:mm* (that is, + or - and five characters). The shorter form +|-*hh* is not supported. Use + for time zones east of the zero meridian and use - for time zones west of the zero meridian. For example, +02:00 indicates a two hour time difference to the east of the zero meridian, and -06:00 indicates a six hour time differences to the west of the zero meridian.

When SAS reads a UTC time by using the B8601TZ informat, and the adjusted time is greater than 24 hours or less than 00 hours, SAS adjusts the value so that the time is between 00:00:00 and 24:00:00. If the E8601TZ format attempts to format a time outside of this time range, the time is formatted with stars to indicate that the value is out of range.

Comparisons

For time values between 00:00:00 and 24:00:00, the E8601TZ format adjusts the time value to be the time at the zero meridian and writes it in the international standard extended time notation. The E8601LZ format makes no adjustment to the time and writes time values in the international standard extended time notation, using a UTC time zone offset for the local SAS session.

Examples

```
put etz e8601tz.;
```

Value of etz	Results
`73441`	`20:24:01+00:00`
`62641`	`17:24:01+00:00`

See Also

"Working with Dates and Times Using the ISO 8601 Basic and Extended Notations" on page 94

FLOAT*w.d* Format

Generates a native single-precision, floating-point value by multiplying a number by 10 raised to the *d*th power.

Category: Numeric
Alignment: left

Syntax

FLOAT*w.d*

Syntax Description

w
specifies the width of the output field.
Requirement: width must be 4

d
specifies the power of 10 by which to multiply the value. This argument is optional.
Default: 0
Range: 0-31

Details

This format is useful in operating environments where a float value is not the same as a truncated double. Values that are written by FLOAT4. typically are values that are meant to be read by some other external program that runs in your operating environment and that expects these single-precision values.

Note: If the value that is to be formatted is a missing value, or if it is out-of-range for a native single-precision, floating-point value, a single-precision value of zero is generated. △

On IBM mainframe systems, a four-byte floating-point number is the same as a truncated eight-byte floating-point number. However, in operating environments using the IEEE floating-point standard, such as IBM PC-based operating environments and most UNIX operating environments, a four-byte floating-point number is not the same as a truncated double. Hence, the RB4. format does not produce the same results as the FLOAT4. format. Floating-point representations other than IEEE might have this same characteristic.

Comparisons

The following table compares the names of float notation in several programming languages:

Language	Float Notation
SAS	FLOAT4
Fortran	REAL+4
C	float
IBM 370 ASM	E
PL/I	FLOAT BIN(21)

Examples

```
put x float4.;
```

Value of **x**	Results*
1	**3F800000**

* The result is a hexadecimal representation of a binary number that is stored in IEEE form.

FRACT*w*. Format

Converts numeric values to fractions.

Category: Numeric
Alignment: right

Syntax

FRACT*w*.

Syntax Description

w
specifies the width of the output field.
Default: 10
Range: 4–32

Details

Dividing the number 1 by 3 produces the value 0.33333333. To write this value as 1/3, use the FRACT*w*. format. FRACT*w*. writes fractions in reduced form, that is, 1/2 instead of 50/100.

Examples

```
put x fract8.;
```

Value of **x**	Results
	`----+----1`
`0.6666666667`	`2/3`
`0.2784`	`174/625`

HEX*w*. Format

Converts real binary (floating-point) values to hexadecimal representation.

Category: Numeric

Alignment: left

See: HEX*w*. Format in the documentation for your operating environment.

Syntax

HEX*w*.

Syntax Description

w

specifies the width of the output field.

Default: 8

Range: 1–16

Tip: If *w*< 16, the HEX*w*. format converts real binary numbers to fixed-point integers before writing them as hexadecimal characters. It also writes negative numbers in two's complement notation, and right aligns digits. If *w* is 16, HEX*w*. displays floating-point values in their hexadecimal form.

Details

In any operating environment, the least significant byte written by HEX*w*. is the rightmost byte. Some operating environments store integers with the least significant digit as the first byte. The HEX*w*. format produces consistent results in any operating environment regardless of the order of significance by byte.

Note: Different operating environments store floating-point values in different ways. However, the HEX16. format writes hexadecimal representations of floating-point values with consistent results in the same way that your operating environment stores them. △

Comparisons

The HEX*w*. numeric format and the $HEX*w*. character format both generate the hexadecimal equivalent of values.

Examples

```
put @8 x hex8.;
```

Value of **x**	Results
	----+----1----+----2
35.4	**00000023**
88	**00000058**
2.33	**00000002**
-150	**FFFFFF6A**

HHMM*w.d* Format

Writes time values as hours and minutes in the form *hh:mm*.

Category: Date and Time
Alignment: right

Syntax

HHMM*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 5

Range: 2–20

d
specifies the number of digits to the right of the decimal point in the minutes value. The digits to the right of the decimal point specify a fraction of a minute. This argument is optional.

Default: 0

Range: 0–19

Requirement: must be less than *w*

Details

The HHMM*w.d* format writes SAS time values in the form *hh:mm*, where

hh
is an integer.

Note: If *hh* is a single digit, HHMM*w.d* places a leading blank before the digit. For example, the HHMM*w.d.* format writes 9:00 instead of 09:00. △

mm
is the number of minutes that range from 00 through 59.

SAS rounds hours and minutes that are based on the value of seconds in a SAS time value.

The HHMM format uses asterisks to format values that are outside the time range 0–24 hours, such as datetime values.

Comparisons

The HHMM*w.d* format is similar to the TIME*w.d* format except that the HHMM*w.d* format does not print seconds.

The HHMM*w.d* format writes a leading blank for a single-hour digit. The TOD*w.d* format writes a leading zero for a single-hour digit.

Examples

The example table uses the input value of 46796, which is the SAS time value that corresponds to 12:59:56 p. m.

SAS Statement	Results
	`----+----1`
`put time hhmm.;`	`13:00`
`put time hhmm8.2;`	`12:59.93`

In the first example, SAS rounds up the time value four seconds based on the value of seconds in the SAS time value. In the second example, by adding a decimal specification of 2 to the format shows that fifty-six seconds is 93% of a minute.

See Also

Formats:

- "HOUR*w.d* Format" on page 188
- "MMSS*w.d* Format" on page 200
- "TIME*w.d* Format" on page 245
- "TOD*w.d* Format" on page 249

Functions:

- "HMS Function" on page 803
- "HOUR Function" on page 807
- "MINUTE Function" on page 928
- "SECOND Function" on page 1122
- "TIME Function" on page 1161

Informat:

- "TIME*w*. Informat" on page 1393

HOUR*w.d* Format

Writes time values as hours and decimal fractions of hours.

Category: Date and Time
Alignment: right

Syntax

HOUR*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 2

Range: 2–20

d
specifies the number of digits to the right of the decimal point in the hour value. Therefore, SAS prints decimal fractions of the hour. This argument is optional.

Requirement: must be less than *w*

Range: 0-19

Details

SAS rounds hours based on the value of minutes in the SAS time value.

The HOUR format uses asterisks to format values that are outside the time range 0–24 hours, such as datetime values.

Examples

The example table uses the input value of 41400, which is the SAS time value that corresponds to 11:30 a.m.

SAS Statement	Results
	`----+----1`
`put time hour4.1;`	`11.5`

See Also

Formats:

"HHMM*w.d* Format" on page 185

"MMSS*w.d* Format" on page 200
"TIME*w.d* Format" on page 245
"TOD*w.d* Format" on page 249
Functions:
"HMS Function" on page 803
"HOUR Function" on page 807
"MINUTE Function" on page 928
"SECOND Function" on page 1122
"TIME Function" on page 1161
Informat:
"TIME*w*. Informat" on page 1393

IB*w.d* Format

Writes native integer binary (fixed-point) values, including negative values.

Category: Numeric
Alignment: left
See: IB*w.d* Format in the documentation for your operating environment.

Syntax

IB*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 4
Range: 1–8

d
specifies to multiply the number by 10^d. This argument is optional.
Default 0
Range: 0–10

Details

The IB*w.d* format writes integer binary (fixed-point) values, including negative values that are represented in two's complement notation. IB*w.d* writes integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 88. △

Comparisons

The IB*w.d* and PIB*w.d* formats are used to write native format integers. (Native format allows you to read and write values created in the same operating environment.) The IBR*w.d* and PIBR*w.d* formats are used to write little endian integers in any operating environment.

To view a table that shows the type of format to use with big endian and little endian integers, see Table 3.1 on page 88.

To view a table that compares integer binary notation in several programming languages, see Table 3.2 on page 89.

Examples

```
y=put(x,ib4.);
put y $hex8.;
```

Value of **x**	Results on Big Endian Platforms*	Results on Little Endian Platforms*
	----+----1	**----+----1**
128	**00000080**	**80000000**

* The result is a hexadecimal representation of a four-byte integer binary number. Each byte occupies one column of the output field.

See Also

Format:

"IBR*w.d* Format" on page 190

IBR*w.d* Format

Writes integer binary (fixed-point) values in Intel and DEC formats.

Category: Numeric
Alignment: left

Syntax

IBR*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 4

Range: 1–8

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–10

Details

The IBR*w.d* format writes integer binary (fixed-point) values, including negative values that are represented in two's complement notation. IBR*w.d* writes integer binary values that are generated by and for Intel and DEC operating environments. Use IBR*w.d* to write integer binary data from Intel or DEC environments on other operating environments. The IBR*w.d* format in SAS code allows for a portable implementation for writing the data in any operating environment.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 88. △

Comparisons

- The IB*w.d* and PIB*w.d* formats are used to write native format integers. (Native format allows you to read and write values that are created in the same operating environment.)
- The IBR*w.d* and PIBR*w.d* formats are used to write little endian integers, regardless of the operating environment you are writing on.
- In Intel and DEC operating environments, the IB*w.d* and IBR*w.d* formats are equivalent.

To view a table that shows the type of format to use with big endian and little endian integers, see Table 3.1 on page 88.

To view a table that compares integer binary notation in several programming languages, see Table 3.2 on page 89.

Examples

```
y=put(x,ibr4.);
put y $hex8.;
```

Value of **x**	Results
	----+----1
128	**80000000**

* The result is a hexadecimal representation of a 4-byte integer binary number. Each byte occupies one column of the output field.

See Also

Format:

"IB*w.d* Format" on page 189

IEEE*w.d* Format

Generates an IEEE floating-point value by multiplying a number by 10 raised to the *d*th power.

Category: Numeric

Alignment: left

Caution: Large floating-point values and floating-point values that require precision might not be identical to the original SAS value when they are written to an IBM mainframe by using the IEEE format and read back into SAS using the IEE informat.

Syntax

IEEE*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 8

Range: 3–8

Tip: If *w* is 8, an IEEE double-precision, floating-point number is written. If *w* is 5, 6, or 7, an IEEE double-precision, floating-point number is written, which assumes truncation of the appropriate number of bytes. If *w* is 4, an IEEE single-precision floating-point number is written. If *w* is 3, an IEEE single-precision, floating-point number is written, which assumes truncation of one byte.

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–10

Details

This format is useful in operating environments where IEEE*w.d* is the floating-point representation that is used. In addition, you can use the IEEE*w.d* format to create files that are used by programs in operating environments that use the IEEE floating-point representation.

Typically, programs generate IEEE values in single-precision (4 bytes) or double-precision (8 bytes). Programs perform truncation solely to save space on output files. Machine instructions require that the floating-point number be one of the two lengths. The IEEE*w.d* format allows other lengths, which enables you to write data to files that contain space-saving truncated data.

Examples

```
test1=put(x,ieee4.);
put test1 $hex8.;

test2=put(x,ieee5.);
put test2 $hex10.;
```

Value of **x**	Results
1	**3F800000**
	3FF0000000

* The result contains hexadecimal representations of binary numbers stored in IEEE form.

JULDAY*w*. Format

Writes date values as the Julian day of the year.

Category: Date and Time

Alignment: right

Syntax

JULDAY*w*.

Syntax Description

w
specifies the width of the output field.

Default: 3

Range: 3–32

Details

The JULDAY*w*. format writes SAS date values in the form *ddd*, where

ddd
is the number of the day, 1–365 (or 1–366 for leap years).

Examples

The example table uses the input values of 13515, which is the SAS date value that corresponds to January 1, 1997, and 13589, which is the SAS date value that corresponds to March 16, 1997.

SAS Statement	Results
	----+----1
put date julday3.;	**1**
put date julday3.;	**75**

JULIAN*w*. Format

Writes date values as Julian dates in the form *yyddd* or *yyyyddd*.

Category: Date and Time
Alignment: left

Syntax

JULIAN*w*.

Syntax Description

w
specifies the width of the output field.

Default: 5

Range: 5–7

Tip: If *w* is 5, the JULIAN*w*. format writes the date with a two-digit year. If *w* is 7, the JULIAN*w*. format writes the date with a four-digit year.

Details

The JULIAN*w*. format writes SAS date values in the form *yyddd* or *yyyyddd*, where

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

ddd
is the number of the day, 1–365 (or 1–366 for leap years), in that year.

Examples

The example table uses the input value of 16794, which is the SAS date value that corresponds to December 24, 2005 (the 358th day of the year).

SAS Statement	Results
	`----+----1`
`put date julian5.;`	`05358`
`put date julian7.;`	`2005358`

See Also

Functions:

"DATEJUL Function" on page 635
"JULDATE Function" on page 866

Informat:

"JULIAN*w*. Informat" on page 1346

MDYAMPM*w.d* Format

Writes datetime values in the form *mm/dd/yy<yy> hh:mm* AM|PM. The year can be either two or four digits.

Category: Date and Time
Alignment: right
Default Time Period: AM

Syntax

MDYAMPM*w*.

Syntax Description

w
specifies the width of the output field.
Default: 19
Range: 8–40

Details

The MDYAMPM*w.d* format writes SAS datetime values in the following form:

mm/*dd*/*yy*<yy> *hh*:*mm*<AM | PM>

The following list explains the datetime variables:

mm
is an integer from 1 through 12 that represents the month.

dd
is an integer from 1 through 31 that represents the day of the month.

yy or *yyyy*
specifies a two-digit or four-digit integer that represents the year.

hh
is the number of hours that range from 0 through 23.

mm
is the number of minutes that range from 00 through 59.

AM | PM
specifies either the time period 00:01–12:00 noon (AM) or the time period 12:01–12:00 midnight (PM). The default is AM.

date and time separator characters
is one of several special characters, such as the slash (/), colon (:), or a blank character that SAS uses to separate date and time components.

Comparison

The MDYAMPM*w*. format writes datetime values with separators in the form *mm*/*dd*/*yy*<*yy*> *hh:mm* AM | PM, and requires a space between the date and the time.

The DATETIME*w.d* format writes datetime values with separators in the form *ddmmmyy*<*yy*>*: hh:mm:ss.ss*.

Examples

This example uses the input value of 1537113180, which is the SAS datetime value that corresponds to 3:53:00 PM on September 15, 2008.

SAS Statement	Results
`put dt mdyampm25.`	`9/15/2008 3:53 PM`

See Also

Format:

"DATETIME*w.d* Format" on page 154

Informat:

"MDYAMPM*w.d* Informat" on page 1347

MMDDYY*w*. Format

Writes date values in the form *mmdd*<*yy*>*yy* or *mm*/*dd*/<*yy*>*yy*, where a forward slash is the separator and the year appears as either 2 or 4 digits.

Category: Date and Time

Alignment: right

Syntax

MMDDYY*w*.

Syntax Description

w

specifies the width of the output field.

Default: 8

Range: 2–10

Interaction: When *w* has a value of from 2 to 5, the date appears with as much of the month and the day as possible. When *w* is 7, the date appears as a two-digit year without slashes.

Details

The MMDDYY*w*. format writes SAS date values in the form *mmdd*<*yy*>*yy* or *mm*/*dd*/<*yy*>*yy*, where

mm
is an integer that represents the month.

/
is the separator.

dd
is an integer that represents the day of the month.

<*yy*>*yy*
is a two-digit or four-digit integer that represents the year.

Examples

The following examples use the input value of 16734, which is the SAS date value that corresponds to October 25, 2005.

SAS Statement	Results
	`----+----1----+`
`put day mmddyy2.;`	`10`
`put day mmddyy3.;`	`10`
`put day mmddyy4.;`	`1025`
`put day mmddyy5.;`	`10/25`
`put day mmddyy6.;`	`102505`
`put day mmddyy7.;`	`102505`
`put day mmddyy8.;`	`10/25/05`
`put day mmddyy10.;`	`10/25/2005`

See Also

Formats:

"DATE*w*. Format" on page 151
"DDMMYY*w*. Format" on page 157
"MMDDYY*x*w. Format" on page 198
"YYMMDD*w*. Format" on page 273

Functions:

"DAY Function" on page 637
"MDY Function" on page 924
"MONTH Function" on page 936
"YEAR Function" on page 1233

Informats:

"DATE*w*. Informat" on page 1322
"DDMMYY*w*. Informat" on page 1326
"YYMMDD*w*. Informat" on page 1410

MMDDYY*xw*. Format

Writes date values in the form *mmdd<yy>yy* or *mm-dd-<yy>yy*, where the *x* in the format name is a character that represents the special character which separates the month, day, and year. The special character can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.

Category: Date and Time
Alignment: right

Syntax

MMDDYY*xw.*

Syntax Description

x
identifies a separator or specifies that no separator appear between the month, the day, and the year. Valid values for *x* are:

B
separates with a blank

C
separates with a colon

D
separates with a dash

N
indicates no separator

P
separates with a period

S
separates with a slash.

w
specifies the width of the output field.

Default: 8

Range: 2–10

Interaction: When *w* has a value of from 2 to 5, the date appears with as much of the month and the day as possible. When *w* is 7, the date appears as a two-digit year without separators.

Interaction: When *x* has a value of N, the width range changes to 2–8.

Details

The MMDDYY*xw*. format writes SAS date values in the form *mmdd<yy>yy* or *mmxddx<yy>yy*, where

mm
is an integer that represents the month.

x
is a specified separator.

dd
is an integer that represents the day of the month.

<yy>yy
is a two-digit or four-digit integer that represents the year.

Examples

The following examples use the input value of 18031, which is the SAS date value that corresponds to May 14, 2009.

SAS Statement	Results
	`----+----1----+`
`put day mmddyyc5.;`	`05:14`
`put day mmddyyd8.;`	`05-14-09`
`put day mmddyyp10.;`	`05.14.2009`
`put day mmddyyn8.;`	`05142009`

See Also

Formats:

- "DATE*w*. Format" on page 151
- "DDMMYY*xw*. Format" on page 158
- "MMDDYY*w*. Format" on page 196
- "YYMMDD*xw*. Format" on page 274

Functions:

- "DAY Function" on page 637
- "MDY Function" on page 924
- "MONTH Function" on page 936
- "YEAR Function" on page 1233

Informat:

- "MMDDYY*w*. Informat" on page 1349

MMSS*w.d* Format

Writes time values as the number of minutes and seconds since midnight.

Category: Date and Time
Alignment: right

Syntax

MMSS*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 5

Range: 2–20

Tip: Set *w* to a minimum of 5 to write a value that represents minutes and seconds.

d
specifies the number of digits to the right of the decimal point in the seconds value. Therefore, the SAS time value includes fractional seconds. This argument is optional.

Range: 0–19

Restriction: must be less than *w*

Details

The MMSS format uses asterisks to format values that are outside the time range 0–24 hours, such as datetime values.

Examples

The example table uses the input value of 4530.

SAS Statement	Results
	`----+----1`
`put time mmss.;`	`75:30`

See Also

Formats:

"HHMM*w*.*d* Format" on page 185
"TIME*w*.*d* Format" on page 245

Functions:

"HMS Function" on page 803
"MINUTE Function" on page 928
"SECOND Function" on page 1122

Informat:

"TIME*w*. Informat" on page 1393

MMYY*w*. Format

Writes date values in the form *mm*M<*yy*>*yy*, where M is the separator and the year appears as either 2 or 4 digits.

Category: Date and Time
Alignment: right

Syntax

MMYY*w*.

Syntax Description

w
specifies the width of the output field.

Default: 7

Range: 5–32

Interaction: When *w* has a value of 5 or 6, the date appears with only the last two digits of the year. When *w* is 7 or more, the date appears with a four-digit year.

Details

The MMYY*w*. format writes SAS date values in the form *mm*M<*yy*>*yy*, where

mm
is an integer that represents the month.

M
is the character separator.

<*yy*>*yy*
is a two-digit or four-digit integer that represents the year.

Examples

The following examples use the input value of 16734, which is the SAS date value that corresponds to October 25, 2005.

SAS Statement	Results
	`----+----1----+`
`put date mmyy5.;`	`10M05`
`put date mmyy6.;`	`10M05`
`put date mmyy.;`	`10M2005`
`put date mmyy7.;`	`10M2005`
`put date mmyy10.;`	`10M2005`

See Also

Format:

"MMYY*xw*. Format" on page 203
"YYMM*w*. Format" on page 270

MMYY*xw*. Format

Writes date values in the form *mm<yy>yy* or *mm-<yy>yy*, where the *x* in the format name is a character that represents the special character that separates the month and the year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.

Category: Date and Time

Alignment: right

Syntax

MMYY*xw.*

Syntax Description

x

identifies a separator or specifies that no separator appear between the month and the year. Valid values for *x* are

C
separates with a colon

D
separates with a dash

N
indicates no separator

P
separates with a period

S
separates with a forward slash.

w

specifies the width of the output field.

Default: 7

Range: 5–32

Interaction: When *x* is set to N, no separator is specified. The width range is then 4–32, and the default changes to 6.

Interaction: When *x* has a value of C, D, P, or S and *w* has a value of 5 or 6, the date appears with only the last two digits of the year. When *w* is 7 or more, the date appears with a four-digit year.

Interaction: When *x* has a value of *N* and *w* has a value of 4 or 5, the date appears with only the last two digits of the year. When *x* has a value of *N* and *w* is 6 or more, the date appears with a four-digit year.

Details

The MMYY*xw*. format writes SAS date values in the form *mm<yy>yy* or *mmx<yy>yy*, where

mm
is an integer that represents the month.

x
is a specified separator.

<yy>yy
is a two-digit or four-digit integer that represents the year.

Examples

The following examples use the input value of 18031, which is the SAS date value that corresponds to May 14, 2009.

SAS Statement	Results
	`----+----1----+`
`put date mmyyc5.;`	`05:09`
`put date mmyyd.;`	`05-2009`
`put date mmyyn4.;`	`0509`
`put date mmyyp8.;`	`05.2009`
`put date mmyys10.;`	`05/2009`

See Also

Format:

"MMYY*w*. Format" on page 201
"YYMM*xw*. Format" on page 271

MONNAME*w*. Format

Writes date values as the name of the month.

Category: Date and Time
Alignment: right

Syntax

MONNAME*w*.

Syntax Description

w
specifies the width of the output field.

Default: 9

Range: 1–32

Tip: Use MONNAME3. to print the first three letters of the month name.

Details

If necessary, SAS truncates the name of the month to fit the format width.

Examples

The example table uses the input value of 16500, which is the SAS date value that corresponds to March 5, 2005.

SAS Statement	Results
	`----+----1`
`put date monname1.;`	`M`
`put date monname3.;`	`Mar`
`put date monname5.;`	`March`

See Also

Format:

"MONTH*w*. Format" on page 205

MONTH*w*. Format

Writes date values as the month of the year.

Category: Date and Time

Alignment: right

Syntax

MONTH*w*.

Syntax Description

w
specifies the width of the output field.

Default: 2

Range: 1–32

Tip: Use MONTH1. to obtain a hexadecimal value.

Details

The MONTH*w*. format writes the month (1 through 12) of the year from a SAS date value. If the month is a single digit, the MONTH*w*. format places a leading blank before the digit. For example, the MONTH*w*. format writes 4 instead of 04.

Examples

The example table uses the input value of 18031, which is the SAS date value that corresponds to May 14, 2009.

SAS Statement	Results
	`----+----1`
`put date month.;`	`5`

See Also

Format:

"MONNAME*w*. Format" on page 204

MONYY*w*. Format

Writes date values as the month and the year in the form *mmmyy* or *mmmyyyy*.

Category: Date and Time

Alignment: right

Syntax

MONYY*w*.

Syntax Description

w
specifies the width of the output field.

Default: 5

Range: 5–7

Details

The MONYY*w*. format writes SAS date values in the form *mmmyy* or *mmmyyyy*, where

mmm
is the first three letters of the month name.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

Comparisons

The MONYY*w*. format and the DTMONYY*w*. format are similar in that they both write date values. The difference is that MONYY*w*. expects a SAS date value as input, and DTMONYY*w*. expects a datetime value.

Examples

The example table uses the input value of 16794, which is the SAS date value that corresponds to December 24, 2005.

SAS Statement	Results
	`----+----1`
`put date monyy5.;`	`DEC05`
`put date monyy7.;`	`DEC2005`

See Also

Formats:

"DTMONYY*w*. Format" on page 165
"DDMMYY*w*. Format" on page 157
"MMDDYY*w*. Format" on page 196
"YYMMDD*w*. Format" on page 273

Functions:

"MONTH Function" on page 936
"YEAR Function" on page 1233

Informat:

"MONYY*w*. Informat" on page 1351

NEGPARENw.d Format

Writes negative numeric values in parentheses.

Category: Numeric

Alignment: right

Syntax

NEGPAREN*w*.*d*

Syntax Description

w

specifies the width of the output field.

Default: 6

Range: 1–32

d

specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Default: 0

Range: 0–31

Details

The NEGPAREN*w*.*d* format attempts to right-align output values. If the input value is negative, NEGPAREN*w*.*d* displays the output by enclosing the value in parentheses, if the field that you specify is wide enough. Otherwise, it uses a minus sign to represent the negative value. If the input value is non-negative, NEGPAREN*w*.*d* displays the value with a leading and trailing blank to ensure proper column alignment. It reserves the last column for a close parenthesis even when the value is positive.

Comparisons

The NEGPAREN*w*.*d* format is similar to the COMMA*w*.*d* format in that it separates every three digits of the value with a comma.

Examples

```
put @1 sales negparen8.;
```

Value of **sales**	Results
	----+----1----+
100	100
1000	1,000

Value of **sales**	Results
-200	**(200)**
-2000	**(2,000)**

NUMX*w.d* Format

Writes numeric values with a comma in place of the decimal point.

Category: Numeric

Alignment: right

Syntax

NUMX*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 12

Range: 1–32

d
specifies the number of digits to the right of the decimal point (comma) in the numeric value. This argument is optional.

Default: 0

Range: 0–31

Details

The NUMX*w.d* format writes numeric values with a comma in place of the decimal point.

Comparisons

The NUMX*w.d* format is similar to the *w.d* format except that NUMX*w.d* writes numeric values with a comma in place of the decimal point.

Examples

```
put x numx10.2;
```

Value of **x**	Results
	`----+----1----+`
`896.48`	`896,48`
`64.89`	`64,89`
`3064.10`	`3064,10`

See Also

Format:

"*w*.*d* Format" on page 254

Informat:

"NUMX*w*.*d* Informat" on page 1353

OCTAL*w*. Format

Converts numeric values to octal representation.

Category: Numeric
Alignment: left

Syntax

OCTAL*w*.

Syntax Description

w

specifies the width of the output field.

Default: 3

Range: 1–24

Details

If necessary, the OCTAL*w*. format converts numeric values to integers before displaying them in octal representation.

Comparisons

OCTAL*w*. converts numeric values to octal representation. The $OCTAL*w*. format converts character values to octal representation.

Examples

```
put x octal6.;
```

Value of **x**	Results
	`----+----1`
`3592`	`007010`

PD*w.d* Format

Writes data in packed decimal format.

Category: Numeric

Alignment: left

See: PD*w.d* Format in the documentation for your operating environment.

Syntax

PD*w.d*

Syntax Description

w

specifies the width of the output field. The *w* value specifies the number of bytes, not the number of digits. (In packed decimal data, each byte contains two digits.)

Default: 1

Range: 1–16

d

specifies to multiply the number by 10^{d}. This argument is optional.

Default: 0

Range: 0–31

Details

Different operating environments store packed decimal values in different ways. However, the PD*w.d* format writes packed decimal values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

The PD*w.d* format writes missing numerical data as –0. When the PD*w.d* informat reads a –0, it stores it as 0.

Comparisons

The following table compares packed decimal notation in several programming languages:

Language	Notation
SAS	PD4.
COBOL	COMP-3 PIC S9(7)
IBM 370 assembler	PL4
PL/I	FIXED DEC

Examples

```
y=put(x,pd4.);
put y $hex8.;
```

Value of **x**	Results*
	----+----1
128	**00000128**

* The result is a hexadecimal representation of a binary number written in packed decimal format. Each byte occupies one column of the output field.

PDJULG*w*. Format

Writes packed Julian date values in the hexadecimal format *yyyyddd*F for IBM.

Category: Date and Time

Syntax

PDJULG*w*.

Syntax Description

w
specifies the width of the output field.

Default: 4

Range: 3-16

Details

The PDJULG*w*. format writes SAS date values in the form *yyyyddd*F, where

yyyy
is the two-byte representation of the four-digit Gregorian year.

ddd
is the one-and-a-half byte representation of the three-digit integer that corresponds to the Julian day of the year, 1–365 (or 1–366 for leap years).

F
is the half byte that contains all binary 1s, which assigns the value as positive.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

SAS Statement	Results
	`----+----1`

`date = '17mar2005'd;`
`juldate = put(date,pdjulg4.);`
`put juldate $hex8.;`
	`2005076F`

See Also

Formats:

"PDJULI*w*. Format" on page 214
"JULIAN*w*. Format" on page 194
"JULDAY*w*. Format" on page 193

Functions:

"JULDATE Function" on page 866
"DATEJUL Function" on page 635

Informats:

"PDJULI*w*. Informat" on page 1358
"PDJULG*w*. Informat" on page 1357
"JULIAN*w*. Informat" on page 1346

System Option:

"YEARCUTOFF= System Option" on page 2058

PDJULI*w*. Format

Writes packed Julian date values in the hexadecimal format *ccyyddd*F for IBM.

Category: Date and Time

Syntax

PDJULI*w*.

Syntax Description

w
specifies the width of the output field.

Default: 4

Range: 3-16

Details

The PDJULI*w*. format writes SAS date values in the form *ccyyddd*F, where

cc
is the one-byte representation of a two-digit integer that represents the century.

yy
is the one-byte representation of a two-digit integer that represents the year. The PDJULI*w*. format makes an adjustment for the century byte by subtracting 1900 from the 4–digit Gregorian year to produce the correct packed decimal *ccyy* representation. A year value of 1998 is stored in *ccyy* as 0098, and a year value of 2011 is stored as 0111.

ddd
is the one-and-a-half byte representation of the three-digit integer that corresponds to the Julian day of the year, 1–365 (or 1–366 for leap years).

F
is the half byte that contains all binary 1s, which assigns the value as positive.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

SAS Statement	Results
	`----+----1`

`date = '17mar2005'd;`
`juldate = put(date,pdjuli4.);`
`put juldate $hex8.;`
	`0105076F`

`date = '31dec2003'd;`
`juldate = put(date,pdjuli4.);`
`put juldate $hex8.;`
	`0103365F`

See Also

Formats:

"PDJULG*w*. Format" on page 213
"JULIAN*w*. Format" on page 194
"JULDAY*w*. Format" on page 193

Functions:

"DATEJUL Function" on page 635

"JULDATE Function" on page 866

Informats:

"PDJULG*w*. Informat" on page 1357

"PDJULI*w*. Informat" on page 1358

"JULIAN*w*. Informat" on page 1346

System Option:

"YEARCUTOFF= System Option" on page 2058

PERCENT*w.d* Format

Writes numeric values as percentages.

Category: Numeric

Alignment: right

Syntax

PERCENT*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 6

Range: 4–32

Tip: The width of the output field must account for the percent sign (%)and parentheses for negative numbers, whether the number is negative or positive.

d

specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Range: 0–31

Requirement: must be less than *w*

Details

The PERCENT*w.d* format multiplies values by 100, formats them the same as the BEST*w.d* format, and adds a percent sign (%) to the end of the formatted value, while it encloses negative values in parentheses.

Examples

```
put @10 gain percent10.;
```

Value of **x**	Results
	----+----1----+----2
0.1	10%
1.2	120%
-0.05	(5%)

See Also

Format:

"PERCENTN*w.d* Format" on page 217

PERCENTN*w.d* Format

Produces percentages, using a minus sign for negative values.

Category: Numeric

Alignment: right

Syntax

PERCENTN*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 6

Range: 4–32

Tip: The width of the output field must account for the minus sign (–), the percent sign (%), and a trailing blank, whether the number is negative or positive.

d

specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Range: 0–31

Requirement: must be less than *w*

Details

The PERCENTN*w.d* format multiplies negative values by 100, formats them the same as the BEST*w.d* format, adds a minus sign to the beginning of the value, and adds a percent sign (%) to the end of the formatted value.

Comparisons

The PERCENTN*w.d* format produces percents by using a minus sign instead of parentheses for negative values. The PERCENT*w.d* format produces percents by using parentheses for negative values.

Examples

```
put x percentn10.;
```

Value of **x**	Results
--0.1	**-10%**
.2	**20%**
.8	**80%**
--0.05	**-5%**
--6.3	**--630%**

See Also

Format:

"PERCENT*w.d* Format" on page 216

PIB*w.d* Format

Writes positive integer binary (fixed-point) values.

Category: Numeric
Alignment: left
See: PIB*w.d* Format in the documentation for your operating environment.

Syntax

PIB*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 1
Range: 1–8

d
specifies to multiply the number by 10^d. This argument is optional.
Default: 0
Range: 0–31

Details

All values are treated as positive. PIB*w.d* writes positive integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 88. △

Comparisons

- Positive integer binary values are the same as integer binary values except that the sign bit is part of the value, which is always a positive integer. The PIB*w.d* format treats all values as positive and includes the sign bit as part of the value.
- The PIB*w.d* format with a width of 1 results in a value that corresponds to the binary equivalent of the contents of a byte. A value that corresponds to the binary equivalent of the contents of a byte is useful if your data contain values between hexadecimal 80 and hexadecimal FF, where the high-order bit can be misinterpreted as a negative sign.
- The PIB*w.d* format is the same as the IB*w.d* format except that PIB*w.d* treats all values as positive values.
- The IB*w.d* and PIB*w.d* formats are used to write native format integers. (Native format allows you to read and write values that are created in the same operating environment.) The IBR*w.d* and PIBR*w.d* formats are used to write little endian integers in any operating environment.

  To view a table that shows the type of format to use with big endian and little endian integers, see Table 3.1 on page 88.

  To view a table that compares integer binary notation in several programming languages, see Table 3.2 on page 89.

Examples

```
y=put(x,pib1.);
put y $hex2.;
```

Value of **x**	Results
	----+----1
12	**0C**

* The result is a hexadecimal representation of a one-byte binary number written in positive integer binary format, which occupies one column of the output field.

See Also

Format:

"PIBR*w.d* Format" on page 221

PIBR*w.d* Format

Writes positive integer binary (fixed-point) values in Intel and DEC formats.

Category: Numeric

Syntax

PIBR*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–8

d
specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–10

Details

All values are treated as positive. PIBR*w.d* writes positive integer binary values that have been generated by and for Intel and DEC operating environments. Use PIBR*w.d* to write positive integer binary data from Intel or DEC environments on other operating environments. The PIBR*w.d* format in SAS code allows for a portable implementation for writing the data in any operating environment.

Note: Different operating environments store positive integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 88. △

Comparisons

- Positive integer binary values are the same as integer binary values except that the sign bit is part of the value, which is always a positive integer. The PIBR*w.d* format treats all values as positive and includes the sign bit as part of the value.
- The PIBR*w.d* format with a width of 1 results in a value that corresponds to the binary equivalent of the contents of a byte. A value that corresponds to the binary equivalent of the contents of a byte is useful if your data contain values between hexadecimal 80 and hexadecimal FF, where the high-order bit can be misinterpreted as a negative sign.
- On Intel and DEC operating environments, the PIB*w.d* and PIBR*w.d* formats are equivalent.
- The IB*w.d* and PIB*w.d* formats are used to write native format integers. (Native format allows you to read and write values that are created in the same operating

environment.) The IBR*w.d* and PIBR*w.d* formats are used to write little endian integers in any operating environment.

To view a table that shows the type of format to use with big endian and little endian integers, see Table 3.1 on page 88.

To view a table that compares integer binary notation in several programming languages, see Table 3.2 on page 89.

Examples

```
y=put(x,pibr2.);
put y $hex4.;
```

Value of **x**	Results
	----+----1
128	**8000**

* The result is a hexadecimal representation of a two-byte binary number written in positive integer binary format, which occupies one column of the output field.

See Also

Informat:

"PIB*w.d* Informat" on page 1362

PK*w.d* Format

Writes data in unsigned packed decimal format.

Category: Numeric

Alignment: left

Syntax

PK*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 1

Range: 1–16

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–10

Requirement: must be less than *w*

Details

Each byte of unsigned packed decimal data contains two digits.

Comparisons

The PK*w.d* format is similar to the PD*w.d* format except that PK*w.d* does not write the sign in the low-order byte.

Examples

```
y=put(x,pk4.);
put y $hex8.;
```

Value of **x**	Results*
	----+----1
128	**00000128**

* The result is a hexadecimal representation of a four-byte number written in packed decimal format. Each byte occupies one column of the output field.

PVALUE*w.d* Format

Writes *p*-values.

Category: Numeric

Alignment: right

Syntax

PVALUE*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 6

Range: 3–32

d

specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Default: the minimum of 4 and *w*–2

Range: 1–30

Restriction: must be less than *w*

Comparisons

The PVALUE*w*.*d* format follows the rules for the *w*.*d* format, except that

- if the value x is such that 0 <= x < 10^{-d}, x prints as "<.0...01" with *d*-1 zeros
- missing values print as "." unless you specify a different character by using the MISSING= system option

Examples

```
put x pvalue6.4;
```

Value of **x**	Results
	----+----1
.05	**0.0500**
0.000001	**<.0001**
0	**<.0001**
.0123456	**0.0123**

QTR*w*. Format

Writes date values as the quarter of the year.

Category: Date and Time

Alignment: right

Syntax

QTR*w*.

Syntax Description

w

specifies the width of the output field.

Default: 1

Range: 1–32

Examples

The example table uses the input value of 16500, which is the SAS date value that corresponds to March 5, 2005.

SAS Statement	Results
	`----+----1`
`put date qtr.;`	`1`

See Also

Format:

"QTRR*w*. Format" on page 225

QTRR*w*. Format

Writes date values as the quarter of the year in Roman numerals.

Category: Date and Time

Alignment: right

Syntax

QTRR*w*.

Syntax Description

w

specifies the width of the output field.

Default: 3

Range: 3–32

Examples

The example table uses the input value of 16694, which is the SAS date value that corresponds to September 15, 2005.

SAS Statement	Results
	`----+----1`
`put date qtrr.;`	`III`

See Also

Format:

"QTR*w*. Format" on page 224

RB*w.d* Format

Writes real binary data (floating-point) in real binary format.

Category: Numeric

Alignment: left

See: RB*w.d* Format in the documentation for your operating environment.

Syntax

RB*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 4

Range: 2–8

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–10

Details

The RB*w.d* format writes numeric data in the same way that SAS stores them. Because it requires no data conversion, RB*w.d* is the most efficient method for writing data with SAS.

Note: Different operating environments store real binary values in different ways. However, RB*w.d* writes real binary values with consistent results in the same type of operating environment that you use to run SAS. △

CAUTION:

Using RB4. to write real binary data on equipment that conforms to the IEEE standard for floating-point numbers results in a truncated eight-byte (double-precision) number rather than a true four-byte (single-precision) floating-point number. △

Comparisons

The following table compares the names of real binary notation in several programming languages:

Language	4 Bytes	8 Bytes
SAS	RB4.	RB8.
Fortran	REAL*4	REAL*8
C	float	double
COBOL	COMP-1	COMP-2
IBM 370 assembler	E	D

Examples

```
y=put(x,rb8.);
put y $hex16.;
```

Value of **x**	Results
	----+---1----+----2
128	**4280000000000000**

* The result is a hexadecimal representation of an eight-byte real binary number as it looks on an IBM mainframe. Each byte occupies one column of the output field.

ROMAN*w*. Format

Writes numeric values as roman numerals.

Category: Numeric
Alignment: left

Syntax

ROMAN*w*.

Syntax Description

w
specifies the width of the output field.
Default: 6
Range: 2–32

Details

The ROMAN*w*. format truncates a floating-point value to its integer component before the value is written.

Examples

```
put @5 year roman10.;
```

Value of **year**	Results
1998	**MCMXCVIII**

S370FF*w.d* Format

Writes native standard numeric data in IBM mainframe format.

Category: Numeric

Syntax

S370FF*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 12

Range: 1–32

d
specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–31

Details

The S370FF*w.d* format writes numeric data in IBM mainframe format (EBCDIC). The EBCDIC numeric values are represented with one byte per digit. If EBCDIC is the native format, S370FF*w.d* performs no conversion.

If a value is negative, an EBCDIC minus sign precedes the value. A missing value is represented as a single EBCDIC period.

Comparisons

On an EBCDIC system, S370FF*w.d* behaves like the *w.d* format.

On all other systems, S370FF*w.d* performs the same role for numeric data that the $EBCDIC*w*. format does for character data.

Examples

```
y=put(x,s370ff5.);
put y $hex10.;
```

Value of **x**	Results*
----+----1	
12345	**F1F2F3F4F5**

* The result is the hexadecimal representation for the integer.

See Also

Formats:

"$EBCDIC*w*. Format" on page 112
"*w.d* Format" on page 254

S370FIB*w.d* Format

Writes integer binary (fixed-point) values, including negative values, in IBM mainframe format.

Category: Numeric
Alignment: left

Syntax

S370FIB*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 4
Range: 1–8

d
specifies to multiply the number by 10^d. This argument is optional.
Default: 0
Range: 0–10

Details

The S370FIB*w.d* format writes integer binary (fixed-point) values that are stored in IBM mainframe format, including negative values that are represented in two's complement notation. S370FIB*w.d* writes integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Use S370FIB*w.d* to write integer binary data in IBM mainframe format from data that are created in other operating environments.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 88. △

Comparisons

- If you use SAS on an IBM mainframe, S370FIB*w.d* and IB*w.d* are identical.
- S370FPIB*w.d*, S370FIBU*w.d*, and S370FIB*w.d* are used to write big endian integers in any operating environment.

  To view a table that shows the type of format to use with big endian and little endian integers, see Table 3.1 on page 88.

  To view a table that compares integer binary notation in several programming languages, see Table 3.2 on page 89.

Examples

```
y=put(x,s370fib4.);
put y $hex8.;
```

Value of **x**	Results
	----+----1
128	**00000080**

* The result is a hexadecimal representation of a 4-byte integer binary number. Each byte occupies one column of the output field.

See Also

Formats:

"S370FIBU*w.d* Format" on page 231
"S370FPIB*w.d* Format" on page 236

S370FIBU*w.d* Format

Writes unsigned integer binary (fixed-point) values in IBM mainframe format.

Category: Numeric
Alignment: left

Syntax

S370FIBU*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 4

Range: 1–8

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–10

Details

The S370FIBU*w.d* format writes unsigned integer binary (fixed-point) values that are stored in IBM mainframe format, including negative values that are represented in two's complement notation. Unsigned integer binary values are the same as integer binary values, except that all values are treated as positive. S370FIBU*w.d* writes integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Use S370FIBU*w.d* to write unsigned integer binary data in IBM mainframe format from data that are created in other operating environments.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 88. △

Comparisons

- The S370FIBU*w.d* format is equivalent to the COBOL notation PIC 9(*n*) BINARY, where *n* is the number of digits.
- The S370FIBU*w.d* format is the same as the S370FIB*w.d* format except that the S370FIBU*w.d* format always uses the absolute value instead of the signed value.
- The S370FPIB*w.d* format writes all negative numbers as FFs, while the S370FIBU*w.d* format writes the absolute value.
- S370FPIB*w.d*, S370FIBU*w.d*, and S370FIB*w.d* are used to write big endian integers in any operating environment.

  To view a table that shows the type of format to use with big endian and little endian integers, see Table 3.1 on page 88.

  To view a table that compares integer binary notation in several programming languages, see Table 3.2 on page 89.

Examples

```
y=put(x,s370fibu1.);
put y $hex2.;
```

Value of **x**	Results*
245	**F5**
-245	**F5**

* The result is a hexadecimal representation of a one-byte integer binary number. Each byte occupies one column of the output field.

See Also

Formats:

"S370FIB*w.d* Format" on page 230
"S370FPIB*w.d* Format" on page 236

S370FPD*w.d* Format

Writes packed decimal data in IBM mainframe format.

Category: Numeric
Alignment: left

Syntax

S370FPD*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 1
Range: 1–16

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–31

Details

Use S370FPD*w.d* in other operating environments to write packed decimal data in the same format as on an IBM mainframe computer.

Comparisons

The following table shows the notation for equivalent packed decimal formats in several programming languages:

Language	Packed Decimal Notation
SAS	S370FPD4.
PL/I	FIXED DEC(7,0)
COBOL	COMP-3 PIC S9(7)
IBM 370 assembler	PL4

Examples

```
y=put(x,s370fpd4.);
put y $hex8.;
```

Value of **x**	Results
	----+----1
128	**0000128C**

* The result is a hexadecimal representation of a binary number written in packed decimal format. Each byte occupies one column of the output field.

S370FPDU*w.d* Format

Writes unsigned packed decimal data in IBM mainframe format.

Category: Numeric
Alignment: left

Syntax

S370FPDU*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 1
Range: 1–16

d
specifies to multiply the number by 10^{d}. This argument is optional.
Default: 0
Range: 0–31

Details

Use S370FPDU*w.d* in other operating environments to write unsigned packed decimal data in the same format as on an IBM mainframe computer.

Comparisons

- The S370FPDU*w.d* format is similar to the S370FPD*w.d* format except that the S370FPD*w.d* format always uses the absolute value instead of the signed value.
- The S370FPDU*w.d* format is equivalent to the COBOL notation PIC 9(*n*) PACKED-DECIMAL, where the *n* value is the number of digits.

Examples

```
y=put(x,s370fpdu2.);
put y $hex4.;
```

Value of **x**	Results
123	**123F**
-123	**123F**

* The result is a hexadecimal representation of a binary number written in packed decimal format. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one column of the output field.

S370FPIB*w.d* Format

Writes positive integer binary (fixed-point) values in IBM mainframe format.

Category: Numeric

Alignment: left

Syntax

S370FPIB*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 4

Range: 1–8

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–10

Details

Positive integer binary values are the same as integer binary values, except that all values are treated as positive. S370FPIB*w.d* writes integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Use S370FPIB*w.d* to write positive integer binary data in IBM mainframe format from data that are created in other operating environments.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 88. △

Comparisons

- If you use SAS on an IBM mainframe, S370FPIB*w.d* and PIB*w.d* are identical.
- The S370FPIB*w.d* format is the same as the S370FIB*w.d* format except that the S370FPIB*w.d* format treats all values as positive values.
- S370FPIB*w.d*, S370FIBU*w.d*, and S370FIB*w.d* are used to write big endian integers in any operating environment.

  To view a table that shows the type of format to use with big endian and little endian integers, see Table 3.1 on page 88.

  To view a table that compares integer binary notation in several programming languages, see Table 3.2 on page 89.

Examples

```
y=put(x,s370fpib1.);
put y $hex2.;
```

Value of **x**	Results*
	----+----1
12	**0C**

*The result is a hexadecimal representation of a one-byte binary number written in positive integer binary format, which occupies one column of the output field.

See Also

Formats:

"S370FIB*w.d* Format" on page 230

"S370FIBU*w.d* Format" on page 231

S370FRB*w.d* Format

Writes real binary (floating-point) data in IBM mainframe format.

Category: Numeric

Alignment: left

Syntax

S370FRB*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 4
Range: 2–8

d
specifies to multiply the number by 10^d. This argument is optional.
Default: 0
Range: 0–10

Details

A floating-point value consists of two parts: a mantissa that gives the value and an exponent that gives the value's magnitude.

Use S370FRB*w.d* in other operating environments to write floating-point binary data in the same format as on an IBM mainframe computer.

Comparisons

The following table shows the notation for equivalent floating-point formats in several programming languages:

Language	4 Bytes	8 Bytes
SAS	S370FRB4.	S370FRB8.
PL/I	FLOAT BIN(21)	FLOAT BIN(53)
Fortran	REAL*4	REAL*8
COBOL	COMP-1	COMP-2
IBM 370 assembler	E	D
C	float	double

Examples

```
y=put(x,s370frb6.);
put y $hex8.;
```

Value of **x**	Results*
128	**42800000**
-123	**C2800000**

* The result is a hexadecimal representation of a binary number in zoned decimal format on an IBM mainframe computer. Each two hexadecimal digits correspond to one byte of binary data, and each byte corresponds to one column of the output field.

S370FZD*w.d* Format

Writes zoned decimal data in IBM mainframe format.

Category: Numeric

Alignment: left

Syntax

S370FZD*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 8

Range: 1–32

d
specifies to multiply the number by 10^{d}. This argument is optional.

Default: 0

Range: 0–31

Details

Use S370FZD*w.d* in other operating environments to write zoned decimal data in the same format as on an IBM mainframe computer.

Comparisons

The following table shows the notation for equivalent zoned decimal formats in several programming languages:

Language	Zoned Decimal Notation
SAS	S370FZD3.
PL/I	PICTURE '99T'
COBOL	PIC S9(3) DISPLAY
assembler	ZL3

Examples

```
y=put(x,s370fzd3.);
put y $hex6.;
```

Value of **x**	Results*
123	**F1F2C3**
-123	**F1F2D3**

* The result is a hexadecimal representation of a binary number in zoned decimal format on an IBM mainframe computer. Each two hexadecimal digits correspond to one byte of binary data, and each byte corresponds to one column of the output field.

S370FZDL*w.d* Format

Writes zoned decimal leading–sign data in IBM mainframe format.

Category: Numeric
Alignment: left

Syntax

S370FZDL*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 8
Range: 1–32

d
specifies to multiply the number by 10^{d}. This argument is optional.
Default: 0
Range: 0–31

Details

Use S370FZDL*w.d* in other operating environments to write zoned decimal leading-sign data in the same format as on an IBM mainframe computer.

Comparisons

- The S370FZDL*w.d* format is similar to the S370FZD*w.d* format except that the S370FZDL*w.d* format displays the sign of the number in the first byte of the formatted output.
- The S370FZDL*w.d* format is equivalent to the COBOL notation PIC S9(*n*) DISPLAY SIGN LEADING, where the *n* value is the number of digits.

Examples

```
y=put(x,s370fzdl3.);
put y $hex6.;
```

Value of **x**	Results*
123	**C1F2F3**
-123	**D1F2F3**

* The result is a hexadecimal representation of a binary number in zoned decimal format on an IBM mainframe computer. Each two hexadecimal digits correspond to one byte of binary data, and each byte corresponds to one column of the output field.

S370FZDS*w.d* Format

Writes zoned decimal separate leading-sign data in IBM mainframe format.

Category: Numeric

Alignment: left

Syntax

S370FZDS*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 8

Range: 2–32

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–31

Details

Use S370FZDS*w.d* in other operating environments to write zoned decimal separate leading-sign data in the same format as on an IBM mainframe computer.

Comparisons

- The S370FZDS*w.d* format is similar to the S370FZDL*w.d* format except that the S370FZDS*w.d* format does not embed the sign of the number in the zoned output.
- The S370FZDS*w.d* format is equivalent to the COBOL notation PIC S9(*n*) DISPLAY SIGN LEADING SEPARATE, where the *n* value is the number of digits.

Examples

```
y=put (x,s370fzds4.);
put y $hex8.;
```

Value of **x**	Results*
123	**4EF1F2F3**
-123	**60F1F2F3**

* The result is a hexadecimal representation of a binary number in zoned decimal format on an IBM mainframe computer. Each two hexadecimal digits correspond to one byte of binary data, and each byte corresponds to one column of the output field.

S370FZDT*w.d* Format

Writes zoned decimal separate trailing-sign data in IBM mainframe format.

Category: Numeric
Alignment: left

Syntax

S370FZDT*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 8

Range: 2–32

d
specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–31

Details

Use S370FZDT*w.d* in other operating environments to write zoned decimal separate trailing-sign data in the same format as on an IBM mainframe computer.

Comparisons

- The S370FZDT*w.d* format is similar to the S370FZDS*w.d* format except that the S370FZDT*w.d* format displays the sign of the number at the end of the formatted output.
- The S370FZDT*w.d* format is equivalent to the COBOL notation PIC S9(*n*) DISPLAY SIGN TRAILING SEPARATE, where the *n* value is the number of digits.

Examples

```
y=put (x,s370fzdt4.); ;
put y $hex8.;
```

Value of **x**	Results*
123	**F1F2F34E**
-123	**F1F2F360**

* The result is a hexadecimal representation of a binary number in zoned decimal format on an IBM mainframe computer. Each two hexadecimal digits correspond to one byte of binary data, and each byte corresponds to one column of the output field.

S370FZDU*w.d* Format

Writes unsigned zoned decimal data in IBM mainframe format.

Category: Numeric

Alignment: left

Syntax

S370FZDU*w.d*

Syntax Description

w
specifies the width of the output field.
Default: 8
Range: 1–32

d
specifies to multiply the number by 10^d. This argument is optional.
Default: 0
Range: 0–31

Details

Use S370FZDU*w*.*d* in other operating environments to write unsigned zoned decimal data in the same format as on an IBM mainframe computer.

Comparisons

- The S370FZDU*w*.*d* format is similar to the S370FZD*w*.*d* format except that the S370FZDU*w*.*d* format always uses the absolute value of the number.
- The S370FZDU*w*.*d* format is equivalent to the COBOL notation PIC 9(*n*) DISPLAY, where the *n* value is the number of digits.

Examples

```
y=put (x,s370fzdu3.);
put y $hex6.;
```

Value of **x**	Results*
123	**F1F2F3**
-123	**F1F2F3**

* The result is a hexadecimal representation of a binary number in zoned decimal format on an IBM mainframe computer. Each pair of hexadecimal characters (such as F1) corresponds to one byte of binary data, and each byte corresponds to one column of the output field.

SSN*w*. Format

Writes Social Security numbers.

Category: Numeric
Alignment: none

Syntax

SSN*w*.

Syntax Description

w
specifies the width of the output field.

Default: 11

Restriction: *w* must be 11

Details

If the value is missing, SAS writes nine single periods with dashes between the third and fourth periods and between the fifth and sixth periods. If the value contains fewer than nine digits, SAS right aligns the value and pads it with zeros on the left. If the value has more than nine digits, SAS writes it as a missing value.

Examples

```
put id ssn11.;
```

Value of **id**	Results
	`----+----1----+`
`263878439`	`263-87-8439`

TIME*w.d* Format

Writes time values as hours, minutes, and seconds in the form *hh:mm:ss.ss*.

Category: Date and Time

Alignment: right

Syntax

TIME*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 8

Range: 2–20

Tip: Make *w* large enough to produce the desired results. To obtain a complete time value with three decimal places, you must allow at least 12 spaces: eight spaces to

the left of the decimal point, one space for the decimal point itself, and three spaces for the decimal fraction of seconds.

d

specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0–19

Requirement: must be less than *w*

Details

The TIME*w.d* format writes SAS time values in the form *hh:mm:ss.ss*, where

hh

is an integer.

Note: If *hh* is a single digit, TIME*w.d* places a leading blank before the digit. For example, the TIME*w.d.* format writes 9:00 instead of 09:00. △

mm

is the number of minutes, ranging from 00 through 59.

ss.ss

is the number of seconds, ranging from 00 through 59, with the fraction of a second following the decimal point.

Comparisons

The TIME*w.d* format is similar to the HHMM*w.d* format except that TIME*w.d* includes seconds.

The TIME*w.d* format writes a leading blank for a single-hour digit. The TOD*w.d* format writes a leading zero for a single-hour digit.

Examples

The example table uses the input value of 59083, which is the SAS time value that corresponds to 4:24:43 p.m.

SAS Statement	Results
	`----+----1`
`put begin time.;`	`16:24:43`

See Also

Formats:

- "HHMM*w.d* Format" on page 185
- "HOUR*w.d* Format" on page 188
- "MMSS*w.d* Format" on page 200
- "TOD*w.d* Format" on page 249

Functions:

- "HOUR Function" on page 807
- "MINUTE Function" on page 928
- "SECOND Function" on page 1122
- "TIME Function" on page 1161

Informat:

- "TIME*w*. Informat" on page 1393

TIMEAMPM*w.d* Format

Writes time and datetime values as hours, minutes, and seconds in the form *hh:mm:ss.ss* with AM or PM.

Category: Date and Time

Alignment: right

Syntax

TIMEAMPM*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 11

Range: 2–20

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0–19

Requirement: must be less than *w*

Details

The TIMEAMPM*w.d* format writes SAS time values and SAS datetime values in the form *hh:mm:ss.ss* with AM or PM, where

hh
is an integer that represents the hour.

mm
is an integer that represents the minutes.

ss.ss
is the number of seconds to two decimal places.

Times greater than 23:59:59 PM appear as the next day.

Make *w* large enough to produce the desired results. To obtain a complete time value with three decimal places and AM or PM, you must allow at least 11 spaces (*hh:mm:ss* PM). If *w* is less than 5, SAS writes AM or PM only.

Comparisons

- The TIMEAMPMM*w.d* format is similar to the TIMEM*w.d* format except, that TIMEAMPMM*w.d* prints AM or PM at the end of the time.
- TIME*w.d* writes hours greater than 23:59:59 PM, and TIMEAMPM*w.d* does not.

Examples

The example table uses the input value of 59083, which is the SAS time value that corresponds to 4:24:43 p.m.

SAS Statement	Results
	`----+----1----+`
`put begin timeampm3.;`	` PM`
`put begin timeampm5.;`	` 4 PM`
`put begin timeampm7.;`	`4:24 PM`
`put begin timeampm11.;`	` 4:24:43 PM`

See Also

Format:

"TIME*w.d* Format" on page 245

TOD*w.d* Format

Writes SAS time values and the time portion of SAS datetime values in the form *hh:mm:ss.ss*.

Category: Date and Time
Alignment: right

Syntax

TOD*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 8

Range: 2–20

Tip: SAS writes a zero for a zero hour if the specified width is sufficient. For example, 02:30 or 00:30.

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0–19

Requirement: must be less than *w*

Details

The TOD*w.d* format writes SAS time and datetime values in the form *hh:mm:ss.ss*, where

hh
is an integer that represents the hour.

mm
is an integer that represents the minutes.

ss.ss
is the number of seconds to two decimal places.

Comparisons

The TOD*w.d* format writes a leading zero for a single-hour digit. The TIME*w.d* format and the HHMM*w.d* format write a leading blank for a single-hour digit.

Examples

In the following example, the SAS datetime value 1566742823 corresponds to August 24, 2009 at 2:20:23 p.m.

SAS Statement	Results
	`----+----1`
`begin = '1:30't;`	
`put begin tod5.;`	`01:30`
`begin = 1566742823;`	
	`14:20:23`
`put begin tod9.;`	

See Also

Formats:

"HHMM*w.d* Format" on page 185

"TIME*w.d* Format" on page 245

"TIMEAMPM*w.d* Format" on page 247

Function:

"TIMEPART Function" on page 1162

Informat:

"TIME*w*. Informat" on page 1393

VAXRB*w.d* Format

Writes real binary (floating-point) data in VMS format.

Category: Numeric

Alignment: right

Syntax

VAXRB*w.d*

Syntax Description

w
specifies the width of the output field.

Default: 8

Range: 2-8

d
specifies the power of 10 by which to divide the value. This argument is optional.

Default: 0

Range: 0–31

Details

Use the VAXRB*w.d* format to write data in native VAX/VMS floating-point notation.

Comparisons

If you use SAS that is running under VAX/VMS, then the VAXRB*w.d* and the RB*w.d* formats are identical.

Example

```
x=1;
y=put(x,vaxrb8.);
put y=$hex16.;
```

Value of **x**	Results*
----+----1	
1	**8040000000000000**

* The result is the hexadecimal representation for the integer.

VMSZN*w.d* Format

Generates VMS and MicroFocus COBOL zoned numeric data.

Category: Numeric

Alignment: left

Syntax

VMSZN*w.d*

w
specifies the width of the output field

Default: 1

Range: 1–32

d
specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Details

The VMSZN*w.d* format is similar to the ZD*w.d* format. Both generate a string of ASCII digits, and the last digit is a special character that denotes the magnitude of the last digit and the sign of the entire number. The difference between these formats is in the special character that is used for the last digit. The following table shows the special characters that are used by the VMSZN*w.d* format.

Desired Digit	Special Character	Desired Digit	Special Character
0	0	–0	p
1	1	–1	q
2	2	–2	r
3	3	–3	s
4	4	–4	t
5	5	–5	u
6	6	–6	v
7	7	–7	w
8	8	–8	x
9	9	–9	y

Data formatted using the VMSZN*w.d* format are ASCII strings.

If the value to be formatted is too large to fit in a field of the specified width, then the VMSZN*w.d* format does the following:

- For positive values, it sets the output to the largest positive number that fits in the given width.
- For negative values, it sets the output to the negative number of greatest magnitude that fits in the given width.

Example

SAS Statements	Results
	`------+------1`
`x=1234;` `put x vmszn4.;`	`1234`
`x=1234;` `put x vmszn5.1;`	`12340`
`x=1234;` `put x vmszn6.2;`	`123400`
`-1234;` `put x vmszn5.;`	`0123t`

See Also

Format:

"ZD*w.d* Format" on page 284

Informat:

"VMSZN*w.d* Informat" on page 1399

w.d Format

Writes standard numeric data one digit per byte.

Category: Numeric

Alignment: right

Alias: F*w.d*

See: *w.d* Format in the documentation for your operating environment.

Syntax

w.d

Syntax Description

w
specifies the width of the output field.

Range: 1–32

Tip: Allow enough space to write the value, the decimal point, and a minus sign, if necessary.

d
specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Range: 0–31

Requirement: must be less than *w*

Tip: If *d* is 0 or you omit *d*, *w*.*d* writes the value without a decimal point.

Details

The *w*.*d* format rounds to the nearest number that fits in the output field. If *w*.*d* is too small, SAS might shift the decimal to the BEST*w*. format. The *w*.*d* format writes negative numbers with leading minus signs. In addition, *w*.*d* right aligns before writing and pads the output with leading blanks.

Comparisons

The Z*w*.*d* format is similar to the *w*.*d* format except that Z*w*.*d* pads right-aligned output with 0s instead of blanks.

Examples

```
put @7 x 6.3;
```

Value of **x**	Results
	`----+----1----+`
23.45	**23.450**

WEEKDATE*w*. Format

Writes date values as the day of the week and the date in the form *day-of-week*, *month-name dd*, *yy* (or *yyyy*).

Category: Date and Time

Alignment: right

Syntax

WEEKDATE*w*.

Syntax Description

w
specifies the width of the output field.

Default: 29

Range: 3–37

Details

The WEEKDATE*w*. format writes SAS date values in the form *day-of-week*, *month-name dd*, *yy* (or *yyyy*), where

dd
is an integer that represents the day of the month.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

If *w* is too small to write the complete day of the week and month, SAS abbreviates as needed.

Comparisons

The WEEKDATE*w*. format is the same as the WEEKDATX*w*. format except that WEEKDATX*w*. prints *dd* before the month's name.

Examples

The example table uses the input value of 16601 which is the SAS date value that corresponds to June 14, 2005.

SAS Statement	Results
	`----+----1----+----2`
`put date weekdate3.;`	`Tue`
`put date weekdate9.;`	`Tuesday`
`put date weekdate15.;`	`Tue, Jun 14, 05`
`put date weekdate17.;`	`Tue, Jun 14, 2005`

See Also

Formats:

"DATE*w*. Format" on page 151
"DDMMYY*w*. Format" on page 157
"MMDDYY*w*. Format" on page 196

"TOD*w.d* Format" on page 249
"WEEKDATX*w*. Format" on page 257
"YYMMDD*w*. Format" on page 273
Functions:
"JULDATE Function" on page 866
"MDY Function" on page 924
"WEEKDAY Function" on page 1229
Informats:
"DATE*w*. Informat" on page 1322
"DDMMYY*w*. Informat" on page 1326
"MMDDYY*w*. Informat" on page 1349
"YYMMDD*w*. Informat" on page 1410

WEEKDATX*w*. Format

Writes date values as the day of the week and date in the form *day-of-week*, *dd month-name yy* (or *yyyy*).

Category: Date and Time
Alignment: right

Syntax

WEEKDATX*w*.

Syntax Description

w
specifies the width of the output field.
Default: 29
Range: 3–37

Details

The WEEKDATX*w*. format writes SAS date values in the form *day-of-week*, *dd month-name*, *yy* (or *yyyy*), where

dd
is an integer that represents the day of the month.

yy or *yyyy*
is a two-digit or a four-digit integer that represents the year.

If *w* is too small to write the complete day of the week and month, then SAS abbreviates as needed.

Comparisons

The WEEKDATE*w*. format is the same as the WEEKDATX*w*. format, except that WEEKDATE*w*. prints *dd* after the month's name.

The WEEKDATX*w*. format is the same as the DTWKDATX*w*. format, except that DTWKDATX*w*. expects a datetime value as input.

Examples

The example table uses the input value of 16490, which is the SAS date value that corresponds to February 23, 2005.

SAS Statement	Results
	`----+----1----+----2----+----3`
`put date weekdatx.;`	`Wednesday, 23 February 2005`

See Also

Formats:

- "DTWKDATX*w*. Format" on page 166
- "DATE*w*. Format" on page 151
- "DDMMYY*w*. Format" on page 157
- "MMDDYY*w*. Format" on page 196
- "TOD*w.d* Format" on page 249
- "WEEKDATE*w*. Format" on page 255
- "YYMMDD*w*. Format" on page 273

Functions:

- "JULDATE Function" on page 866
- "MDY Function" on page 924
- "WEEKDAY Function" on page 1229

Informats:

- "DATE*w*. Informat" on page 1322
- "DDMMYY*w*. Informat" on page 1326
- "MMDDYY*w*. Informat" on page 1349
- "YYMMDD*w*. Informat" on page 1410

WEEKDAY*w*. Format

Writes date values as the day of the week.

Category: Date and Time

Alignment: right

Syntax

WEEKDAY*w*.

Syntax Description

w
specifies the width of the output field.

Default: 1

Range: 1–32

Details

The WEEKDAY*w*. format writes a SAS date value as the day of the week (where 1=Sunday, 2=Monday, and so on).

Examples

The example table uses the input value of 16469, which is the SAS date value that corresponds to February 2, 2005.

SAS Statement	Result
	`----+----1`
`put date weekday.;`	`4`

See Also

Format:

"DOWNAME*w*. Format" on page 163

WEEKU*w*. Format

Writes a week number in decimal format by using the U algorithm.

Category: Date and Time

Alignment: left

Syntax

WEEKU*w*.

Syntax Description

w
specifies the width of the output field.

Default: 11

Range: 3–200

Details

The WEEKU*w*. format writes a week-number format. The WEEKU*w*. format writes the various formats depending on the specified width. Algorithm U calculates the SAS date value by using the number of the week within the year (Sunday is considered the first day of the week). The number-of-the-week value is represented as a decimal number in the range 0–53, with a leading zero and maximum value of 53. For example, the fifth week of the year would be represented as 05.

Refer to the following table for widths, formats, and examples:

Widths	Formats	Examples
3-4	Www	w01
5-6	yyWww	03W01
7-8	yyWwwdd	03W0101
9-10	yyyyWwwdd	2003W0101
11-200	yyyy-Www-dd	2003-W01-01

Comparisons

The WEEKV*w*. format writes the week number as a decimal number in the range 01–53, with weeks beginning on a Monday and week 1 of the year including both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. The WEEKW*w*. format writes the week number of the year as a decimal number in the range 00–53, with Monday as the first day of week 1. The WEEKU*w*. format writes the week number of the year (Sunday as the first day of the week) as a decimal number in the range 0–53, with a leading zero.

Examples

```
sasdate = '01JAN2003'd;
```

Statements	Results
	`----+----1----+`
`v=put(sasdate,weeku3.);`	
`w=put(sasdate,weeku5.);`	
`x=put(sasdate,weeku7.);`	
`y=put(sasdate,weeku9.);`	
`z=put(sasdate,weeku11.);`	
`put v;`	`W00`
`put w;`	`03W00`
`put x;`	`03W0004`
`put y;`	`2003W0004`
`put z;`	`2003-W00-04`

See Also

Formats:

"WEEKV*w*. Format" on page 261
"WEEKW*w*. Format" on page 263

Functions:

"WEEK Function" on page 1226

Informats:

"WEEKU*w*. Informat" on page 1400
"WEEKV*w*. Informat" on page 1402
"WEEKW*w*. Informat" on page 1404

WEEKV*w*. Format

Writes a week number in decimal format by using the V algorithm.

Category: Date and Time
Alignment: left

Syntax

WEEKV*w*.

Syntax Description

w
specifies the width of the output field.

Default: 11

Range: 3–200

Details

The WEEKV*w*. format writes the various formats depending on the specified width. Algorithm V calculates the SAS date value, with the number-of-the-week value represented as a decimal number in the range 01–53, with a leading zero and maximum value of 53. Weeks begin on a Monday and week 1 of the year is the week that includes both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. For example, the fifth week of the year would be represented as 06.

Refer to the following table for widths, formats, and examples:

Widths	Formats	Examples
3-4	Www	w01
5-6	yyWww	03W01
7-8	yyWwwdd	03W0101
9-10	yyyyWwwdd	2003W0101
11-200	yyyy-Www-dd	2003-W01-01

Comparisons

The WEEKV*w*. format writes the week number as a decimal number in the range 01–53, with weeks beginning on a Monday and week 1 of the year including both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. The WEEKW*w*. format writes the week number of the year as a decimal number in the range 00–53, with Monday as the first day of week 1. The WEEKU*w*. format writes the week number of the year (Sunday as the first day of the week) as a decimal number in the range 0–53, with a leading zero.

Examples

```
sasdate='01JAN2003'd;
```

Statements	Results
	`----+----1----+`
`v=put(sasdate,weekv3.);`	
`w=put(sasdate,weekv5.);`	
`x=put(sasdate,weekv7.);`	
`y=put(sasdate,weekv9.);`	
`z=put(sasdate,weekv11.);`	
`put v;`	`W01`
`put w;`	`03W01`
`put x;`	`03W0103`
`put y;`	`2003W0103`
`put z;`	`2003-W01-03`

See Also

Formats:

"WEEKU*w*. Format" on page 259
"WEEKW*w*. Format" on page 263

Functions:

"WEEK Function" on page 1226

Informats:

"WEEKU*w*. Informat" on page 1400
"WEEKV*w*. Informat" on page 1402
"WEEKW*w*. Informat" on page 1404

WEEKW*w*. Format

Writes a week number in decimal format by using the W algorithm.

Category: Date and Time
Alignment: left

Syntax

WEEKW*w*.

Syntax Description

w
specifies the width of the output field.
Default: 11
Range: 3–200

Details

The WEEKW*w*. format writes the various formats depending on the specified width. Algorithm W calculates the SAS date value using the number of the week within the year (Monday is considered the first day of the week). The number-of-the-week value is represented as a decimal number in the range 0–53, with a leading zero and maximum value of 53. For example, the fifth week of the year would be represented as 05.

Refer to the following table for widths, formats, and examples:

Widths	Formats	Examples
3-4	Www	w01
5-6	yyWww	03W01
7-8	yyWwwdd	03W0101

Widths	Formats	Examples
9-10	yyyyWwwdd	2003W0101
11-200	yyyy-Www-dd	2003-W01-01

Comparisons

The WEEKV*w*. format writes the week number as a decimal number in the range 01–53. Weeks beginning on a Monday and on week 1 of the year include both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. The WEEKW*w*. format writes the week number of the year as a decimal number in the range 00–53, with Monday as the first day of week 1. The WEEKU*w*. format writes the week number of the year (Sunday as the first day of the week) as a decimal number in the range 0–53, with a leading zero.

Examples

```
sasdate = '01JAN2003'd;
```

Statements	Results
	`----+----1----+`
`v=put(sasdate,weekw3.);`	
`w=put(sasdate,weekw5.);`	
`x=put(sasdate,weekw7.);`	
`y=put(sasdate,weekw9.);`	
`z=put(sasdate,weekw11.);`	
`put v;`	`W03`
`put w;`	`03W03`
`put x;`	`03W0003`
`put y;`	`2003W0003`
`put z;`	`2003-W00-03`

See Also

Formats:

"WEEKU*w*. Format" on page 259
"WEEKV*w*. Format" on page 261

Functions:

"WEEK Function" on page 1226

Informats:

"WEEKU*w*. Informat" on page 1400
"WEEKV*w*. Informat" on page 1402
"WEEKW*w*. Informat" on page 1404

WORDDATE*w*. Format

Writes date values as the name of the month, the day, and the year in the form *month-name dd*, *yyyy*.

Category: Date and Time
Alignment: right

Syntax

WORDDATE*w*.

Syntax Description

w
specifies the width of the output field.
Default: 18
Range: 3–32

Details

The WORDDATE*w*. format writes SAS date values in the form *month-name dd*, *yyyy*, where

dd
is an integer that represents the day of the month.

yyyy
is a four-digit integer that represents the year.

If the width is too small to write the complete month, SAS abbreviates as necessary.

Comparisons

The WORDDATE*w*. format is the same as the WORDDATX*w*. format except that WORDDATX*w*. prints *dd* before the month's name.

Examples

The example table uses the input value of 16601, which is the SAS date value that corresponds to June 14, 2005.

SAS Statement	Result
	`----+----1----+----2`
`put term worddate3.;`	`Jun`
`put term worddate9.;`	`June`
`put term worddate12.;`	`Jun 14, 2005`
`put term worddate20.;`	`June 14, 2005`

See Also

Format:

"WORDDATX*w*. Format" on page 266

WORDDATX*w*. Format

Writes date values as the day, the name of the month, and the year in the form *dd month-name yyyy*.

Category: Date and Time

Alignment: right

Syntax

WORDDATX*w*.

Syntax Description

w
specifies the width of the output field.

Default: 18

Range: 3–32

Details

The WORDDATX*w*. format writes SAS date values in the form *dd month-name*, *yyyy*, where

dd
is an integer that represents the day of the month.

yyyy
is a four-digit integer that represents the year.

If the width is too small to write the complete month, SAS abbreviates as necessary.

Comparisons

The WORDDATX*w*. format is the same as the WORDDATE*w*. format except that WORDDATE*w*. prints *dd* after the month's name.

Examples

The example table uses the input value of 16500, which is the SAS date value that corresponds to March 5, 2005.

SAS Statement	Results
	`----+----1----+----2`
`put term worddatx.;`	`05 March 2005`

See Also

Format:

"WORDDATE*w*. Format" on page 265

WORDF*w*. Format

Writes numeric values as words with fractions that are shown numerically.

Category: Numeric
Alignment: left

Syntax

WORDF*w*.

Syntax Description

w
specifies the width of the output field.
Default: 10
Range: 5–32767

Details

The WORDF*w*. format converts numeric values to their equivalent in English words, with fractions that are represented numerically in hundredths. For example, 8.2 prints as eight and 20/100.

Negative numbers are preceded by the word minus. When the value's equivalent in words does not fit into the specified field, it is truncated on the right and the last character prints as an asterisk.

Comparisons

The WORDF*w*. format is similar to the WORDS*w*. format except that WORDF*w*. prints fractions as numbers instead of words.

Examples

```
put price wordf15.;
```

Value of **price**	Results
	`----+----1----+`
`2.5`	`two and 50/100`

See Also

Format:
"WORDS*w*. Format" on page 268

WORDS*w*. Format

Writes numeric values as words.

Category: Numeric
Alignment: left

Syntax

WORDS*w*.

Syntax Description

w
specifies the width of the output field.
Default: 10
Range: 5–32767

Details

You can use the WORDS*w*. format to print checks with the amount written out below the payee line.

Negative numbers are preceded by the word minus. If the number is not an integer, the fractional portion is represented as hundredths. For example, 5.3 prints as five and thirty hundredths. When the value's equivalent in words does not fit into the specified field, it is truncated on the right and the last character prints as an asterisk.

Comparisons

The WORDS*w*. format is similar to the WORDF*w*. format except that WORDS*w*. prints fractions as words instead of numbers.

Examples

```
put price words23.;
```

Value of **price**	Results
	----+----1----+----2----+
2.1	**two and ten hundredths**

See Also

Format:

"WORDF*w*. Format" on page 267

YEAR*w*. Format

Writes date values as the year.

Category: Date and Time

Alignment: right

Syntax

YEAR*w*.

Syntax Description

w

specifies the width of the output field.

Default: 4

Range: 2–32

Tip: If *w* is less than 4, the last two digits of the year print. Otherwise, the year value prints as four digits.

Comparisons

The YEAR*w*. format is similar to the DTYEAR*w*. format in that they both write date values. The difference is that YEAR*w*. expects a SAS date value as input, and DTYEAR*w*. expects a datetime value.

Examples

The example table uses the input value of 16601, which is the SAS date value that corresponds to June 14, 2005.

SAS Statement	Results
	`----+----1`
`put date year2.;`	`05`
`put date year4.;`	`2005`

See Also

Format:

"DTYEAR*w*. Format" on page 168

YYMM*w*. Format

Writes date values in the form <*yy*>*yy*M*mm*, where M is a character separator to indicate that the month number follows the M and the year appears as either 2 or 4 digits.

Category: Date and Time

Alignment: right

Syntax

YYMM*w*.

Syntax Description

w

specifies the width of the output field.

Default: 7

Range: 5–32

Interaction: When *w* has a value of 5 or 6, the date appears with only the last two digits of the year. When *w* is 7 or more, the date appears with a four-digit year.

Details

The YYMM*w*. format writes SAS date values in the form <*yy*>*yy*M*mm*, where

<*yy*>*yy*
is a two-digit or four-digit integer that represents the year.

M
is the character separator to indicate that the number of the month follows..

mm
is an integer that represents the month.

Examples

The following examples use the input value of 16734, which is the SAS date value that corresponds to October 25, 2005.

SAS Statement	Result
	`----+----1----+`
`put date yymm5.;`	`05M10`
`put date yymm6.;`	`05M10`
`put date yymm.;`	`2005M10`
`put date yymm7.;`	`2005M10`
`put date yymm10.;`	`2005M10`

See Also

Format:

"MMYY*w*. Format" on page 201
"YYMM*xw*. Format" on page 271

YYMM*xw*. Format

Writes date values in the form <*yy*>*yymm* or <*yy*>*yy-mm*, where the *x* in the format name is a character that represents the special character that separates the year and the month, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.

Category: Date and Time

Alignment: right

Syntax

YYMM*xw.*

Syntax Description

x

identifies a separator or specifies that no separator appear between the year and the month. Valid values for *x* are:

C

separates with a colon

D
separates with a dash

N
indicates no separator

P
separates with a period

S
separates with a forward slash.

w
specifies the width of the output field.

Default: 7

Range: 5–32

Interaction: When *x* is set to *N*, no separator is specified. The width range is then 4–32, and the default changes to 6.

Interaction: When *x* has a value of C, D, P, or S and *w* has a value of 5 or 6, the date appears with only the last two digits of the year. When *w* is 7 or more, the date appears with a four-digit year.

Interaction: When *x* has a value of *N* and *w* has a value of 4 or 5, the date appears with only the last two digits of the year. When *x* has a value of *N* and *w* is 6 or more, the date appears with a four-digit year.

Details

The YYMM*xw*. format writes SAS date values in the form <*yy*>*yymm* or <*yy*>*yy*X*mm*, where

<*yy*>*yy*
is a two-digit or four-digit integer that represents the year.

x
is a specified separator.

mm
is an integer that represents the month.

Examples

The following examples use the input value of 18031, which is the SAS date value that corresponds to May 14, 2009.

SAS Statement	Results
	`----+----1----+`
`put date yymmc5.;`	`09:05`
`put date yymmd.;`	`2009-05`
`put date yymmn4.;`	`0905`
`put date yymmp8.;`	`2009.05`
`put date yymms10.;`	`2009/05`

See Also

Format:

"MMYY*xw*. Format" on page 203
"YYMM*w*. Format" on page 270

YYMMDD*w*. Format

Writes date values in the form *yymmdd* or *<yy>yy-mm-dd*, where a dash is the separator and the year appears as either 2 or 4 digits.

Category: Date and Time
Alignment: right

Syntax

YYMMDD*w*.

Syntax Description

w
specifies the width of the output field.

Default: 8

Range: 2–10

Interaction: When *w* has a value of from 2 to 5, the date appears with as much of the year and the month as possible. When *w* is 7, the date appears as a two-digit year without dashes.

Details

The YYMMDD*w*. format writes SAS date values in the form *yymmdd* or *<yy>yy–mm–dd*, where

<yy>yy	is a two-digit or four-digit integer that represents the year.
–	is the separator.
mm	is an integer that represents the month.
dd	is an integer that represents the day of the month.

To format a date that has a four-digit year and no separators, use the YYMMDD*x*. format.

Examples

The following examples use the input value of 16529, which is the SAS date value that corresponds to April 3, 2005.

SAS Statement	Results
	`----+----1----+`
`put day yymmdd2.;`	`05`
`put day yymmdd3.;`	` 05`
`put day yymmdd4.;`	`0504`
`put day yymmdd5.;`	`05-04`
`put day yymmdd6.;`	`050403`
`put day yymmdd7.;`	` 050403`
`put day yymmdd8.;`	`05-04-03`
`put day yymmdd10.;`	`2005-04-03`

See Also

Formats:

“DATE*w*. Format” on page 151
“DDMMYY*w*. Format” on page 157
“MMDDYY*w*. Format” on page 196
“YYMMDD*xw*. Format” on page 274

Functions:

“DAY Function” on page 637
“MDY Function” on page 924
“MONTH Function” on page 936
“YEAR Function” on page 1233

Informats:

“DATE*w*. Informat” on page 1322
“DDMMYY*w*. Informat” on page 1326
“MMDDYY*w*. Informat” on page 1349

YYMMDD*xw*. Format

Writes date values in the form *yymmdd* or *<yy>yy-mm-dd*, where the *x* in the format name is a character that represents the special character which separates the year, month, and day. The special character can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.

Category: Date and Time
Alignment: right

Syntax

YYMMDD*xw.*

Syntax Description

x

identifies a separator or specifies that no separator appear between the year, the month, and the day. Valid values for *x* are:

B
separates with a blank

C
separates with a colon

D
separates with a dash

N
indicates no separator

P
separates with a period

S
separates with a slash.

w

specifies the width of the output field.

Default: 8

Range: 2–10

Interaction: When *w* has a value of from 2 to 5, the date appears with as much of the year and the month. When *w* is 7, the date appears as a two-digit year without separators.

Interaction: When *x* has a value of N, the width range is 2–8.

Details

The YYMMDD*xw*. format writes SAS date values in the form *yymmdd* or *<yy>yyxmmxdd*, where

<yy>yy
is a two-digit or four-digit integer that represents the year.

x
is a specified separator.

mm
is an integer that represents the month.

dd
is an integer that represents the day of the month.

Examples

The following examples use the input value of 18031, which is the SAS date value that corresponds to May 14, 2009.

SAS Statement	Results
	`----+----1----+`
`put day yymmddc5.;`	`09:05`
`put day yymmddd8.;`	`09-05-14`
`put day yymmddp10.;`	`2009.05.14`
`put day yymmddn8.;`	`20090514`

See Also

Formats:

"DATE*w*. Format" on page 151
"DDMMYY*xw*. Format" on page 158
"MMDDYY*xw*. Format" on page 198
"YYMMDD*w*. Format" on page 273

Functions:

"DAY Function" on page 637
"MDY Function" on page 924
"MONTH Function" on page 936
"YEAR Function" on page 1233

Informat:

"YYMMDD*w*. Informat" on page 1410

YYMON*w*. Format

Writes date values in the form *yymmm* or *yyyymmm*.

Category: Date and Time
Alignment: right

Syntax

YYMON*w*.

Syntax Description

w
specifies the width of the output field. If the format width is too small to print a four-digit year, only the last two digits of the year are printed.

Default: 7

Range: 5–32

Details

The YYMON*w*. format abbreviates the month's name to three characters.

Examples

The example table uses the input value of 16601, which is the SAS date value that corresponds to June 14, 2005.

SAS Statement	Results
	`----+----1`
`put date yymon6.;`	`05JUN`
`put date yymon7.;`	`2005JUN`

See Also

Format:

"MMYY*w*. Format" on page 201

YYQ*w*. Format

Writes date values in the form *<yy>yy*Q*q*, where Q is the separator, the year appears as either 2 or 4 digits, and *q* is the quarter of the year.

Category: Date and Time

Alignment: right

Syntax

YYQ*w*.

Syntax Description

w

specifies the width of the output field.

Default: 6

Range: 4–32

Interaction: When *w* has a value of 4 or 5, the date appears with only the last two digits of the year. When *w* is 6 or more, the date appears with a four-digit year.

Details

The YYQ*w*. format writes SAS date values in the form <*yy*>*yy*Q*q*, where

<*yy*>*yy*
is a two-digit or four-digit integer that represents the year.

Q
is the character separator.

q
is an integer (1,2,3, or 4) that represents the quarter of the year.

Examples

The following examples use the input value of 16601, which is the SAS date value that corresponds to June 14, 2005.

SAS Statements	Results
	`----+----1----+`
`put date yyq4.;`	`05Q2`
`put date yyq5.;`	`05Q2`
`put date yyq.;`	`2005Q2`
`put date yyq6.;`	`2005Q2`
`put date yyq10.;`	`2005Q2`

See Also

Formats:

"YYQ*xw*. Format" on page 278
"YYQR*w*. Format" on page 280

YYQ*xw*. Format

Writes date values in the form <*yy*>*yyq* or <*yy*>*yy-q*, where the *x* in the format name is a character that represents the special character that separates the year and the quarter or the year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits.

Category: Date and Time
Alignment: right

Syntax

YYQ*xw*.

Syntax Description

x

identifies a separator or specifies that no separator appear between the year and the quarter. Valid values for *x* are:

C
separates with a colon

D
separates with a dash

N
indicates no separator

P
separates with a period

S
separates with a forward slash.

w

specifies the width of the output field.

Default: 6

Range: 4–32

Interaction: When *x* is set to *N*, no separator is specified. The width range is then 3–32, and the default changes to 5.

Interaction: When *w* has a value of 4 or 5, the date appears with only the last two digits of the year. When *w* is 6 or more, the date appears with a four-digit year.

Interaction: When *x* has a value of *N* and *w* has a value of 3 or 4, the date appears with only the last two digits of the year. When *x* has a value of *N* and *w* is 5 or more, the date appears with a four-digit year.

Details

The YYQ*xw*. format writes SAS date values in the form <*yy*>*yyq* or <*yy*>*yyxq*, where

<*yy*>*yy*
is a two-digit or four-digit integer that represents the year.

x
is a specified separator.

q
is an integer (1,2,3, or 4) that represents the quarter of the year.

Examples

The following examples use the input value of 18031, which is the SAS date value that corresponds to July 14, 2005.

SAS Statement	Results
	`----+----1----+`
`put date yyqc4.;`	`09:2`
`put date yyqd.;`	`2009-2`

SAS Statement	Results
`put date yyqn3.;`	`092`
`put date yyqp6.;`	`2009.2`
`put date yyqs8.;`	`2009/2`

See Also

Formats:

"YYQ*w*. Format" on page 277
"YYQR*xw*. Format" on page 281

YYQR*w*. Format

Writes date values in the form <*yy*>*yy*Q*qr*, where Q is the separator, the year appears as either 2 or 4 digits, and *qr* is the quarter of the year expressed in roman numerals.

Category: Date and Time

Alignment: right

Syntax

YYQR*w*.

Syntax Description

w

specifies the width of the output field.

Default: 8

Range: 6–32

Interaction: When the value of *w* is too small to write a four-digit year, the date appears with only the last two digits of the year.

Details

The YYQR*w*. format writes SAS date values in the form <*yy*>*yy*Q*qr*, where

<*yy*>*yy*
is a two-digit or four-digit integer that represents the year.

Q
is the character separator.

qr
is a roman numeral (I, II, III, or IV) that represents the quarter of the year.

Examples

The following examples use the input value of 16601, which is the SAS date value that corresponds to June 14, 2005.

SAS Statement	Result
	`----+----1----+`
`put date yyqr6.;`	`05QII`
`put date yyqr7.;`	`2005QII`
`put date yyqr.;`	`2005QII`
`put date yyqr8.;`	`2005QII`
`put date yyqr10.;`	`2005QII`

See Also

Format:

"YYQ*w*. Format" on page 277
"YYQR*xw*. Format" on page 281

YYQR*xw*. Format

Writes date values in the form <*yy*>*yyqr* or <*yy*>*yy-qr*, where the *x* in the format name is a character that represents the special character that separates the year and the quarter or the year, which can be a hyphen (-), period (.), blank character, slash (/), colon (:), or no separator; the year can be either 2 or 4 digits and *qr* is the quarter of the year expressed in roman numerals.

Category: Date and Time
Alignment: right

Syntax

YYQR*xw.*

Syntax Description

x
identifies a separator or specifies that no separator appear between the year and the quarter. Valid values for *x* are:

C
separates with a colon

D

separates with a dash

N
indicates no separator

P
separates with a period

S
separates with a forward slash.

w
specifies the width of the output field.

Default: 8

Range: 6–32

Interaction: When *x* is set to *N*, no separator is specified. The width range is then 5–32, and the default changes to 7.

Interaction: When the value of *w* is too small to write a four-digit year, the date appears with only the last two digits of the year.

Details

The YYQR*xw*. format writes SAS date values in the form *<yy>yyqr* or *<yy>yyxqr*, where

<yy>yy
is a two-digit or four-digit integer that represents the year.

x
is a specified separator.

qr
is a roman numeral (I, II, III, or IV) that represents the quarter of the year.

Examples

The following examples use the input value of 18031, which is the SAS date value that corresponds to May 14, 2009.

SAS Statement	Result
	`----+----1----+`
`put date yyqrc6.;`	`09:II`
`put date yyqrd.;`	`2009-II`
`put date yyqrn5.;`	`09II`
`put date yyqrp8.;`	`2009.II`
`put date yyqrs10.;`	`2009/II`

See Also

Format:

"YYQ*xw*. Format" on page 278

"YYQR*w*. Format" on page 280

Z*w.d* Format

Writes standard numeric data with leading 0s.

Category: Numeric

Alignment: right

Syntax

Z*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 1

Range: 1–32

Tip: Allow enough space to write the value, the decimal point, and a minus sign, if necessary.

d

specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Default: 0

Range: 0–31

Tip: If *d* is 0 or you omit *d*, Z*w.d* writes the value without a decimal point.

Details

The Z*w.d* format writes standard numeric values one digit per byte and fills in 0s to the left of the data value.

The Z*w.d* format rounds to the nearest number that will fit in the output field. If *w.d* is too large to fit, SAS might shift the decimal to the BEST*w*. format. The Z*w.d* format writes negative numbers with leading minus signs. In addition, it right aligns before writing and pads the output with leading zeros.

Comparisons

The Z*w.d* format is similar to the *w.d* format except that Z*w.d* pads right-aligned output with 0s instead of blanks.

Examples

```
put @5 seqnum z8.;
```

Value of **seqnum**	Results
	`----+----1`
`1350`	`00001350`

ZD*w.d* Format

Writes numeric data in zoned decimal format .

Category: Numeric

Alignment: left

See: ZD*w.d* Format in the documentation for your operating environment.

Syntax

ZD*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 1

Range: 1–32

d

specifies to multiply the number by 10^d. This argument is optional.

Default: 0

Range: 0–31

Details

The zoned decimal format is similar to standard numeric format in that every digit requires one byte. However, the value's sign is in the last byte, along with the last digit.

Note: Different operating environments store zoned decimal values in different ways. However, the ZD*w.d* format writes zoned decimal values with consistent results if the values are created in the same type of operating environment that you use to run SAS. △

Comparisons

The following table compares the zoned decimal format with notation in several programming languages:

Language	Zoned Decimal Notation
SAS	ZD3.
PL/I	PICTURE '99T'
COBOL	DISPLAY PIC S 999
IBM 370 assembler	ZL3

Examples

```
y=put(x,zd4.);
put y $hex8.;
```

Value of **x**	Results
120	**F0F1F2C0**

* The result is a hexadecimal representation of a binary number in zoned decimal format on an IBM mainframe computer. Each byte occupies one column of the output field.

Formats Documented in Other SAS Publications

The main references for SAS formats are *SAS Language Reference: Dictionary* and the *SAS National Language Support (NLS): Reference Guide*. See the documentation for your operating environment for host-specific information about formats.

SAS National Language Support (NLS): Reference Guide

Table 3.5

Category	Formats for NLS	Description
BIDI text handling	$BIDIw. Format	Converts between a logically ordered string and a visually ordered string, by reversing the order of Hebrew and Arabic characters while preserving the order of Latin words and numbers.
	$LOGVSw.Format	Processes a character string that is in left-to-right-logical order, and then writes the character string in visual order.
	$LOGVSRw. Format	Processes a character string that is in right-to-left-logical order, and then writes the character string in visual order.

	$VSLOGw. Format	Processes a character string that is in visual order, and then writes the character string in left-to-right logical order.
	$VSLOGRw. Format	Processes a character string that is in visual order, and then writes the character string in right-to-left logical order.
Character	$UCS2Bw. Format	Processes a character string that is in the encoding of the current SAS session, and then writes the character string in big-endian, 16-bit, UCS2, Unicode encoding.
	$UCS2BEw. Format	Processes a character string that is in big-endian, 16-bit, UCS2, Unicode encoding, and then writes the character string in the encoding of the current SAS session.
	$UCS2Lw. Format	Processes a character string that is in the encoding of the current SAS session, and then writes the character string in little-endian, 16-bit, UCS2, Unicode encoding.
	$UCS2LEw. Format	Processes a character string that is in little-endian, 16-bit, UCS2, Unicode encoding, and then writes the character string in the encoding of the current SAS session.
	$UCS2Xw. Format	Processes a character string that is in the encoding of the current SAS session, and then writes the character string in native-endian, 16-bit, UCS2, Unicode encoding.
	$UCS2XEw. Format	Processes a character string that is in native-endian, 16-bit, UCS2, Unicode encoding, and then writes the character string in the encoding of the current SAS session.
	$UCS4Bw. Format	Processes a character string that is in the encoding of the current SAS session, and then writes the character string in big-endian, 32-bit, UCS4, Unicode encoding.
	$UCS4BEw. Format	Processes a character string that is in big-endian, 32-bit, UCS4, Unicode encoding, and then writes the character string in the encoding of the current SAS session.
	$UCS4Lw. Format	Processes a character string that is in the encoding of the current SAS session, and then writes the character string in little-endian, 32-bit, UCS4, Unicode encoding.
	%UCS4LEw. Format	Processes a character string that is in little-endian, 32-bit, UCS4, Unicode encoding, and then writes the character string in the encoding of the current SAS session.
	$UCS4Xw. Format	Processes a character string that is in the encoding of the current SAS session, and then writes the character string in native-endian, 32-bit, UCS4, Unicode encoding.
	$UCS4XEw. Format	Processes a character string that is in native-endian, 32-bit, UCS4, Unicode encoding, and then writes the character string in the encoding of the current SAS session.

	$UESCw. Format	Processes a character string that is encoded in the current SAS session, and then writes the character string in Unicode escape (UESC) representation.
	$UESCEw. Format	Processes a character string that is in Unicode escape (UESC) representation, and then writes the character string in the encoding of the current SAS session.
	$UNCRw. Format	Processes a character string that is encoded in the current SAS session, and then writes the character string in numeric character representation (NCR).
	$UNCREw. Format	Processes a character string that is in numeric character representation (NCR), and then writes the character string in the encoding of the current SAS session.
	$UPARENw. Format	Processes a character string that is encoded in the current SAS session, and then writes the character string in Unicode parenthesis (UPAREN) representation.
	$UPARENEw. Format	Processes a character string that is in Unicode parenthesis (UPAREN), and then writes the character string in the encoding of the current SAS session.
	$UTF8Xw. Format	Processes a character string that is in the encoding of the current SAS session, and then writes the character string in universal transformation format (UTF-8) encoding.
DBCS	$KANJIw. Format	Adds shift-code data to DBCS data.
	$KANJIXw. Format	Removes shift-code data from DBCS data.
Date and Time	HDATEw. Format	Writes date values in the form *yyyy mmmmm dd* where *dd* is the day-of-the-month, *mmmmm* represents the month's name in Hebrew, and *yyyy* is the year.
	HEBDATEw. Format	Writes date values according to the Jewish calendar.
	MINGUOw. Format	Writes date values as Taiwanese dates in the form *yyyymmdd* .
	NENGOw. Format	Writes date values as Japanese dates in the form *e.yymmdd* .
	NLDATEw. Format	Converts a SAS date value to the date value of the specified locale, and then writes the date value as a date.
	NLDATEMDw. Format	Converts the SAS date value to the date value of the specified locale, and then writes the value as the name of the month and the day of the month.
	NLDATEMNw. Format	Converts a SAS date value to the date value of the specified locale, and then writes the value as the name of the month.
	NLDATEWw. Format	Converts a SAS date value to the date value of the specified locale, and then writes the value as the date and the day of the week.
	NLDATEWNw. Format	Converts the SAS date value to the date value of the specified locale, and then writes the date value as the day of the week.

NLDATEYMw. Format	Converts the SAS date value to the date value of the specified locale, and then writes the date value as the year and the name of the month.
NLDATEYQw. Format	Converts the SAS date value to the date value of the specified locale, and then writes the date value as the year and the quarter.
NLDATEYRw. Format	Converts the SAS date value to the date value of the specified locale, and then writes the date value as the year.
NLDATEYWw. Format	Converts the SAS date value to the date value of the specified locale, and then writes the date value as the year and the week.
NLDATMw. Format	Converts a SAS datetime value to the datetime value of the specified locale, and then writes the value as a datetime.
NLDATMAPw. Format	Converts a SAS datetime value to the datetime value of the specified locale, and then writes the value as a datetime with a.m. or p.m.
NLDATMDTw. Format	Converts the SAS datetime value to the datetime value of the specified locale, and then writes the value as the name of the month, day of the month and year.
NLDATMMDw. Format	Converts the SAS datetime value to the datetime value of the specified locale, and then writes the value as the name of the month and the day of the month.
NLDATMMNw. Format	Converts the SAS datetime value to the datetime value of the specified locale, and then writes the value as the name of the month.
NLDATMTMw. Format	Converts the time portion of a SAS datetime value to the time-of-day value of the specified locale, and then writes the value as a time of day.
NLDATMWNw. Format	Converts a SAS datetime value to the datetime value of the specified locale, and then writes the value as the day of the week.
NLDATMWw. Format	Converts SAS datetime values to the locale sensitive datetime string as the day of the week and the datetime.
NLDATMYMw. Format	Converts the SAS datetime value to the datetime value of the specified locale, and then writes the value as the year and the name of the month.
NLDATMYQw. Format	Converts the SAS datetime value to the datetime value of the specified locale, and then writes the value as the year and the quarter of the year.
NLDATMYRw. Format	Converts the SAS datetime value to the datetime value of the specified locale, and then writes the value as the year.
NLDATMYWw. Format	Converts the SAS datetime value to the datetime value of the specified locale, and then writes the value as the year and the name of the week.

	NLTIMEw. Format	Converts a SAS time value to the time value of the specified locale, and then writes the value as a time value.
	NLTIMAPw. Format	Converts a SAS time value to the time value of a specified locale, and then writes the value as a time value with a.m. or p.m.
	WEEKUw. Format“WEEKU*w*. Format” on page 259	Writes a week number in decimal format by using the U algorithm.
	WEEKVw. Format“WEEKV*w*. Format” on page 261	Writes a week number in decimal format by using the V algorithm.
	WEEKWw.Format“WEEKW*w*. Format” on page 263	Writes a week number in decimal format by using the W algorithm.
	YYWEEKUw. Format	Writes a week number in decimal format by using the U algorithm, excluding day–of–the–week information.
	YYWEEKVw. Format	Writes a week number in decimal format by using the V algorithm, excluding day-of-the-week information.
	YYWEEKWw. Format	Writes a week number in decimal format by using the W algorithm, excluding the day-of-week information.
Hebrew text handling	$CPTDWw. Format	Processes a character string that is in Hebrew text, encoded in IBM-PC (cp862), and then writes the character string in Windows Hebrew encoding (cp 1255).
	$CPTWDw. Format	Processes a character string that is encoded in Windows (cp1255), and then writes the character string in Hebrew DOS (cp862) encoding.
Numeric	EUROw.d Format	Writes numeric values with a leading euro symbol (E), a comma that separates every three digits, and a period that separates the decimal fraction.
	EUROXw.d Format	Writes numeric values with a leading euro symbol (E), a period that separates every three digits, and a comma that separates the decimal fraction.
	NLBESTw. Format	Writes the best numerical notation based on the locale.
	NLMNIAEDw.d Format	Writes the monetary format of the international expression for the United Arab Emirates.
	NLMNIAUDw.d Format	Writes the monetary format of the international expression for Australia.
	NLMNIBGNw.d Format	Writes the monetary format of the international expression for Bulgaria.
	NLMNIBRLw.d Format	Writes the monetary format of the international expression for Brazil.
	NLMNICADw.d Format	Writes the monetary format of the international expression for Canada.
	NLMNICHFw.d Format	Writes the monetary format of the international expression for Liechtenstein and Switzerland.

NLMNICNYw.d Format	Writes the monetary format of the international expression for China.
NLMNICZKw.d Format	Writes the monetary format of the international expression for the Czech Republic.
NLMNIDKKw.d Format	Writes the monetary format of the local expression for Denmark, Faroe Island, and Greenland.
NLMNIEEKw.d Format	Writes the monetary format of the international expression for Estonia.
NLMNIEGPw.d Format	Writes the monetary format of the international expression for Egypt.
NLMNIEURw.d Format	Writes the monetary format of the international expression for Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, the Netherlands, Portugal, Slovenia, and Spain.
NLMNIGBPw.d Format	Writes the monetary format of the international expression for the United Kingdom.
NLMNIHKDw.d Format	Writes the monetary format of the international expression for Hong Kong.
NLMNIHRKw.d Format	Writes the monetary format of the international expression for Croatia.
NLMNIHUFw.d Format	Writes the monetary format of the international expression for Hungary.
NLMNIIDRw.d Format	Writes the monetary format of the international expression for Indonesia.
NLMNIILSw.d Format	Writes the monetary format of the international expression for Israel.
NLMNIINRw.d Format	Writes the monetary format of the international expression for India.
NLMNIJPYw.d Format	Writes the monetary format of the international expression for Japan.
NLMNIKRWw.d Format	Writes the monetary format of the international expression for South Korea.
NLMNILTLw.d Format	Writes the monetary format of the international expression for Lithuania.
NLMNILVLw.d Format	Writes the monetary format of the international expression for Latvia.
NLMNIMOPw.d Format	Writes the monetary format of the international expression for Macau.
NLMNIMXNw.d Format	Writes the monetary format of the international expression for Mexico.
NLMNIMYRw.d Format	Writes the monetary format of the international expression for Malaysia.
NLMNINOKw.d Format	Writes the monetary format of the international expression for Norway.

NLMNINZDw.d Format	Writes the monetary format of the international expression for New Zealand.
NLMNIPLNw.d Format	Writes the monetary format of the international expression for Poland.
NLMNIRUBw.d Format	Writes the monetary format of the international expression for Russia.
NLMNISEKw.d Format	Writes the monetary format of the international expression for Sweden.
NLMNISGDw.d Format	Writes the monetary format of the international expression for Singapore.
NLMNITHBw.d Format	Writes the monetary format of the international expression for Thailand.
NLMNITRYw.d Format	Writes the monetary format of the international expression for Turkey.
NLMNITWDw.d Format	Writes the monetary format of the international expression for Taiwan.
NLMNIUSDw.d Format	Writes the monetary format of the international expression for Puerto Rico and the United States.
NLMNIZARw.d Format	Writes the monetary format of the international expression for South Africa.
NLMNLAEDw.d Format	Writes the monetary format of the local expression for the United Arab Emirates.
NLMNLAUDw.d Format	Writes the monetary format of the local expression for Australia.
NLMNLBGNw.d Format	Writes the monetary format of the local expression for Bulgaria.
NLMNLBRLw.d Format	Writes the monetary format of the local expression for Brazil.
NLMNLCADw.d Format	Writes the monetary format of the local expression for Canada.
NLMNLCHFw.d Format	Writes the monetary format of the local expression for Liechtenstein and Switzerland.
NLMNLCNYw.d Format	Writes the monetary format of the local expression for China.
NLMNLCZKw.d Format	Writes the monetary format of the local expression for the Czech Republic.
NLMNLDKKw.d Format	Writes the monetary format of the local expression for Denmark, Faroe Island, and Greenland.
NLMNLEEKw.d Format	Writes the monetary format of the local expression for Estonia.
NLMNLEGPw.d Format	Writes the monetary format of the local expression for Egypt.

NLMNLEURw.d Format	Writes the monetary format of the local expression for Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, the Netherlands, Portugal, Slovenia, and Spain.
NLMNLGBPw.d Format	Writes the monetary format of the local expression for the United Kingdom.
NLMNLHKDw.d Format	Writes the monetary format of the local expression for Hong Kong.
NLMNLHRKw.d Format	Writes the monetary format of the local expression for Croatia.
NLMNLHUFw.d Format	Writes the monetary format of the local expression for Hungary.
NLMNLIDRw.d Format	Writes the monetary format of the local expression for Indonesia.
NLMNLILSw.d Format	Writes the monetary format of the local expression for Israel.
NLMNLINRw.d Format	Writes the monetary format of the local expression for India.
NLMNLJPYw.d Format	Writes the monetary format of the local expression for Japan.
NLMNLKRWw.d Format	Writes the monetary format of the local expression for South Korea.
NLMNLLTLw.d Format	Writes the monetary format of the local expression for Lithuania.
NLMNLLVLw.d Format	Writes the monetary format of the local expression for Latvia.
NLMNLMOPw.d Format	Writes the monetary format of the local expression for Macau.
NLMNLMXNw.d Format	Writes the monetary format of the local expression for Mexico.
NLMNLMYRw.d Format	Writes the monetary format of the local expression for Malaysia.
NLMNLNOKw.d Format	Writes the monetary format of the local expression for Norway.
NLMNLNZDw.d Format	Writes the monetary format of the local expression for New Zealand.
NLMNLPLNw.d Format	Writes the monetary format of the local expression for Poland.
NLMNLRUBw.d Format	Writes the monetary format of the local expression for Russia.
NLMNSEKw.d Format	Writes the monetary format of the local expression for Sweden.
NLMNLSGDw.d Format	Writes the monetary format of the local expression for Singapore.

NLMNLTHBw.d Format	Writes the monetary format of the local expression for Thailand.
NLMNLTRYw.d Format	Writes the monetary format of the local expression for Turkey.
NLMNLTWDw.d Format	Writes the monetary format of the local expression for Taiwan.
NLMNLUSDw.d Format	Writes the monetary format of the local expression for Puerto Rico, and the United States.
NLMNLZARw.d Format	Writes the monetary format of the local expression for South Africa.
NLMNYw.d Format	Writes the monetary format of the local expression in the specified locale using local currency.
NLMNYIw.d Format	Writes the monetary format of the international expression in the specified locale.
NLNUMw.d Format	Writes the numeric format of the local expression in the specified locale.
NLNUMIw.d Format	Writes the numeric format of the international expression in the specified locale.
NLPCTw.d Format	Writes percentage data of the local expression in the specified locale.
NLPCTIw.d Format	Writes percentage data of the international expression in the specified locale.
NLPCTNw.d Format	Produces percentages, using a minus sign for negative values.
NLPCTPw.d Format	Writes locale-specific numeric values as percentages.
NLPVALUEw.d Format	Writes p-values of the local expression in the specified locale.
NLSTRMONw.d Format	Writes a numeric value as a day-of-the-month in the specified locale.
NLSTRQTRw.d Format	Writes a numeric value as the quarter-of-the-year in the specified locale.
NLSTRWKw.d Format	Writes a numeric value as the day-of-the-week in the specified locale.
YENw.d	Writes numeric values with yen signs, commas, and decimal points.

CHAPTER

4

Functions and CALL Routines

Definitions of Functions and CALL Routines

Definition of Functions

A *SAS function* performs a computation or system manipulation on arguments, and returns a value that can be used in an assignment statement or elsewhere in expressions.

In Base SAS software, you can use SAS functions in DATA step programming statements, in a WHERE expression, in macro language statements, in PROC REPORT, and in Structured Query Language (SQL).

Some statistical procedures also use SAS functions. In addition, some other SAS software products offer functions that you can use in the DATA step. Refer to the documentation that pertains to the specific SAS software product for additional information about these functions.

Definition of CALL Routines

A *CALL routine* alters variable values or performs other system functions. CALL routines are similar to functions, but differ from functions in that you cannot use them in assignment statements or expressions.

All SAS CALL routines are invoked with CALL statements. That is, the name of the routine must appear after the keyword CALL in the CALL statement.

Syntax

Syntax of Functions

The syntax of a function has one of the following forms:

function-name (*argument-1*<, *...argument-n*>)

function-name (OF *variable-list*)

function-name (<*argument* | OF *variable-list* | OF *array-name*[*]><..., <*argument* | OF *variable-list* | OF *array-name*[*]>>)

where

function-name
names the function.

argument
can be a variable name, constant, or any SAS expression, including another function. The number and type of arguments that SAS allows are described with individual functions. Multiple arguments are separated by a comma.

Tip: If the value of an argument is invalid (for example, missing or outside the prescribed range), SAS writes a note to the log indicating that the argument is invalid, sets _ERROR_ to 1, and sets the result to a missing value.

Examples:

- `x=max(cash,credit);`
- `x=sqrt(1500);`
- `NewCity=left(upcase(City));`
- `x=min(YearTemperature-July,YearTemperature-Dec);`
- `s=repeat('----+',16);`
- `x=min((enroll-drop),(enroll-fail));`
- `dollars=int(cash);`
-
```
if sum(cash,credit)>1000 then
   put 'Goal reached';
```

variable-list
can be any form of a SAS variable list, including individual variable names. If more than one variable list appears, separate them with a space or with a comma and another OF.

Examples:

- `a=sum(of x  y  z);`
- The following two examples are equivalent.
  -
```
a=sum(of x1-x10 y1-y10 z1-z10);
a=sum(of x1-x10, of y1-y10, of z1-z10);
```
- `z=sum(of y1-y10);`
- `z=msplint(x0,5,of x1-x5,of y1-y5,-2,2);`

array-name{*}
names a currently defined array. Specifying an array with an asterisk as a subscript causes SAS to treat each element of the array as a separate argument.

The OF operator has been extended to accept temporary arrays. You can use temporary arrays in OF lists for most SAS functions just as you can use regular variable arrays, but there are some restrictions. For a list of these restrictions, see "Using the OF Operator with Temporary Arrays" on page 308.

Syntax of CALL Routines

The syntax of a CALL routine has one of the following forms:

CALL *routine-name* (*argument-1*<, *...argument-n*>);

CALL *routine-name* (OF *variable-list*);

CALL *routine-name* (*argument-1* | OF *variable-list-1* <, *...argument-n* | OF *variable-list-n*>);

where

routine-name
names a SAS CALL routine.

argument
can be a variable name, a constant, any SAS expression, an external module name, an array reference, or a function. Multiple arguments are separated by a comma. The number and type of arguments that are allowed are described with individual CALL routines in the dictionary section.

Examples:

- `call prxsubstr(prx,string,position);`
- `call prxchange('/old/new',1+k,trim(string),result,length);`
- `call set(dsid);`
- `call ranbin(Seed_1,n,p,X1);`
- `call label(abc{j},lab);`
- `call cats(result,'abc',123);`

variable-list
can be any form of a SAS variable list, including variable names. If more than one variable list appears, separate them with a space or with a comma and another OF.

Examples:

- `call cats(inventory, of y1-y15, of z1-z15);`
- `call catt(of item17-item23 pack17-pack23);`

Using Functions and CALL Routines

Restrictions Affecting Function Arguments

If the value of an argument is invalid, SAS writes a note or error message to the log and sets the result to a missing value. Here are some common restrictions for function arguments:

- Some functions require that their arguments be restricted within a certain range. For example, the argument of the LOG function must be greater than 0.
- When a numeric argument has a missing value, many functions write a note to the SAS log and return a missing value. Exceptions include some of the descriptive statistics functions and financial functions.
- For some functions, the allowed range of the arguments is platform-dependent, such as with the EXP function.

Using the OF Operator with Temporary Arrays

You can use the OF operator with temporary arrays. This capability enables the passing of temporary arrays to most functions whose arguments contain a varying number of parameters. You can use temporary arrays in OF lists in some functions, just as you can use temporary arrays in OF lists in regular variable arrays.

There are some limitations in using temporary arrays. These limitations are listed after the example.

The following example shows how you can use temporary arrays:

```
data _null_;
   array y[10] _temporary_ (1,2,3,4,5,6,7,8,9,10);
   x = sum(of y{*});
   put x=;
run;

data _null_;
   array y[10] $10 _temporary_ ('1','2','3','4','5',
                                '6','7','8','9','10');
   x = max(of y{*});
   put x=;
run;
```

Output 4.1 Log Output for the Example of Using Temporary Arrays

```
x=55
x=10
```

The following limitations affect temporary array OF lists:

- cannot be used as array indices
- can be used in functions where the number of parameters matches the number of elements in the OF list, as with regular variable arrays
- can be used in functions that take a varying number of parameters
- cannot be used with the DIF, LAG, SUBSTR, LENGTH, TRIM, or MISSING functions, nor with any of the variable information functions such as VLENGTH

Characteristics of Target Variables

Some character functions produce resulting variables, or *target variables*, with a default length of 200 bytes. Numeric target variables have a default length of 8 bytes. Character functions to which the default target variable lengths do not apply are shown in the following table. These functions obtain the length of the return argument based on the length of the first argument.

Table 4.1 Functions Whose Return Argument Is Based on the Length of the First Argument

Functions	
COMPBL	RIGHT
COMPRESS	STRIP
DEQUOTE	SUBSTR
INPUTC	SUBSTRN
LEFT	TRANSLATE
LOWCASE	TRIM

Functions	
PUTC	TRIMN
REVERSE	UPCASE

The following list of functions shows the length of the target variable if the target variable has not been assigned a length:

BYTE — target variable is assigned a default length of 1.

INPUT — length of the target variable is determined by the width of the informat.

PUT — length of the target variable is determined by the width of the format.

VTYPE — target variable is assigned a default length of 1.

VTYPEX — target variable is assigned a default length of 1.

Notes about Descriptive Statistic Functions

SAS provides functions that return descriptive statistics. Many of these functions correspond to the statistics produced by the MEANS and UNIVARIATE procedures. The computing method for each statistic is discussed in the elementary statistics procedures section of the *Base SAS Procedures Guide*. SAS calculates descriptive statistics for the nonmissing values of the arguments.

Notes about Financial Functions

SAS provides a group of functions that perform financial calculations. The functions are grouped into the following types:

Table 4.2 Types of Financial Functions

Function Type	Functions	Description
Cashflow	CONVX, CONVXP	calculates convexity for cashflows
	DUR, DURP	calculates modified duration for cashflows
	PVP, YIELDP	calculates present value and yield-to-maturity for a periodic cashflow
Parameter calculations	COMPOUND	calculates compound interest parameters
	MORT	calculates amortization parameters
Internal rate of return	INTRR, IRR	calculates the internal rate of return
Net present and future value	NETPV, NPV	calculates net present and future values
	SAVING	calculates the future value of periodic saving
Depreciation	DACCxx	calculates the accumulated depreciation up to the specified period
	DEPxxx	calculates depreciation for a single period

Function Type	Functions	Description
Pricing	BLKSHCLPRC, BLKSHPTPRC	calculates call prices and put prices for European options on stocks, based on the Black-Scholes model
	BLACKCLPRC, BLACKPTPRC	calculates call prices and put prices for European options on futures, based on the Black model
	GARKHCLPRC, GARKHPTPRC	calculates call prices and put prices for European options on stocks, based on the Garman-Kohlhagen model
	MARGRCLPRC, MARGRPTPRC	calculates call prices and put prices for European options on stocks, based on the Margrabe model

Using Pricing Functions

A pricing model is used to calculate a theoretical market value (price) for a financial instrument. This value is referred to as a mark-to-market (MTM) value. Typically, a pricing function has the following form:

$$price = function\,(rf1, rf2, rf3, ...)$$

In the pricing function, *rf1*, *rf2*, and *rf3* are risk factors such as interest rates or foreign exchange rates. The specific values of the risk factors that are used to calculate the MTM value are the base case values. The set of base case values is known as the base case market state.

After determining the MTM value, you can perform the following tasks with the base case values of the risk factors (*rf1*, *rf2*, and *rf3*):

- Set the base case values to specific values to perform scenario analyses.
- Set the base case values to a range of values to perform profit/loss curve analyses and profit/loss surface analyses.
- Automatically set the base case values to different values to calculate sensitivities—that is, to calculate the delta and gamma values of the risk factors.
- Perturb the base case values to create many possible market states so that many possible future prices can be calculated, and simulation analyses can be performed. For Monte Carlo simulation, the values of the risk factors are generated using mathematical models and the copula methodology.

A list of pricing functions and their descriptions are included in Table 4.2 on page 310.

Using DATA Step Functions within Macro Functions

The macro functions %SYSFUNC and %QSYSFUNC can call most DATA step functions to generate text in the macro facility. %SYSFUNC and %QSYSFUNC have one difference: %QSYSFUNC masks special characters and mnemonics and %SYSFUNC does not. For more information about these functions, see %QSYSFUNC and %SYSFUNC in *SAS Macro Language: Reference*.

%SYSFUNC arguments are a single DATA step function and an optional format, as shown in the following examples:

```
%sysfunc(date(),worddate.)
%sysfunc(attrn(&dsid,NOBS))
```

You cannot nest DATA step functions within %SYSFUNC. However, you can nest %SYSFUNC functions that call DATA step functions. For example:

```
%sysfunc(compress(%sysfunc(getoption(sasautos)),
   %str(%)%(%')));
```

All arguments in DATA step functions within %SYSFUNC must be separated by commas. You cannot use argument lists that are preceded by the word OF.

Because %SYSFUNC is a macro function, you do not need to enclose character values in quotation marks as you do in DATA step functions. For example, the arguments to the OPEN function are enclosed in quotation marks when you use the function alone, but the arguments do not require quotation marks when used within %SYSFUNC.

```
dsid=open("sasuser.houses","i");
dsid=open("&mydata","&mode");
%let dsid=%sysfunc(open(sasuser.houses,i));
%let dsid=%sysfunc(open(&mydata,&mode));
```

Using CALL Routines and the %SYSCALL Macro Statement

When the %SYSCALL macro statement invokes a CALL routine, the value of each macro variable argument is retrieved and passed *unresolved* to the CALL routine. Upon completion of the CALL routine, the value for each argument is written back to the respective macro variable. If %SYSCALL encounters an error condition, the execution of the CALL routine terminates without updating the macro variable values and an error message is written to the log.

When %SYSCALL invokes a CALL routine, the argument value is passed unresolved to the CALL routine. The unresolved argument value might have been quoted using macro quoting functions and might contain delta characters. The argument value in its quoted form can cause unpredictable results when character values are compared. Some CALL routines unquote their arguments when they are called by %SYSCALL and return the unquoted values. Other CALL routines do not need to unquote their arguments. The following is a list of CALL routines that unquote their arguments when called by %SYSCALL:

- “CALL COMPCOST Routine” on page 447
- “CALL LEXCOMB Routine” on page 458
- “CALL LEXPERK Routine” on page 465
- “CALL LEXPERM Routine” on page 469
- “CALL PRXCHANGE Routine” on page 479
- “CALL PRXNEXT Routine” on page 485
- “CALL PRXSUBSTR Routine” on page 490
- “CALL SCAN Routine” on page 516
- “CALL SORTC Routine” on page 528
- “CALL STDIZE Routine” on page 531
- “CALL SYSTEM Routine” on page 538

In comparison, %SYSCALL invokes a CALL routine and returns an unresolved value, which contains delta characters. %SYSFUNC invokes a function and returns a resolved value, which does not contain delta characters. For more information, see “How Macro Quoting Works”, “%SYSCALL Macro Statement”, and “%SYSFUNC Macro Function” in *SAS Macro Language: Reference*.

Using Functions to Manipulate Files

SAS manipulates files in different ways, depending on whether you use functions or statements. If you use functions such as FOPEN, FGET, and FCLOSE, you have more opportunity to examine and manipulate your data than when you use statements such as INFILE, INPUT, and PUT.

When you use external files, the FOPEN function allocates a buffer called the File Data Buffer (FDB) and opens the external file for reading or updating. The FREAD function reads a record from the external file and copies the data into the FDB. The FGET function then moves the data to the DATA step variables. The function returns a value that you can check with statements or other functions in the DATA step to determine how to further process your data. After the records are processed, the FWRITE function writes the contents of the FDB to the external file, and the FCLOSE function closes the file.

When you use SAS data sets, the OPEN function opens the data set. The FETCH and FETCHOBS functions read observations from an open SAS data set into the Data Set Data Vector (DDV). The GETVARC and GETVARN functions then move the data to DATA step variables. The functions return a value that you can check with statements or other functions in the DATA step to determine how you want to further process your data. After the data is processed, the CLOSE function closes the data set.

For a complete listing of functions and CALL routines, see "Functions and CALL Routines by Category" on page 345. For complete descriptions and examples, see the dictionary section of this book.

Function Compatibility with SBCS, DBCS, and MBCS Character Sets

Overview

SAS string functions and CALL routines can be categorized by level numbers that are used in internationalization. I18N is the abbreviation for internationalization, and indicates string functions that can be adapted to different languages and locales without program changes.

I18N recognizes the following three levels that identify the character sets that you can use:

- "I18N Level 0" on page 313
- "I18N Level 1" on page 314
- "I18N Level 2" on page 314

For more information about function compatibility, see K Functions Compatibility in the *SAS National Language Support (NLS): Reference Guide*.

I18N Level 0

I18N Level 0 functions are designed for use with Single Byte Character Sets (SBCS) only.

I18N Level 1

I18N Level 1 functions should be avoided, if possible, if you are using a non-English language. The I18N Level 1 functions might not work correctly with Double Byte Character Set (DBCS) or Multi-Byte Character Set (MBCS) encodings under certain circumstances.

I18N Level 2

I18N Level 2 functions are designed for use with SBCS, DBCS, and MBCS (UTF8).

Using Random-Number Functions and CALL Routines

Types of Random-Number Functions

Two types of random-number functions are available in SAS. The newest random-number function is the RAND function. It uses the Mersenne-Twister pseudo-random number generator (RNG) that was developed by Matsumoto and Nishimura (1998). This RNG has a very long period of $2^{19937} - 1$, and has very good statistical properties. (A period is the number of occurrences before the pseudo-random number sequence repeats.)

The RAND function is started with a single seed. However, the state of the process cannot be captured by a single seed, which means that you cannot stop and restart the generator from its stopping point. Use the STREAMINIT function to produce a sequence of values that begins at the beginning of a stream. For more information, see the Details section of the "RAND Function" on page 1069.

The older random-number generators include the UNIFORM, NORMAL, RANUNI, RANNOR, and other functions that begin with RAN. These functions have a period of only $2^{31} - 2$ or less. The pseudo-random number stream is started with a single seed, and the state of the process can be captured in a new seed. This means that you can stop and restart the generator from its stopping point by providing the proper seed to the corresponding CALL routines. You can use the random-number functions to produce a sequence of values that begins in the middle of a stream.

Seed Values

Random-number functions and CALL routines generate streams of pseudo-random numbers from an initial starting point, called a *seed*, that either the user or the computer clock supplies. A seed must be a nonnegative integer with a value less than $2^{31}–1$ (or 2,147,483,647). If you use a positive seed, you can always replicate the stream of random numbers by using the same DATA step. If you use zero as the seed, the computer clock initializes the stream, and the stream of random numbers cannot be replicated.

Understanding How Functions Generate a Random-Number Stream

Using the DATA Step to Generate a Single Stream of Random Numbers

The DATA steps in this section illustrate several properties of the random-number functions. Each of the DATA steps that call a function generates a single stream of pseudo-random numbers based on a seed value of 7, because that is the first seed for the first call for every step. Some of the DATA steps change the seed value in various ways. Some of the steps have single function calls and others have multiple function calls. None of these DATA steps change the seed. The only seed that is relevant to the function calls is the seed that was used with the first execution of the first random-number function. There is no way to create separate streams with functions (CALL routines are used for this purpose), and the only way you can restart the function random-number stream is to start a new DATA step.

The following example executes multiple DATA steps:

```
options nodate pageno=1 linesize=80 pagesize=60;

   /* This DATA step produces a single stream of random numbers */
   /* based on a seed value of 7.                               */
data a;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
   a = ranuni (7); output;
run;

   /* This DATA step uses a DO statement to produce a single */
   /* stream of random numbers based on a seed value of 7.   */
data b (drop = i);
   do i = 7 to 18;
      b = ranuni (i);
      output;
   end;
run;

   /* This DATA step uses a DO statement to produce a single */
   /* stream of random numbers based on a seed value of 7.   */
data c (drop = i);
   do i = 1 to 12;
      c = ranuni (7);
      output;
   end;
run;
```

```
   /* This DATA step calls the RANUNI and the RANNOR functions */
   /* and produces a single stream of random numbers based on  */
   /* a seed value of 7.                                       */
data d;
   d = ranuni (7); f = ' '; output;
   d = ranuni (8); f = ' '; output;
   d = rannor (9); f = 'n'; output;
   d = .;          f = ' '; output;
   d = ranuni (0); f = ' '; output;
   d = ranuni (1); f = ' '; output;
   d = rannor (2); f = 'n'; output;
   d = .;          f = ' '; output;
   d = ranuni (3); f = ' '; output;
   d = ranuni (4); f = ' '; output;
   d = rannor (5); f = 'n'; output;
   d = .;          f = ' '; output;
run;

   /* This DATA step calls the RANNOR function and produces a     */
   /* single stream of random numbers based on a seed value of 7. */
data e (drop = i);
   do i = 1 to 6;
      e = rannor (7); output;
      e = .;          output;
   end;
run;

   /* This DATA step merges the output data sets that were */
   /* created from the previous five DATA steps.           */
data five;
   merge a b c d e;
run;

   /* This procedure writes the output from the merged data sets. */
proc print label data=five;
   options missing = ' ';
   label f = '00'x;
   title 'Single Random Number Streams';
run;
```

The following output shows the program results.

Output 4.2 Results from Generating a Single Random-Number Stream

Single Random Number Streams 1

Obs	a	b	c	d		e
1	0.29474	0.29474	0.29474	0.29474		0.39464
2	0.79062	0.79062	0.79062	0.79062		
3	0.79877	0.79877	0.79877	0.26928	n	0.26928
4	0.81579	0.81579	0.81579			
5	0.45122	0.45122	0.45122	0.45122		0.27475
6	0.78494	0.78494	0.78494	0.78494		
7	0.80085	0.80085	0.80085	-0.11729	n	-0.11729
8	0.72184	0.72184	0.72184			
9	0.34856	0.34856	0.34856	0.34856		-1.41879
10	0.46597	0.46597	0.46597	0.46597		
11	0.73523	0.73523	0.73523	-0.39033	n	-0.39033
12	0.66709	0.66709	0.66709			

The pseudo-random number streams in output data sets A, B, and C are identical. The stream in output data set D mixes calls to the RANUNI and the RANNOR functions. In observations 1, 2, 5, 6, 9, and 10, the values that are returned by RANUNI exactly match the values in the previous streams. Observations 3, 7, and 11, which are flagged by "n", contain the values that are returned by the RANNOR function. The mix of the function calls does not affect the generation of the pseudo-random number stream. All of the results are based on a single stream of uniformly distributed values, some of which are transformed and returned from other functions such as RANNOR. The results of the RANNOR function are produced from two internal calls to RANUNI. The DATA step that creates output data set D executes the following steps three times to create 12 observations:

- call to RANUNI
- call to RANUNI
- call to RANNOR (which internally calls RANUNI twice)
- skipped line to compensate for the second internal call to RANUNI

In the DATA step that creates data set E, RANNOR is called six times, each time skipping a line to compensate for the fact that two internal calls to RANUNI are made for each call to RANNOR. Notice that the three values that are returned from RANNOR in the DATA step that creates data set D match the corresponding values in data set E.

Using the %SYSFUNC Macro to Generate a Single Stream of Random Numbers

When the RANUNI function is called through the macro language by using %SYSFUNC, one pseudo-random number stream is created. You cannot change the seed value unless you close SAS and start a new SAS session. The %SYSFUNC macro produces the same pseudo-random number stream as the DATA steps that generated the data sets A, B, and C for the first macro invocation only. Any subsequent macro calls produce a continuation of the single stream.

```
%macro ran;
   %do i = 1 %to 12;
      %let x = %sysfunc (ranuni (7));
      %put &x;
   %end;
   %mend;

%ran;
```

SAS writes the following output to the log:

Output 4.3 Results of Execution with the %SYSFUNC Macro

```
10   %macro ran;
11      %do i = 1 %to 12;
12         %let x = %sysfunc (ranuni (7));
13         %put &x;
14      %end;
15   %mend;
16   %ran;
0.29473798875451
0.79062100955779
0.79877014262544
0.81579051763554
0.45121804506109
0.78494144826426
0.80085421204606
0.72184205973606
0.34855818345609
0.46596586120592
0.73522999404707
0.66709365028287
```

Comparison of Seed Values in Random-Number Functions and CALL Routines

Each random-number function and CALL routine generates pseudo-random numbers from a specific statistical distribution. Each random-number function requires a seed value expressed as an integer constant or a variable that contains the integer constant. Each CALL routine calls a variable that contains the seed value. Additionally, every CALL routine requires a variable that contains the generated pseudo-random numbers.

The seed variable must be initialized before the first execution of the function or CALL routine. After each execution of a function, the current seed is updated internally, but the value of the seed argument remains unchanged. However, after each iteration of the CALL routine the seed variable contains the current seed in the stream that generates the next pseudo-random number. With a function, it is not possible to control the seed values, and, therefore, the pseudo-random numbers after the initialization.

Except for the NORMAL and UNIFORM functions, which are equivalent to the RANNOR and RANUNI functions, respectively, SAS provides a CALL routine that has the same name as each random-number function. Using CALL routines gives you greater control over the seed values.

Generating Multiple Streams from Multiple Seeds in Random-Number CALL Routines

Overview of Random-Number CALL Routines and Streams

You can use the random-number CALL routines to generate multiple streams of pseudo-random numbers within a single DATA step. If you supply a different seed value to initialize each of the seed variables, the streams of the generated pseudo-random numbers are computationally independent, but they might not be statistically independent unless you select the seed values carefully.

Note: Although you can create multiple streams with multiple seeds, this practice is not recommended. It is always safer to create a single stream. With multiple streams, as the streams become longer, the chances of the stream overlapping increase. △

The following two examples deliberately select seeds to illustrate worst-case scenarios. The examples show how to produce multiple streams by using multiple seeds. Although this practice is not recommended, you can use the random-number CALL routines with multiple seeds.

Example 1: Using Multiple Seeds to Generate Multiple Streams

This example shows that you can use multiple seeds to generate multiple streams of pseudo-randomly distributed values by using the random-number CALL routines. The first DATA step creates a data set with three variables that are normally distributed. The second DATA step creates variables that are uniformly distributed. The SGSCATTER procedure (see the *SAS ODS Graphics: Procedures Guide*) is used to show the relationship between each pair of variables for each of the two distributions.

```
options pageno = 1 nodate ls = 80 ps = 64;

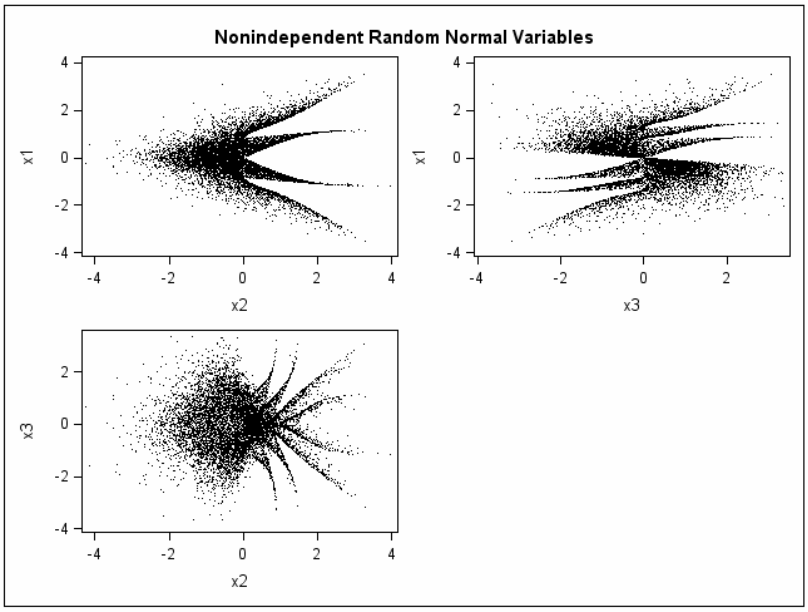
data normal;
   seed1 = 11111;
   seed2 = 22222;
   seed3 = 33333;
   do i = 1 to 10000;
      call rannor(seed1, x1);
      call rannor(seed2, x2);
      call rannor(seed3, x3);
      output;
   end;
run;

data uniform;
   seed1 = 11111;
   seed2 = 22222;
   seed3 = 33333;
   do i = 1 to 10000;
      call ranuni(seed1, x1);
      call ranuni(seed2, x2);
      call ranuni(seed3, x3);
      output;
   end;
run;

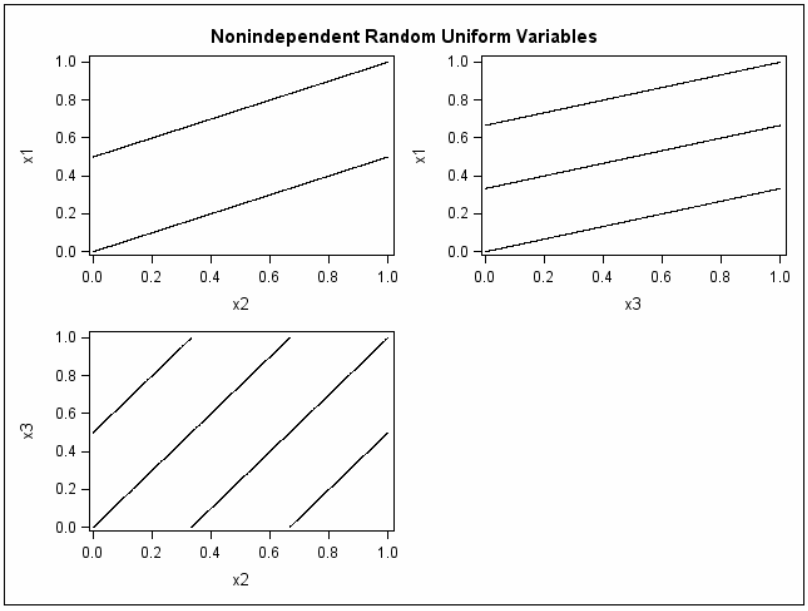
proc sgscatter data = normal;
   title 'Nonindependent Random Normal Variables';
   plot x1*x2 x1*x3 x3*x2 / markerattrs = (size = 1);
run;

proc sgscatter data = uniform;
   title 'Nonindependent Random Uniform Variables';
   plot x1*x2 x1*x3 x3*x2 / markerattrs = (size = 1);
run;
```

Display 4.1 Multiple Streams from Multiple Seeds: Nonindependent Random Normal Variables



Display 4.2 Multiple Streams from Multiple Seeds: Nonindependent Random Uniform Variables



The first plot (Display 4.1 on page 321) shows that normal variables appear to be linearly uncorrelated, but they are obviously not independent. The second plot (Display 4.2 on page 321) shows that uniform variables are clearly related. With this class of random-number generators, there is never any guarantee that the streams will be independent.

Example 2: Using Different Seeds with the CALL RANUNI Routine

The following example uses three different seeds and the CALL RANUNI routine to produce multiple streams.

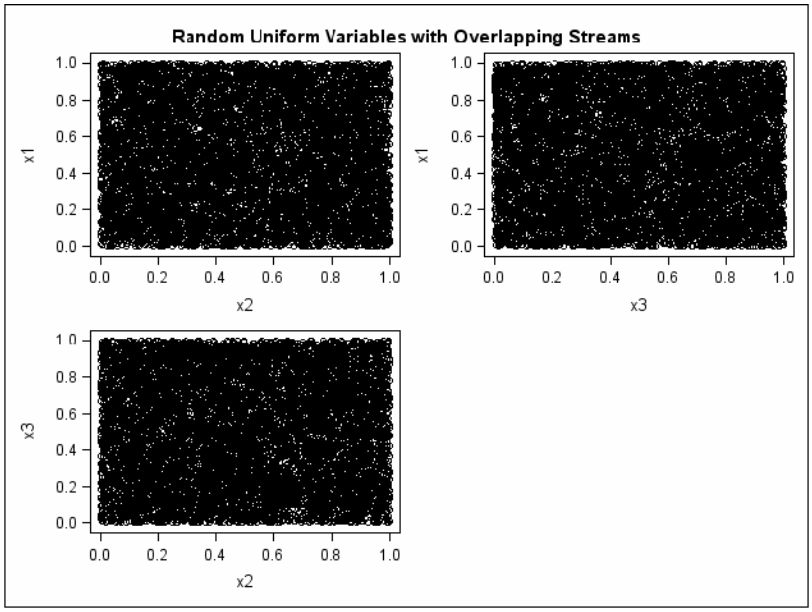
```
data uniform(drop=i);
   seed1 = 255793849;
   seed2 =1408147117;
   seed3 = 961782675;
   do i=1 to 10000;
      call ranuni(seed1, x1);
      call ranuni(seed2, x2);
      call ranuni(seed3, x3);
      i2 = lag(x2);
      i3 = lag2(x3);
      output;
   end;
label i2='Lag(x2)' i3='Lag2(x3)';
run;

title 'Random Uniform Variables with Overlapping Streams';
proc sgscatter data=uniform;
   plot x1*x2 x1*x3 x3*x2 / markerattrs = (size = 1);
run;

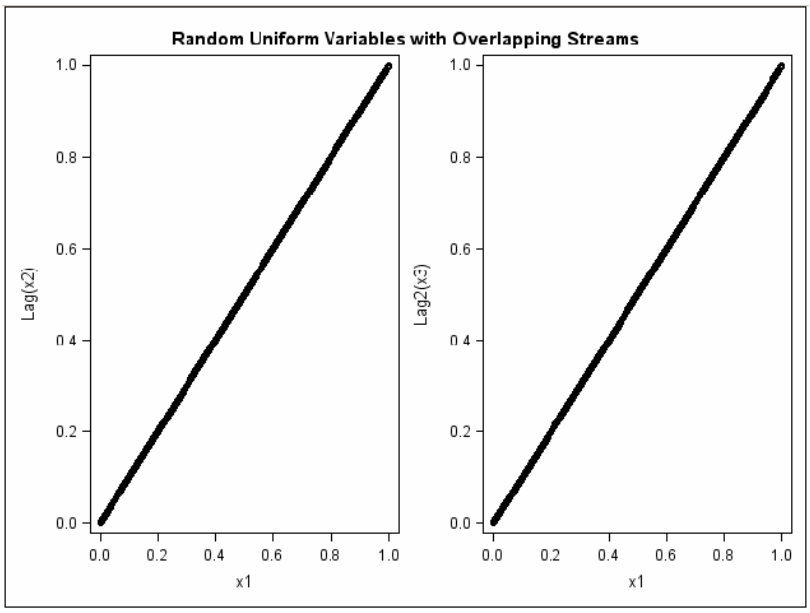
proc sgscatter data=uniform;
   plot i2*x1 i3*x1 / markerattrs = (size = 1);
run;

proc print noobs data=uniform(obs=10);
run;
```

Display 4.3 Using Different Seeds with CALL RANUNI: Random Uniform Variables with Overlapping Streams, Plot 1



Display 4.4 Using Different Seeds with CALL RANUNI: Random Uniform Variables with Overlapping Streams, Plot 2



Output 4.4 Random Uniform Variables with Overlapping Streams

Random Uniform Variables with Overlapping Streams 2

seed1	seed2	seed3	x1	x2	x3	i2	i3
1408147117	961782675	383001085	0.65572	0.44786	0.17835	.	.
961782675	383001085	1989090982	0.44786	0.17835	0.92624	0.44786	.
383001085	1989090982	1375749095	0.17835	0.92624	0.64063	0.17835	0.17835
1989090982	1375749095	89319994	0.92624	0.64063	0.04159	0.92624	0.92624
1375749095	89319994	1345897251	0.64063	0.04159	0.62673	0.64063	0.64063
89319994	1345897251	561406336	0.04159	0.62673	0.26143	0.04159	0.04159
1345897251	561406336	1333490358	0.62673	0.26143	0.62095	0.62673	0.62673
561406336	1333490358	963442111	0.26143	0.62095	0.44864	0.26143	0.26143
1333490358	963442111	1557707418	0.62095	0.44864	0.72536	0.62095	0.62095
963442111	1557707418	137842443	0.44864	0.72536	0.06419	0.44864	0.44864

The first plot (Display 4.3 on page 322) shows expected results: the variables appear to be statistically independent. However, the second plot (Display 4.4 on page 323) and the listing of the first 10 observations show that there is almost complete overlap between the two streams. The last 9999 values in x1 match the first 9999 values in x2, and the last 9998 values in x1 match the first 9998 values in x3. In other words, there is perfect agreement between the non-missing parts of x1 and lag(x2) and also x1 and lag2(x3). Even if the streams appear to be independent at first glance as in the first plot, there might be overlap, which might be undesirable depending on how the streams are used.

In practice, if you make multiple small streams with separate and randomly selected seeds, you probably will not encounter the problems that are shown in the first two

examples. “Example 2: Using Different Seeds with the CALL RANUNI Routine” on page 322 deliberately selects seeds to illustrate worst-case scenarios.

It is always safer to create a single stream. With multiple streams, as the streams get longer, the chances of the streams overlapping increase.

Generating Multiple Variables from One Seed in Random-Number Functions

Overview of Functions and Streams

If you use functions in your program, you cannot generate more than one stream of pseudo-random numbers by supplying multiple seeds within a DATA step.

The following example uses the RANUNI function to show the safest way to create multiple variables from the same stream with a single seed.

Example: Generating Random Uniform Variables with Overlapping Streams

In the following example, the RANUNI function is used to create random uniform variables with overlapping streams. The example shows the safest way to create multiple variables by using the RANUNI function. All variables are created from the same stream with a single seed.

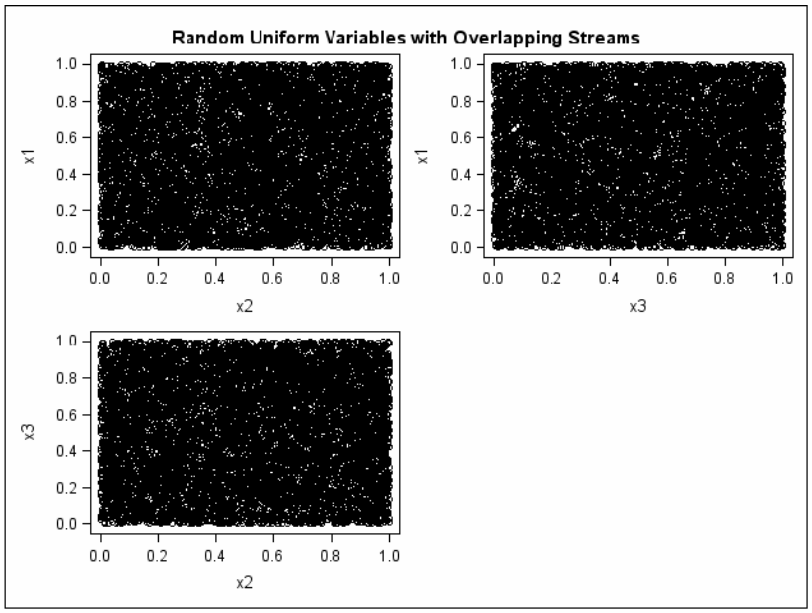
```
options pageno=1 nodate ls=80 ps=64;

data uniform(drop=i);
   do i = 1 to 10000;
      x1 = ranuni(11111);
      x2 = ranuni(11111);
      x3 = ranuni(11111);
      i2 = lag(x2);
      i3 = lag2(x3);
      output;
   end;
label i2 = 'Lag(x2)' i3 = 'Lag2(x3)';
run;

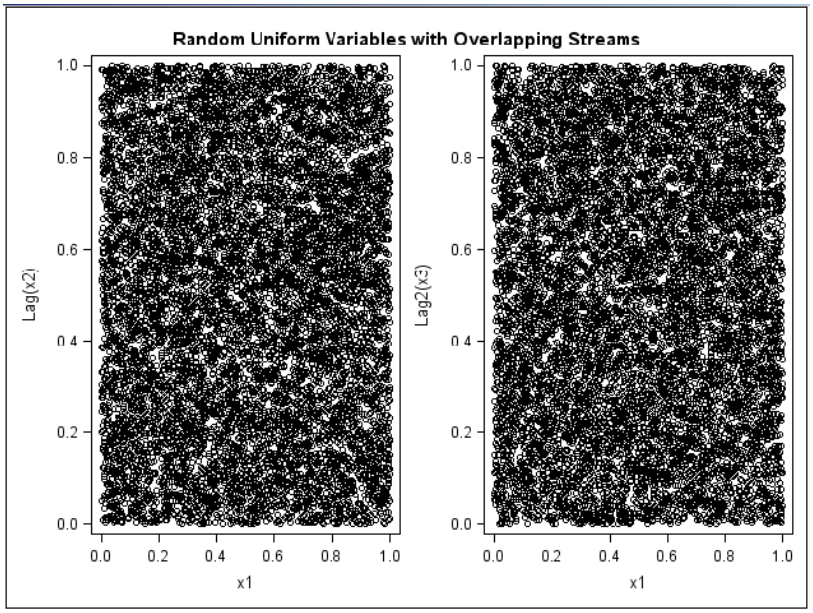
title 'Random Uniform Variables with Overlapping Streams';
proc sgscatter data = uniform;
   plot x1*x2 x1*x3 x3*x2 / markerattrs = (size = 1);
run;

proc sgscatter data = uniform;
   plot i2*x1 i3*x1 / markerattrs = (size = 1);
run;
```

Display 4.5 Random Uniform Variables with Overlapping Streams: Plot 1



Display 4.6 Random Uniform Variables with Overlapping Streams: Plot 2



In "Example: Generating Random Uniform Variables with Overlapping Streams" on page 324, it appears that the variables are independent. However, even this programming approach might not work well in general. The random-number functions and CALL routines have a period of only 2^{31} - 2 or less (approximately 2.1 billion). When this limit is reached, the stream repeats. Modern computers performing complicated simulations can easily exhaust the entire stream in minutes.

Using the RAND Function as an Alternative

A better approach to generating random uniform variables is to use the RAND function, where multiple streams are not permitted. The RAND function has a period of $2^{19937} - 1$. This limit will never be reached, at least with computers of the early 21st century. The number $2^{19937} - 1$ is approximately 10^{6000} (1 followed by 6000 zeros). In comparison, the largest value that can be represented in eight bytes on most computers that run SAS is approximately 10^{307}.

The RAND function, which is the latest random-number function that was designed, does not allow multiple streams. The RAND function uses a different algorithm from the random-number CALL routines, which allow you to create multiple streams with multiple seeds. Because the state of the RAND process cannot be captured by a single seed, you cannot stop and restart the generator from its stopping point. Therefore, the RAND function allows only a single stream of numbers, but it can be used to make multiple streams, just as the RANUNI function can.

Effectively Using the Random-Number CALL Routines

Starting, Stopping, and Restarting a Stream

A reasonable use of the random-number CALL routines is starting and stopping a single stream, provided the stream never exhausts the RANUNI stream. For example, you might want SAS to perform iterations, stop, evaluate the results, and then restart the stream at the point it stopped. The following example illustrates this principle.

Example: Starting, Stopping, and Restarting a Stream

This example generates a stream of five numbers, stops, restarts, generates five more numbers from the same stream, combines the results, and generates the full stream for comparison. In the first DATA step, the state of the random-number seed is stored in a macro variable seed for use as the starting seed in the next step. The separate streams in the example output match the full stream.

```
options pageno=1 nodate ls=80 ps=64;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call ranuni(seed, x);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
      call ranuni(seed, x);
      output;
   end;
run;
```

```
data all;
   set u1 u2;
   z = ranuni(104);
run;

proc print label;
   title 'Random Uniform Variables with Overlapping Streams';
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.5 Starting, Stopping, and Restarting a Stream

Random Uniform Variables with Overlapping Streams 1

Obs	Separate Streams	Single Stream
1	0.23611	0.23611
2	0.88923	0.88923
3	0.58173	0.58173
4	0.97746	0.97746
5	0.84667	0.84667
6	0.80484	0.80484
7	0.46983	0.46983
8	0.29594	0.29594
9	0.17858	0.17858
10	0.92292	0.92292

Comparison of Changing the Seed in a CALL Routine and in a Function

Example: Changing Seeds in a CALL Routine and in a Function

If you use a CALL routine to change the seed, the results are different from using a function to change the seed. The following example shows the difference.

```
data seeds;
   retain Seed1 Seed2 Seed3 104;
   do i = 1 to 10;
      call ranuni(Seed1,X1);
      call ranuni(Seed2,X2);
      X3 = ranuni(Seed3);
      if i = 5 then do;
         Seed2 = 17;
         Seed3 = 17;
      end;
      output;
   end;
run;

proc print data = seeds;
   title 'Random Uniform Variables with Overlapping Streams';
   id i;
run;
```

Output 4.6 Changing Seeds in a CALL Routine and in a Function

Random Uniform Variables with Overlapping Streams 3

i	Seed1	Seed2	Seed3	X1	X2	X3
1	507036483	507036483	104	0.23611	0.23611	0.23611
2	1909599212	1909599212	104	0.88923	0.88923	0.88923
3	1249251009	1249251009	104	0.58173	0.58173	0.58173
4	2099077474	2099077474	104	0.97746	0.97746	0.97746
5	1818205895	17	17	0.84667	0.84667	0.84667
6	1728390132	310018657	17	0.80484	0.14436	0.80484
7	1008960848	1055505749	17	0.46983	0.49151	0.46983
8	635524535	1711572821	17	0.29594	0.79701	0.29594
9	383494893	879989345	17	0.17858	0.40978	0.17858
10	1981958542	1432895200	17	0.92292	0.66724	0.92292

Changing Seed2 in the CALL RANUNI statement when i=5, forces the stream for X2 to deviate from the stream for X1. However, changing Seed3 in the RANUNI function has no effect. The X3 stream continues on as if nothing has changed, and the X1 and X3 streams are the same.

Date and Time Intervals

Definition of a Date and Time Interval

An interval is a unit of measurement that SAS counts within an elapsed period of time, such as days, months or hours. SAS determines date and time intervals based on fixed points on the calendar or clock. The starting point of an interval calculation defaults to the beginning of the period in which the beginning value falls, which might not be the actual beginning value that is specified. For example, if you are using the INTCK function to count the months between two dates, regardless of the actual day of the month that is specified by the date in the beginning value, SAS treats the beginning value as the first day of that month.

Interval Names and SAS Dates

Specific interval names are used with SAS date values, while other interval names are used with SAS time and datetime values. The interval names that are used with SAS date values are YEAR, SEMIYEAR, QTR, MONTH, SEMIMONTH, TENDAY, WEEK, WEEKDAY, and DAY. The interval names that are used with SAS time and datetime values are HOUR, MINUTE, and SECOND.

Interval names that are used with SAS date values can be prefixed with 'DT' to construct interval names for use with SAS datetime values. The interval names DTYEAR, DTSEMIYEAR, DTQTR, DTMONTH, DTSEMIMONTH, DTTENDAY, DTWEEK, DTWEEKDAY, and DTDAY are used with SAS time or datetime values.

Incrementing Dates and Times by Using Multipliers and by Shifting Intervals

SAS provides date, time, and datetime intervals for counting different periods of elapsed time. By using multipliers and shift indexes, you can create multiples of intervals and shift their starting point to construct more complex interval specifications.

The general form of an interval name is

name<*multiplier*><.*shift-index*>

Both the *multiplier* and the *shift–index* arguments are optional and default to 1. For example, YEAR, YEAR1, YEAR.1, and YEAR1.1 are all equivalent ways of specifying ordinary calendar years that begin in January. If you specify other values for *multiplier* and for *shift-index*, you can create multiple intervals that begin in different parts of the year. For example, the interval WEEK6.11 specifies six-week intervals starting on second Wednesdays.

For more information, see "Single-unit Intervals", "Multi-unit Intervals", and "Shifted Intervals" in *SAS Language Reference: Concepts*.

Commonly Used Time Intervals

Time intervals that do not nest within years or days are aligned relative to the SAS date or datetime value 0. SAS uses the arbitrary reference time of midnight on January 1, 1960, as the origin for non-shifted intervals. Shifted intervals are defined relative to January 1, 1960.

For example, MONTH13 defines the intervals January 1, 1960, February 1, 1961, March 1, 1962, and so on, and the intervals December 1, 1958, November 1, 1957, and so on, before the base date January 1, 1960.

As another example, the interval specification WEEK6.13 defines six-week periods starting on second Fridays. The convention of alignment relative to the period that contains January 1, 1960, determines where to start counting to determine which dates correspond to the second Fridays of six-week intervals.

The following table lists time intervals that are commonly used.

Table 4.3 Commonly Used Intervals with Optional Multiplier and Shift Indexes

Interval	Description
DAY3	Three-day intervals
WEEK	Weekly intervals starting on Sundays
WEEK.7	Weekly intervals starting on Saturdays
WEEK6.13	Six-week intervals starting on second Fridays
WEEK2	Biweekly intervals starting on first Sundays
WEEK1.1	Same as WEEK
WEEK.2	Weekly intervals starting on Mondays
WEEK6.3	Six-week intervals starting on first Tuesdays
WEEK6.11	Six-week intervals starting on second Wednesdays
WEEK4	Four-week intervals starting on first Sundays
WEEKDAY	Five-day work week with a Saturday-Sunday weekend

Interval	Description
WEEKDAY1W	Six-day week with Sunday as a weekend day
WEEKDAY35W	Five-day week with Tuesday and Thursday as weekend days (W indicates that day 3 and day 5 are weekend days)
WEEKDAY17W	Same as WEEKDAY
WEEKDAY67W	Five-day week with Friday and Saturday as weekend days
WEEKDAY3.2	Three-weekday intervals with Saturday and Sunday as weekend days (The intervals are aligned with respect to Jan. 1, 1960. For intervals that nest within a year, it is not necessary to go back to Jan. 1, 1960 to determine the alignment.)
TENDAY4.2	Four ten-day periods starting at the second TENDAY period
SEMIMONTH2.2	Intervals from the sixteenth of one month through the fifteenth of the next month
MONTH2.2	February–March, April–May, June–July, August–September, October–November, and December–January of the following year
MONTH2	January–February, March–April, May–June, July–August, September–October, November–December
QTR3.2	Nine-month intervals starting on February 1, 1960, November 1, 1960, August 1, 1961, May 1, 1962, and so on.
SEMIYEAR.3	Six-month intervals, March–August and September–February
YEAR.10	Fiscal years starting in October
YEAR2.7	Biennial intervals starting in July of even years
YEAR2.19	Biennial intervals starting in July of odd years
YEAR4.11	Four-year intervals starting in November of leap years (frequency of U.S. presidential elections)
YEAR4.35	Four-year intervals starting in November of even years between leap years (frequency of U.S. midterm elections)
DTMONTH13	Thirteen-month intervals starting at midnight of January 1, 1960, such as November 1, 1957, December 1, 1958, January 1, 1960, February 1, 1961, and March 1, 1962
HOUR8.7	Eight-hour intervals starting at 6 a.m., 2 p.m., and 10 p.m. (might be used for work shifts)

For a complete list of the valid values for *interval*, see the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Retail Calendar Intervals: ISO 8601 Compliant

The retail industry often accounts for its data by dividing the yearly calendar into four 13-week periods, based on one of the following formats: 4-4-5, 4-5-4, or 5-4-4. The first, second, and third numbers specify the number of weeks in the first, second, and third months of each period, respectively.

The intervals that are created from the formats can be used in any of the following functions: INTCINDEX, INTCK, INTCYCLE, INTFIT, INTFMT, INTGET, INTINDEX, INTNX, INTSEAS, INTSHIFT, and INTTEST.

For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Best Practices for Custom Interval Names

The following items list best practices to use when you are creating custom interval names:

- Custom interval names should not conflict with existing SAS interval names. For example, if BASE is a SAS interval name, do not use the following formats for the name of a custom interval:

  BASE

  BASE*m*

  BASE*m*.*n*

  DTBASE

  DTBASE*m*

  DTBASE*m*.*n*

  where

  m
  specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

  n
  specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods that are shifted to start on the first day of March of each calendar year and end in February of the following year.

  If you define a custom interval such as CUSTBASE, then you can use CUSTBASE*m*.*n*.

  Because of these rules, do not begin a custom interval name with DT, and do not end the custom interval name with a number.

- To ensure that custom intervals work reliably, always include one of the following formats:

  date-format with beginning and ending values
  specifies intervals that are used with SAS date values.

  datetime-format with beginning and ending values
  specifies intervals that are used with SAS datetime values.

  number-format with beginning and ending values
  specifies intervals that are used with SAS observation numbers.

- Beginning and ending values should be of the same type. Both values should be date values, datetime values, or observation numbers.
- Calculations for custom intervals cannot be performed before the first begin value or after the last end value. If you use the begin variable only, then the last end value you can calculate is the last begin value –1. If you forecast or backcast the time series, be sure to include time definitions for the forecast and backcast values.

□

CUSTBASE*m*.2 is never able to calculate a beginning period for the first date value in a data set because, by definition, the beginning of the first interval starts before the data set begins (at the – (*m*– 2) th observation). For example, you might have an interval called CUSTBASE4.2 with the first interval beginning before the first observation:

```
OBS
-2    Start of partial CUSTBASE4.2 interval observation:  -(4-2) = -2.
-1
0
1     End of partial CUSTBASE4.2 interval observation: This is the first
       observation in the data set.
2     Start of first complete CUSTBASE4.2 interval.
3
4
5     End of first complete CUSTBASE4.2 interval.
6     Start of 2nd CUSTBASE4.2 interval.
```

If you execute the INTNX function, the result must return the date that is associated with OBS –2, which does not exist:

```
INTNX('CUSTBASE4.2', date-at-obs1, 0, 'B');
```

□ Include a variable named *season* in the custom interval data set to define the seasonal index. This result is similar to the result of **INTINDEX ('*interval*', *date*);**

In the following example, the data set is associated with the custom interval CUSTWEEK:

Obs	begin	season
1	27DEC59	52
2	03JAN60	1
3	10JAN60	2
4	17JAN60	3
5	24JAN60	4
6	31JAN60	5

The following examples show the results of using custom interval functions:

INTINDEX ('CUSTWEEK', '*03JAN60'D*);
returns a value of 1.

INTSEAS ('CUSTWEEK');
returns a value of 52, which is the largest value of the season.

INTCYCLE ('CUSTWEEK');
returns CUSTWEEK52, which is CUSTBASE*max*(*season*).

INTCINDEX ('CUSTWEEK', '27DEC59'D);
returns a value of 1.

INTCINDEX('CUSTWEEK', '03JAN60'D)
returns a value of 2.

A new cycle begins when the season is less than the previous value of *season*.

□ Seasonality occurs when seasons are identified, such as season1, season2, season3, and so forth. If all seasons are identified as season1, then there is no seasonality. No seasonality is also called trivial seasonality.

Only trivial seasonality is available for intervals of the form CUSTBASE*m*. If *season* is not included in the data set, then trivial seasonality is valid.

- If a format for the begin variable is included in a data set, then a message generated by INTFMT ('CUSTBASE', 'l') or INTFMT ('CUSTBASE', 's') appears. The message recommends a format based on the format that is specified in the data set.
- Executing **`INTSHIFT ('CUSTBASE');`** or **`INTSHIFT ('CUSTBASE`*`m.s`*`');`** returns the value of CUSTBASE.
- With INTNX, INTCK, and INTTEST, the intervals CUSTBASE, CUSTBASE*m*, and CUSTBASE*m.s* work as expected.

Pattern Matching Using Perl Regular Expressions (PRX)

Definition of Pattern Matching

Pattern matching enables you to search for and extract multiple matching patterns from a character string in one step. Pattern matching also enables you to make several substitutions in a string in one step. You do this by using the PRX functions and CALL routines in the DATA step.

For example, you can search for multiple occurrences of a string and replace those strings with another string. You can search for a string in your source file and return the position of the match. You can find words in your file that are doubled.

Definition of Perl Regular Expression (PRX) Functions and CALL Routines

Perl regular expression (PRX) functions and CALL routines refers to a group of functions and CALL routines that use a modified version of Perl as a pattern-matching language to parse character strings. You can do the following:

- search for a pattern of characters within a string
- extract a substring from a string
- search and replace text with other text
- parse large amounts of text, such as Web logs or other text data

Perl regular expressions comprise the character string matching category for functions and CALL routines. For a short description of these functions and CALL routines, see the "Functions and CALL Routines by Category" on page 345.

Benefits of Using Perl Regular Expressions in the DATA Step

Using Perl regular expressions in the DATA step enhances search-and-replace options in text. You can use Perl regular expressions to perform the following tasks:

- validate data
- replace text
- extract a substring from a string

You can write SAS programs that do not use regular expressions to produce the same results as you do when you use Perl regular expressions. However, the code without the

regular expressions requires more function calls to handle character positions in a string and to manipulate parts of the string.

Perl regular expressions combine most, if not all, of these steps into one expression. The resulting code is less prone to error, easier to maintain, and clearer to read.

Using Perl Regular Expressions in the DATA Step

Syntax of Perl Regular Expressions

The Components of a Perl Regular Expression

Perl regular expressions consist of characters and special characters that are called metacharacters. When performing a match, SAS searches a source string for a substring that matches the Perl regular expression that you specify. Using metacharacters enables SAS to perform special actions. These actions include forcing the match to begin in a particular location, and matching a particular set of characters. Paired forward slashes are the default delimiters. The following two examples show metacharacters and the values they match:

- If you use the metacharacter \d, SAS matches a digit between 0–9.
- If you use /\dt/, SAS finds the digits in the string "Raleigh, NC 27506".

You can see lists of PRX metacharacters in "Tables of Perl Regular Expression (PRX) Metacharacters" on page 2205.

Basic Syntax for Finding a Match in a String

You use the PRXMATCH function to find the position of a matched value in a source string. PRXMATCH has the following general form:

```
/search-string/source-string/
```

The following example uses the PRXMATCH function to find the position of *search-string* in *source-string*:

```
prxmatch('world', 'Hello world!');
```

The result of PRXMATCH is the value 7, because *world* occurs in the seventh position of the string *Hello world!*.

Basic Syntax for Searching and Replacing Text: Example 1

The basic syntax for searching and replacing text has the following form:

```
s/regular-expression/replacement-string/
```

The following example uses the PRXCHANGE function to show how substitution is performed:

```
prxchange('s/world/planet/', 1, 'Hello world!');
```

where

s specifies the metacharacter for substitution.

world	specifies the regular expression.
planet	specifies the replacement value for *world*.
1	specifies that the search ends when one match is found.
Hello world!	specifies the source string to be searched.

The result of the substitution is **`Hello planet`**.

Basic Syntax for Searching and Replacing Text: Example 2

Another example of using the PRXCHANGE function changes the value *Jones, Fred* to *Fred Jones*:

```
prxchange('s/(\w+), (\w+)/$2 $1',-1, 'Jones, Fred');
```

In this example, the Perl regular expression is **`s/(\w+), (\w+)/$2 $1`**. The number of times to search for a match is –1. The source string is 'Jones, Fred'. The value –1 specifies that matching patterns continue to be replaced until the end of the source is reached.

The Perl regular expression can be divided into its elements:

s	specifies a substitution regular expression.
(\w+)	matches one or more word characters (alphanumeric and underscore). The parentheses indicate that the value is stored in capture buffer 1.
,<space>	matches a comma and a space.
(\w+)	matches one or more word characters (alphanumeric and underscore). The parentheses indicate that the value is stored in capture buffer 2.
/	separator between the regular expression and the replacement string.
$2	part of the replacement string that substitutes the value in capture buffer 2, which in this case is the word after the comma, puts the substitution in the results.
<space>	puts a space in the result.
$1	puts capture buffer 1 into the result. In this case, it is the word before the comma.

Replacing Text: Example 3

The following example uses the \u and \L metacharacters to replace the second character in MCLAUREN with a lower case letter:

```
data _null_;
   x = 'MCLAUREN';
   x = prxchange("s/ (MC) /\u\L$1/i", -1, x);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=McLAUREN
```

Example 1: Validating Data

You can test for a pattern of characters within a string. For example, you can examine a string to determine whether it contains a correctly formatted telephone number. This type of test is called data validation.

The following example validates a list of phone numbers. To be valid, a phone number must have one of the following forms: **(XXX) XXX-XXXX** or **XXX-XXX-XXXX**.

```
data _null_;  ❶
   if _N_ = 1 then
      do;
         paren = "\([2-9]\d\d\) ?[2-9]\d\d-\d\d\d\d";  ❷
         dash = "[2-9]\d\d-[2-9]\d\d-\d\d\d\d";  ❸
         expression = "/(" || paren || ")|(" || dash || ")/";  ❹
         retain re;
         re = prxparse(expression);  ❺
         if missing(re) then  ❻
            do;
               putlog "ERROR: Invalid expression " expression;  ❼
               stop;
            end;
      end;

   length first last home business $ 16;
   input first last home business;

   if ^prxmatch(re, home) then  ❽
      putlog "NOTE: Invalid home phone number for " first last home;

   if ^prxmatch(re, business) then  ❾
      putlog "NOTE: Invalid business phone number for " first last business;

   datalines;
Jerome Johnson (919)319-1677 (919)846-2198
Romeo Montague 800-899-2164 360-973-6201
Imani Rashid (508)852-2146 (508)366-9821
Palinor Kent . 919-782-3199
Ruby Archuleta . .
Takei Ito 7042982145 .
Tom Joad 209/963/2764 2099-66-8474
;
run;
```

The following items correspond to the lines that are numbered in the DATA step that is shown above.

❶ Create a DATA step.

❷ Build a Perl regular expression to identify a phone number that matches (XXX)XXX-XXXX, and assign the variable PAREN to hold the result. Use the following syntax elements to build the Perl regular expression:

\(	matches the open parenthesis in the area code.
[2–9]	matches the digits 2–9, which is the first number in the area code.

\d	matches a digit, which is the second number in the area code.
\d	matches a digit, which is the third number in the area code.
\)	matches the closed parenthesis in the area code.
<space>?	matches the space (which is the preceding subexpression) zero or one time. Spaces are significant in Perl regular expressions. They match a space in the text that you are searching. If a space precedes the question mark metacharacter (as it does in this case), the pattern matches either zero spaces or one space in this position in the phone number.

❸ Build a Perl regular expression to identify a phone number that matches XXX-XXX-XXXX, and assign the variable DASH to hold the result.

❹ Build a Perl regular expression that concatenates the regular expressions for (XXX)XXX-XXXX and XXX—XXX—XXXX. The concatenation enables you to search for both phone number formats from one regular expression.

The PAREN and DASH regular expressions are placed within parentheses. The bar metacharacter (|) that is located between PAREN and DASH instructs the compiler to match either pattern. The slashes around the entire pattern tell the compiler where the start and end of the regular expression is located.

❺ Pass the Perl regular expression to PRXPARSE and compile the expression. PRXPARSE returns a value to the compiled pattern. Using the value with other Perl regular expression functions and CALL routines enables SAS to perform operations with the compiled Perl regular expression.

❻ Use the MISSING function to check whether the regular expression was successfully compiled.

❼ Use the PUTLOG statement to write an error message to the SAS log if the regular expression did not compile.

❽ Search for a valid home phone number. PRXMATCH uses the value from PRXPARSE along with the search text and returns the position where the regular expression was found in the search text. If there is no match for the home phone number, the PUTLOG statement writes a note to the SAS log.

❾ Search for a valid business phone number. PRXMATCH uses the value from PRXPARSE along with the search text and returns the position where the regular expression was found in the search text. If there is no match for the business phone number, the PUTLOG statement writes a note to the SAS log.

Output 4.7 Output from Validating Data

```
NOTE: Invalid home phone number for Palinor Kent
NOTE: Invalid home phone number for Ruby Archuleta
NOTE: Invalid business phone number for Ruby Archuleta
NOTE: Invalid home phone number for Takei Ito 7042982145
NOTE: Invalid business phone number for Takei Ito
NOTE: Invalid home phone number for Tom Joad 209/963/2764
NOTE: Invalid business phone number for Tom Joad 2099-66-8474
```

Example 2: Replacing Text

This example uses a Perl regular expression to find a match and replace the matching characters with other characters. PRXPARSE compiles the regular expression and uses PRXCHANGE to find the match and perform the replacement. The example replaces all occurrences of a less than sign with **<**, a common substitution when converting text to HTML.

```
data _null_;  ❶
   input;  ❷
   _infile_ = prxchange('s/</&lt;/', -1, _infile_); ❸
   put _infile_; ❹
   datalines;  ❺
x + y < 15
x < 10 < y
y < 11
;
run;
```

The following items correspond to the numbered lines in the DATA step that is shown above.

❶ Create a DATA step.

❷ Bring an input data record into the input buffer without creating any SAS variables.

❸ Call the PRXCHANGE routine to perform the pattern exchange. The format for the regular expression is **s/regular-expression/replacement-text/**. The **s** before the regular expression signifies that this is a substitution regular expression. The –1 is a special value that is passed to PRXCHANGE and indicates that all possible replacements should be made.

❹ Write the current output line to the log by using the _INFILE_ option with the PUT statement.

❺ Identify the input file.

Output 4.8 Output from Replacing Text

```
x + y &lt; 15
x &lt; 10 &lt; y
y &lt; 11
```

The ability to pass a regular expression to PRXCHANGE and return a result enables calling PRXCHANGE from a PROC SQL query. The following query produces a column with the same character substitution as in the preceding example. From the input table the query reads **text_lines**, changes the text for the column **line**, and places the results in a column named **html_line**:

```
proc sql;
   select prxchange('s/</&lt;/', -1, line)
   as html_line
   from text_lines;
quit;
```

Example 3: Extracting a Substring from a String

You can use Perl regular expressions to find and easily extract text from a string. In this example, the DATA step creates a subset of North Carolina business phone numbers. The program extracts the area code and checks it against a list of area codes for North Carolina.

```
data _null_;  ❶
   if _N_ = 1 then
      do;
         paren = "\(([2-9]\d\d)\) ?[2-9]\d\d-\d\d\d\d";  ❷
         dash = "([2-9]\d\d)-[2-9]\d\d-\d\d\d\d";  ❸
         regexp = "/(" || paren || ")|(" || dash || ")/";  ❹
         retain re;
         re = prxparse(regexp);  ❺
         if missing(re) then  ❻
            do;
               putlog "ERROR: Invalid regexp " regexp;  ❼
               stop;
            end;

         retain areacode_re;
         areacode_re = prxparse("/828|336|704|910|919|252/");  ❽
         if missing(areacode_re) then
            do;
               putlog "ERROR: Invalid area code regexp";
               stop;
            end;
      end;

   length first last home business $ 16;
   length areacode $ 3;
   input first last home business;

   if ^prxmatch(re, home) then
      putlog "NOTE: Invalid home phone number for " first last home;

   if prxmatch(re, business) then  ❾
      do;
         which_format = prxparen(re);  ❿
         call prxposn(re, which_format, pos, len);  ⓫
         areacode = substr(business, pos, len);
         if prxmatch(areacode_re, areacode) then  ⓬
            put "In North Carolina: " first last business;
      end;
      else
         putlog "NOTE: Invalid business phone number for " first last business;
   datalines;
Jerome Johnson (919)319-1677 (919)846-2198
Romeo Montague 800-899-2164 360-973-6201
Imani Rashid (508)852-2146 (508)366-9821
Palinor Kent 704-782-4673 704-782-3199
Ruby Archuleta 905-384-2839 905-328-3892
```

```
Takei Ito 704-298-2145 704-298-4738
Tom Joad 515-372-4829 515-389-2838
;
```

The following items correspond to the numbered lines in the DATA step that is shown above.

❶ Create a DATA step.

❷ Build a Perl regular expression to identify a phone number that matches (XXX)XXX-XXXX, and assign the variable PAREN to hold the result. Use the following syntax elements to build the Perl regular expression:

\(	matches the open parenthesis in the area code. The open parenthesis marks the start of the submatch.
[2–9]	matches the digits 2–9.
\d	matches a digit, which is the second number in the area code.
\d	matches a digit, which is the third number in the area code.
\)	matches the closed parenthesis in the area code. The closed parenthesis marks the end of the submatch.
?	matches the space (which is the preceding subexpression) zero or one time. Spaces are significant in Perl regular expressions. They match a space in the text that you are searching. If a space precedes the question mark metacharacter (as it does in this case), the pattern matches either zero spaces or one space in this position in the phone number.

❸ Build a Perl regular expression to identify a phone number that matches XXX-XXX-XXXX, and assign the variable DASH to hold the result.

❹ Build a Perl regular expression that concatenates the regular expressions for (XXX)XXX-XXXX and XXX—XXX—XXXX. The concatenation enables you to search for both phone number formats from one regular expression.

The PAREN and DASH regular expressions are placed within parentheses. The bar metacharacter (|) that is located between PAREN and DASH instructs the compiler to match either pattern. The slashes around the entire pattern tell the compiler where the start and end of the regular expression is located.

❺ Pass the Perl regular expression to PRXPARSE and compile the expression. PRXPARSE returns a value to the compiled pattern. Using the value with other Perl regular expression functions and CALL routines enables SAS to perform operations with the compiled Perl regular expression.

❻ Use the MISSING function to check whether the Perl regular expression compiled without error.

❼ Use the PUTLOG statement to write an error message to the SAS log if the regular expression did not compile.

❽ Compile a Perl regular expression that searches a string for a valid North Carolina area code.

❾ Search for a valid business phone number.

❿ Use the PRXPAREN function to determine which submatch to use. PRXPAREN returns the last submatch that was matched. If an area code matches the form (XXX), PRXPAREN returns the value 2. If an area code matches the form XXX, PRXPAREN returns the value 4.

⓫ Call the PRXPOSN routine to retrieve the position and length of the submatch.

⓬ Use the PRXMATCH function to determine whether the area code is a valid North Carolina area code, and write the observation to the log.

Output 4.9 Output from Extracting a Substring from a String

```
In North Carolina: Jerome Johnson (919)846-2198
In North Carolina: Palinor Kent 704-782-3199
In North Carolina: Takei Ito 704-298-4738
```

Example 4: Another Example of Extracting a Substring from a String

In this example, the PRXPOSN function is passed to the original search text instead of to the position and length variables. PRXPOSN returns the text that is matched.

```
data _null_;  ❶
   length first last phone $ 16;
   retain re;
   if _N_ = 1 then do;  ❷
      re=prxparse("/\(([2-9]\d\d)\) ?[2-9]\d\d-\d\d\d\d/");  ❸
   end;

   input first last phone & 16.;
   if prxmatch(re, phone) then do;  ❹
      area_code = prxposn(re, 1, phone);  ❺
      if area_code ^in ("828"
                                  "336"
                                  "704"
                                  "910"
                                  "919"
                                  "252") then
         putlog "NOTE: Not in North Carolina: "
                     first last phone;  ❻
    end;
   datalines;  ❼
Thomas Archer    (919)319-1677
Lucy Mallory       (800)899-2164
Tom Joad              (508)852-2146
Laurie Jorgensen  (252)352-7583
;
run;
```

The following items correspond to the numbered lines in the DATA step that is shown above.

❶ Create a DATA step.

❷ If this is the first record, find the value of *re*.

❸ Build a Perl regular expression for pattern matching. Use the following syntax elements to build the Perl regular expression:

/	is the beginning delimiter for a regular expression.
\(	marks the next character entry as a character or a literal.
(	marks the start of the submatch.

[2–9]	matches the digits 2–9 and identifies the first number in the area code.
\d	matches a digit, which is the second number in the area code.
\d	matches a digit, which is the third number in the area code.
\)	matches the close parenthesis in the area code. The close parenthesis marks the end of the submatch.
?	matches the space (which is the preceding subexpression) zero or one time. Spaces are significant in Perl regular expressions. The spaces match a space in the text that you are searching. If a space precedes the question mark metacharacter (as it does in this case), the pattern matches either zero spaces or one space in this position in the phone number.
||	is the concatenation operator.
[2–9]	matches the digits 2–9 and identifies the first number in the seven-digit phone number.
\d	matches a digit, which is the second number in the seven-digit phone number.
\d	matches a digit, which is the third number in the seven-digit phone number.
–	is the hyphen between the first three and last four digits of the phone number after the area code.
\d	matches a digit, which is the fourth number in the seven-digit phone number.
\d	matches a digit, which is the fifth number in the seven-digit phone number.
\d	matches a digit, which is the sixth number in the seven-digit phone number.
\d	matches a digit, which is the seventh number in the seven-digit phone number.
/	is the ending delimiter for a regular expression.

❹ Return the position at which the string begins.

❺ Identify the position at which the area code begins.

❻ Search for an area code from the list. If the area code is not valid for North Carolina, use the PUTLOG statement to write a note to the SAS log.

❼ Identify the input file.

Output 4.10 Output from Extracting a Substring from a String

```
NOTE: Not in North Carolina: Lucy Mallory (800)899-2164
NOTE: Not in North Carolina: Tom Joad (508)852-2146
```

Writing Perl Debug Output to the SAS Log

The DATA step provides debugging support with the CALL PRXDEBUG routine. CALL PRXDEBUG enables you to turn on and off Perl debug output messages that are sent to the SAS log.

The following example writes Perl debug output to the SAS log.

```
data _null_;

      /* CALL PRXDEBUG(1) turns on Perl debug output. */
   call prxdebug(1);
   putlog 'PRXPARSE: ';
   re = prxparse('/[bc]d(ef*g)+h[ij]k$/');
   putlog 'PRXMATCH: ';
   pos = prxmatch(re, 'abcdefg_gh_');

      /* CALL PRXDEBUG(0) turns off Perl debug output. */
   call prxdebug(0);
run;
```

SAS writes the following output to the log.

Output 4.11 SAS Debugging Output

```
PRXPARSE:
Compiling REx '[bc]d(ef*g)+h[ij]k$'
size 41 first at 1
rarest char g at 0
rarest char d at 0
   1: ANYOF[bc](10)
  10: EXACT <d>(12)
  12: CURLYX[0] {1,32767}(26)
  14:   OPEN1(16)
  16:     EXACT <e>(18)
  18:     STAR(21)
  19:       EXACT <f>(0)
  21:     EXACT <g>(23)
  23:   CLOSE1(25)
  25:   WHILEM[1/1](0)
  26: NOTHING(27)
  27: EXACT <h>(29)
  29: ANYOF[ij](38)
  38: EXACT <k>(40)
  40: EOL(41)
  41: END(0)
anchored 'de' at 1 floating 'gh' at 3..2147483647 (checking floating) stclass
'ANYOF[bc]' minlen 7

PRXMATCH:
Guessing start of match, REx '[bc]d(ef*g)+h[ij]k$' against 'abcdefg_gh_'...
Did not find floating substr 'gh'...
Match rejected by optimizer
```

For a detailed explanation of Perl debug output, see "CALL PRXDEBUG Routine" on page 482.

Perl Artistic License Compliance

Perl regular expressions are supported beginning with SAS®9.

The PRX functions use a modified version of Perl 5.6.1 to perform regular expression compilation and matching. Perl is compiled into a library for use with SAS. This library is shipped with SAS®9. The modified and original Perl 5.6.1 files are freely available in a ZIP file from **http://support.sas.com/rnd/base**. The ZIP file is provided to comply with the Perl Artistic License and is not required in order to use the PRX functions. Each of the modified files has a comment block at the top of the file describing how and when the file was changed. The executables were given nonstandard Perl names. The standard version of Perl can be obtained from **http://www.perl.com**.

Only Perl regular expressions are accessible from the PRX functions. Other parts of the Perl language are not accessible. The modified version of Perl regular expressions does not support the following items:

- Perl variables (except the capture buffer variables $1 - $n, which are supported).
- The regular expression options /c and /g, and the /e option with substitutions.
- The regular expression option /o in SAS 9.0. (It is supported in SAS 9.1 and later.)
- Named characters, which use the \N{name} syntax.
- The metacharacters \pP, \PP, and \X.
- Executing Perl code within a regular expression, which includes the syntax (?{code}), (??{code}), and (?p{code}).
- Unicode pattern matching.
- Using ?PATTERN?. ? is treated like an ordinary regular expression start and end delimiter.
- The metacharacter \G.
- Perl comments between a pattern and replacement text. For example: s{regexp} # perl comment {replacement} is not supported.
- Matching backslashes with m/\\\\/. Instead use m/\\/ to match a backslash.

Base SAS Functions for Web Applications

Four functions that manipulate Web-related content are available in Base SAS software. HTMLENCODE and URLENCODE return encoded strings. HTMLDECODE and URLDECODE return decoded strings. For information about Web-based SAS tools, follow the Communities link on the SAS support page, at **support.sas.com**.

Functions and CALL Routines by Category

Table 4.4 Categories and Descriptions of Functions and CALL Routines

Category	Functions and CALL Routines	Description
	"ANYXDIGIT Function" on page 401	Searches a character string for a hexadecimal character that represents a digit, and returns the first position at which that character is found.
Arithmetic	"DIVIDE Function" on page 658	Returns the result of a division that handles special missing values for ODS output.
Array	"DIM Function" on page 655	Returns the number of elements in an array.
	"HBOUND Function" on page 802	Returns the upper bound of an array.
	"LBOUND Function" on page 878	Returns the lower bound of an array.
Bitwise Logical Operations	"BAND Function" on page 416	Returns the bitwise logical AND of two arguments.
	"BLSHIFT Function" on page 426	Returns the bitwise logical left shift of two arguments.
	"BNOT Function" on page 427	Returns the bitwise logical NOT of an argument.
	"BOR Function" on page 428	Returns the bitwise logical OR of two arguments.
	"BRSHIFT Function" on page 429	Returns the bitwise logical right shift of two arguments.
	"BXOR Function" on page 430	Returns the bitwise logical EXCLUSIVE OR of two arguments.
Character String Matching	"CALL PRXCHANGE Routine" on page 479	Performs a pattern-matching replacement.
	"CALL PRXDEBUG Routine" on page 482	Enables Perl regular expressions in a DATA step to send debugging output to the SAS log.
	"CALL PRXFREE Routine" on page 484	Frees memory that was allocated for a Perl regular expression.
	"CALL PRXNEXT Routine" on page 485	Returns the position and length of a substring that matches a pattern, and iterates over multiple matches within one string.
	"CALL PRXPOSN Routine" on page 487	Returns the start position and length for a capture buffer.
	"CALL PRXSUBSTR Routine" on page 490	Returns the position and length of a substring that matches a pattern.

Category	Functions and CALL Routines	Description
	"PRXCHANGE Function" on page 1039	Performs a pattern-matching replacement.
	"PRXMATCH Function" on page 1045	Searches for a pattern match and returns the position at which the pattern is found.
	"PRXPAREN Function" on page 1049	Returns the last bracket match for which there is a match in a pattern.
	"PRXPARSE Function" on page 1051	Compiles a Perl regular expression (PRX) that can be used for pattern matching of a character value.
	"PRXPOSN Function" on page 1053	Returns a character string that contains the value for a capture buffer.
Character	"ANYALNUM Function" on page 379	Searches a character string for an alphanumeric character, and returns the first position at which the character is found.
	"ANYALPHA Function" on page 381	Searches a character string for an alphabetic character, and returns the first position at which the character is found.
	"ANYCNTRL Function" on page 383	Searches a character string for a control character, and returns the first position at which that character is found.
	"ANYDIGIT Function" on page 384	Searches a character string for a digit, and returns the first position at which the digit is found.
	"ANYFIRST Function" on page 386	Searches a character string for a character that is valid as the first character in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.
	"ANYGRAPH Function" on page 388	Searches a character string for a graphical character, and returns the first position at which that character is found.
	"ANYLOWER Function" on page 390	Searches a character string for a lowercase letter, and returns the first position at which the letter is found.
	"ANYNAME Function" on page 392	Searches a character string for a character that is valid in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.
	"ANYPRINT Function" on page 394	Searches a character string for a printable character, and returns the first position at which that character is found.
	"ANYPUNCT Function" on page 396	Searches a character string for a punctuation character, and returns the first position at which that character is found.
	"ANYSPACE Function" on page 397	Searches a character string for a white-space character (blank, horizontal and vertical tab, carriage return, line feed, and form feed), and returns the first position at which that character is found.
	"ANYUPPER Function" on page 399	Searches a character string for an uppercase letter, and returns the first position at which the letter is found.

Category	Functions and CALL Routines	Description
	"ANYXDIGIT Function" on page 401	Searches a character string for a hexadecimal character that represents a digit, and returns the first position at which that character is found.
	"BYTE Function" on page 431	Returns one character in the ASCII or the EBCDIC collating sequence.
	"CALL CATS Routine" on page 441	Removes leading and trailing blanks, and returns a concatenated character string.
	"CALL CATT Routine" on page 443	Removes trailing blanks, and returns a concatenated character string.
	"CALL CATX Routine" on page 445	Removes leading and trailing blanks, inserts delimiters, and returns a concatenated character string.
	"CALL COMPCOST Routine" on page 447	Sets the costs of operations for later use by the COMPGED function
	"CALL MISSING Routine" on page 473	Assigns missing values to the specified character or numeric variables.
	"CALL SCAN Routine" on page 516	Returns the position and length of the *n*th word from a character string.
	"CAT Function" on page 543	Does not remove leading or trailing blanks, and returns a concatenated character string.
	"CATQ Function" on page 546	Concatenates character or numeric values by using a delimiter to separate items and by adding quotation marks to strings that contain the delimiter.
	"CATS Function" on page 550	Removes leading and trailing blanks, and returns a concatenated character string.
	"CATT Function" on page 552	Removes trailing blanks, and returns a concatenated character string.
	"CATX Function" on page 554	Removes leading and trailing blanks, inserts delimiters, and returns a concatenated character string.
	"CHAR Function" on page 578	Returns a single character from a specified position in a character string.
	"CHOOSEC Function" on page 579	Returns a character value that represents the results of choosing from a list of arguments.
	"CHOOSEN Function" on page 581	Returns a numeric value that represents the results of choosing from a list of arguments.
	"COALESCEC Function" on page 588	Returns the first non-missing value from a list of character arguments.
	"COLLATE Function" on page 589	Returns a character string in ASCII or EBCDIC collating sequence.
	"COMPARE Function" on page 591	Returns the position of the leftmost character by which two strings differ, or returns 0 if there is no difference.
	"COMPBL Function" on page 594	Removes multiple blanks from a character string.

Category	Functions and CALL Routines	Description
	"COMPGED Function" on page 596	Returns the generalized edit distance between two strings.
	"COMPLEV Function" on page 601	Returns the Levenshtein edit distance between two strings.
	"COMPRESS Function" on page 604	Returns a character string with specified characters removed from the original string.
	"COUNT Function" on page 616	Counts the number of times that a specified substring appears within a character string.
	"COUNTC Function" on page 618	Counts the number of characters in a string that appear or do not appear in a list of characters.
	"COUNTW Function" on page 621	Counts the number of words in a character string.
	"DEQUOTE Function" on page 646	Removes matching quotation marks from a character string that begins with a quotation mark, and deletes all characters to the right of the closing quotation mark.
	"FIND Function" on page 735	Searches for a specific substring of characters within a character string.
	"FINDC Function" on page 737	Searches a string for any character in a list of characters.
	"FINDW Function" on page 743	Returns the character position of a word in a string, or returns the number of the word in a string.
	"FIRST Function" on page 755	Returns the first character in a character string.
	"IFC Function" on page 812	Returns a character value based on whether an expression is true, false, or missing.
	"INDEX Function" on page 817	Searches a character expression for a string of characters, and returns the position of the string's first character for the first occurrence of the string.
	"INDEXC Function" on page 819	Searches a character expression for any of the specified characters, and returns the position of that character.
	"INDEXW Function" on page 820	Searches a character expression for a string that is specified as a word, and returns the position of the first character in the word.
	"LEFT Function" on page 881	Left-aligns a character string.
	"LENGTH Function" on page 883	Returns the length of a non-blank character string, excluding trailing blanks, and returns 1 for a blank character string.
	"LENGTHC Function" on page 884	Returns the length of a character string, including trailing blanks.
	"LENGTHM Function" on page 885	Returns the amount of memory (in bytes) that is allocated for a character string.

Category	Functions and CALL Routines	Description
	"LENGTHN Function" on page 887	Returns the length of a character string, excluding trailing blanks.
	"LOWCASE Function" on page 912	Converts all letters in an argument to lowercase.
	"MD5 Function" on page 922	Returns the result of the message digest of a specified string.
	"MISSING Function" on page 929	Returns a numeric result that indicates whether the argument contains a missing value.
	"NLITERAL Function" on page 945	Converts a character string that you specify to a SAS name literal.
	"NOTALNUM Function" on page 948	Searches a character string for a non-alphanumeric character, and returns the first position at which the character is found.
	"NOTALPHA Function" on page 950	Searches a character string for a nonalphabetic character, and returns the first position at which the character is found.
	"NOTCNTRL Function" on page 952	Searches a character string for a character that is not a control character, and returns the first position at which that character is found.
	"NOTDIGIT Function" on page 954	Searches a character string for any character that is not a digit, and returns the first position at which that character is found.
	"NOTFIRST Function" on page 957	Searches a character string for an invalid first character in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.
	"NOTGRAPH Function" on page 959	Searches a character string for a non-graphical character, and returns the first position at which that character is found.
	"NOTLOWER Function" on page 961	Searches a character string for a character that is not a lowercase letter, and returns the first position at which that character is found.
	"NOTNAME Function" on page 963	Searches a character string for an invalid character in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.
	"NOTPRINT Function" on page 965	Searches a character string for a nonprintable character, and returns the first position at which that character is found.
	"NOTPUNCT Function" on page 967	Searches a character string for a character that is not a punctuation character, and returns the first position at which that character is found.
	"NOTSPACE Function" on page 969	Searches a character string for a character that is not a white-space character (blank, horizontal and vertical tab, carriage return, line feed, and form feed), and returns the first position at which that character is found.

Category	Functions and CALL Routines	Description
	"NOTUPPER Function" on page 971	Searches a character string for a character that is not an uppercase letter, and returns the first position at which that character is found.
	"NOTXDIGIT Function" on page 973	Searches a character string for a character that is not a hexadecimal character, and returns the first position at which that character is found.
	"NVALID Function" on page 975	Checks the validity of a character string for use as a SAS variable name.
	"PROPCASE Function" on page 1037	Converts all words in an argument to proper case.
	"QUOTE Function" on page 1066	Adds double quotation marks to a character value.
	"RANK Function" on page 1085	Returns the position of a character in the ASCII or EBCDIC collating sequence.
	"REPEAT Function" on page 1093	Returns a character value that consists of the first argument repeated n+1 times.
	"REVERSE Function" on page 1095	Reverses a character string.
	"RIGHT Function" on page 1097	Right aligns a character expression.
	"SCAN Function" on page 1111	Returns the *n*th word from a character string.
	"SOUNDEX Function" on page 1129	Encodes a string to facilitate searching.
	"SPEDIS Function" on page 1130	Determines the likelihood of two words matching, expressed as the asymmetric spelling distance between the two words.
	"STRIP Function" on page 1138	Returns a character string with all leading and trailing blanks removed.
	"SUBPAD Function" on page 1140	Returns a substring that has a length you specify, using blank padding if necessary.
	"SUBSTR (left of =) Function" on page 1141	Replaces character value contents.
	"SUBSTR (right of =) Function" on page 1143	Extracts a substring from an argument.
	"SUBSTRN Function" on page 1144	Returns a substring, allowing a result with a length of zero.
	"TRANSLATE Function" on page 1166	Replaces specific characters in a character string.
	"TRANSTRN Function" on page 1167	Replaces or removes all occurrences of a substring in a character string.
	"TRANWRD Function" on page 1169	Replaces all occurrences of a substring in a character string.

Category	Functions and CALL Routines	Description
	"TRIM Function" on page 1173	Removes trailing blanks from a character string, and returns one blank if the string is missing.
	"TRIMN Function" on page 1175	Removes trailing blanks from character expressions, and returns a string with a length of zero if the expression is missing.
	"UPCASE Function" on page 1177	Converts all letters in an argument to uppercase.
	"VERIFY Function" on page 1193	Returns the position of the first character in a string that is not in any of several other strings.
Combinatorial	"ALLCOMB Function" on page 374	Generates all combinations of the values of *n* variables taken *k* at a time in a minimal change order.
	"ALLPERM Function" on page 376	Generates all permutations of the values of several variables in a minimal change order.
	"CALL ALLCOMB Routine" on page 432	Generates all combinations of the values of *n* variables taken *k* at a time in a minimal change order.
	"CALL ALLCOMBI Routine" on page 434	Generates all combinations of the indices of *n* objects taken *k* at a time in a minimal change order.
	"CALL ALLPERM Routine" on page 437	Generates all permutations of the values of several variables in a minimal change order.
	"CALL GRAYCODE Routine" on page 450	Generates all subsets of *n* items in a minimal change order.
	"CALL LEXCOMB Routine" on page 458	Generates all distinct combinations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.
	"CALL LEXCOMBI Routine" on page 462	Generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order.
	"CALL LEXPERK Routine" on page 465	Generates all distinct permutations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.
	"CALL LEXPERM Routine" on page 469	Generates all distinct permutations of the non-missing values of several variables in lexicographic order.
	"CALL RANPERK Routine" on page 503	Randomly permutes the values of the arguments, and returns a permutation of *k* out of *n* values.
	"CALL RANPERM Routine" on page 505	Randomly permutes the values of the arguments.
	"COMB Function" on page 590	Computes the number of combinations of *n* elements taken *r* at a time.
	"GRAYCODE Function" on page 797	Generates all subsets of *n* items in a minimal change order.
	"LCOMB Function" on page 880	Computes the logarithm of the COMB function which is the logarithm of the number of combinations of *n* objects taken *r* at a time.

Category	Functions and CALL Routines	Description
	"LEXCOMB Function" on page 889	Generates all distinct combinations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.
	"LEXCOMBI Function" on page 892	Generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order.
	"LEXPERK Function" on page 894	Generates all distinct permutations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.
	"LEXPERM Function" on page 896	Generates all distinct permutations of the non-missing values of several variables in lexicographic order.
	"LFACT Function" on page 899	Computes the logarithm of the FACT (factorial) function.
	"LPERM Function" on page 913	Computes the logarithm of the PERM function which is the logarithm of the number of permutations of *n* objects, with the option of including *r* number of elements.
	"PERM Function" on page 1008	Computes the number of permutations of *n* items that are taken *r* at a time.
Date and Time	"CALL IS8601_CONVERT Routine" on page 454	Converts an ISO 8601 interval to datetime and duration values, or converts datetime and duration values to an ISO 8601 interval.
	"DATDIF Function" on page 632	Returns the number of days between two dates after computing the difference between the dates according to specified day count conventions.
	"DATE Function" on page 634	Returns the current date as a SAS date value.
	"DATEJUL Function" on page 635	Converts a Julian date to a SAS date value.
	"DATEPART Function" on page 636	Extracts the date from a SAS datetime value.
	"DATETIME Function" on page 637	Returns the current date and time of day as a SAS datetime value.
	"DAY Function" on page 637	Returns the day of the month from a SAS date value.
	"DHMS Function" on page 651	Returns a SAS datetime value from date, hour, minute, and second values.
	"HMS Function" on page 803	Returns a SAS time value from hour, minute, and second values.
	"HOLIDAY Function" on page 804	Returns a SAS date value of a specified holiday for a specified year.
	"HOUR Function" on page 807	Returns the hour from a SAS time or datetime value.
	"INTCINDEX Function" on page 830	Returns the cycle index when a date, time, or datetime interval and value are specified.

Category	Functions and CALL Routines	Description
	"INTCK Function" on page 833	Returns the count of the number of interval boundaries between two dates, two times, or two datetime values.
	"INTCYCLE Function" on page 836	Returns the date, time, or datetime interval at the next higher seasonal cycle when a date, time, or datetime interval is specified.
	"INTFIT Function" on page 838	Returns a time interval that is aligned between two dates.
	"INTFMT Function" on page 841	Returns a recommended SAS format when a date, time, or datetime interval is specified.
	"INTGET Function" on page 843	Returns a time interval based on three date or datetime values.
	"INTINDEX Function" on page 845	Returns the seasonal index when a date, time, or datetime interval and value are specified.
	"INTNX Function" on page 848	Increments a date, time, or datetime value by a given time interval, and returns a date, time, or datetime value.
	"INTSEAS Function" on page 855	Returns the length of the seasonal cycle when a date, time, or datetime interval is specified.
	"INTSHIFT Function" on page 857	Returns the shift interval that corresponds to the base interval.
	"INTTEST Function" on page 859	Returns 1 if a time interval is valid, and returns 0 if a time interval is invalid.
	"JULDATE Function" on page 866	Returns the Julian date from a SAS date value.
	"JULDATE7 Function" on page 868	Returns a seven-digit Julian date from a SAS date value.
	"MDY Function" on page 924	Returns a SAS date value from month, day, and year values.
	"MINUTE Function" on page 928	Returns the minute from a SAS time or datetime value.
	"MONTH Function" on page 936	Returns the month from a SAS date value.
	"NWKDOM Function" on page 978	Returns the date for the *n*th occurrence of a weekday for the specified month and year.
	"QTR Function" on page 1063	Returns the quarter of the year from a SAS date value.
	"SECOND Function" on page 1122	Returns the second from a SAS time or datetime value.
	"TIME Function" on page 1161	Returns the current time of day as a numeric SAS time value.
	"TIMEPART Function" on page 1162	Extracts a time value from a SAS datetime value.
	"TODAY Function" on page 1165	Returns the current date as a numeric SAS date value.

Category	Functions and CALL Routines	Description
	"WEEK Function" on page 1226	Returns the week-number value.
	"WEEKDAY Function" on page 1229	From a SAS date value, returns an integer that corresponds to the day of the week.
	"YEAR Function" on page 1233	Returns the year from a SAS date value.
	"YRDIF Function" on page 1235	Returns the difference in years between two dates.
	"YYQ Function" on page 1237	Returns a SAS date value from year and quarter year values.
Descriptive Statistics	"CMISS Function" on page 584	Counts the number of missing arguments.
	"CSS Function" on page 624	Returns the corrected sum of squares.
	"CV Function" on page 626	Returns the coefficient of variation.
	"EUCLID Function" on page 673	Returns the Euclidean norm of the non-missing arguments.
	"GEOMEAN Function" on page 788	Returns the geometric mean.
	"GEOMEANZ Function" on page 790	Returns the geometric mean, using zero fuzzing.
	"HARMEAN Function" on page 799	Returns the harmonic mean.
	"HARMEANZ Function" on page 801	Returns the harmonic mean, using zero fuzzing.
	"IQR Function" on page 864	Returns the interquartile range.
	"KURTOSIS Function" on page 869	Returns the kurtosis.
	"LARGEST Function" on page 877	Returns the *k*th largest non-missing value.
	"LPNORM Function" on page 914	Returns the L*p* norm of the second argument and subsequent non-missing arguments.
	"MAD Function" on page 916	Returns the median absolute deviation from the median.
	"MAX Function" on page 921	Returns the largest value.
	"MEAN Function" on page 925	Returns the arithmetic mean (average).
	"MEDIAN Function" on page 925	Returns the median value.
	"MIN Function" on page 927	Returns the smallest value.

Category	Functions and CALL Routines	Description
	"MISSING Function" on page 929	Returns a numeric result that indicates whether the argument contains a missing value.
	"N Function" on page 943	Returns the number of non-missing numeric values.
	"NMISS Function" on page 947	Returns the number of missing numeric values.
	"ORDINAL Function" on page 982	Returns the k*th* smallest of the missing and nonmissing values.
	"PCTL Function" on page 985	Returns the percentile that corresponds to the percentage.
	"RANGE Function" on page 1084	Returns the range of the nonmissing values.
	"RMS Function" on page 1098	Returns the root mean square of the nonmissing arguments.
	"SKEWNESS Function" on page 1126	Returns the skewness of the nonmissing arguments.
	"SMALLEST Function" on page 1128	Returns the *k*th smallest nonmissing value.
	"STD Function" on page 1133	Returns the standard deviation of the nonmissing arguments.
	"STDERR Function" on page 1134	Returns the standard error of the mean of the nonmissing arguments.
	"SUM Function" on page 1148	Returns the sum of the nonmissing arguments.
	"SUMABS Function" on page 1149	Returns the sum of the absolute values of the non-missing arguments.
	"USS Function" on page 1180	Returns the uncorrected sum of squares of the nonmissing arguments.
	"VAR Function" on page 1182	Returns the variance of the nonmissing arguments.
Distance	"GEODIST Function" on page 786	Returns the geodetic distance between two latitude and longitude coordinates.
	"ZIPCITYDISTANCE Function" on page 1240	Returns the geodetic distance between two ZIP code locations.
External Files	"DCLOSE Function" on page 638	Closes a directory that was opened by the DOPEN function.
	"DCREATE Function" on page 640	Returns the complete pathname of a new, external directory.
	"DINFO Function" on page 657	Returns information about a directory.
	"DNUM Function" on page 660	Returns the number of members in a directory.
	"DOPEN Function" on page 661	Opens a directory, and returns a directory identifier value.

Category	Functions and CALL Routines	Description
	"DOPTNAME Function" on page 662	Returns directory attribute information.
	"DOPTNUM Function" on page 664	Returns the number of information items that are available for a directory.
	"DREAD Function" on page 665	Returns the name of a directory member.
	"DROPNOTE Function" on page 667	Deletes a note marker from a SAS data set or an external file.
	"FAPPEND Function" on page 679	Appends the current record to the end of an external file.
	"FCLOSE Function" on page 680	Closes an external file, directory, or directory member.
	"FCOL Function" on page 681	Returns the current column position in the File Data Buffer (FDB).
	"FDELETE Function" on page 682	Deletes an external file or an empty directory.
	"FEXIST Function" on page 686	Verifies the existence of an external file that is associated with a fileref.
	"FGET Function" on page 687	Copies data from the File Data Buffer (FDB) into a variable.
	"FILEEXIST Function" on page 689	Verifies the existence of an external file by its physical name.
	"FILENAME Function" on page 690	Assigns or deassigns a fileref to an external file, directory, or output device.
	"FILEREF Function" on page 692	Verifies whether a fileref has been assigned for the current SAS session.
	"FINFO Function" on page 749	Returns the value of a file information item.
	"FNOTE Function" on page 760	Identifies the last record that was read, and returns a value that the FPOINT function can use.
	"FOPEN Function" on page 762	Opens an external file and returns a file identifier value.
	"FOPTNAME Function" on page 764	Returns the name of an item of information about a file.
	"FOPTNUM Function" on page 766	Returns the number of information items that are available for an external file.
	"FPOINT Function" on page 767	Positions the read pointer on the next record to be read.
	"FPOS Function" on page 769	Sets the position of the column pointer in the File Data Buffer (FDB).
	"FPUT Function" on page 771	Moves data to the File Data Buffer (FDB) of an external file, starting at the FDB's current column position.

Category	Functions and CALL Routines	Description
	"FREAD Function" on page 772	Reads a record from an external file into the File Data Buffer (FDB).
	"FREWIND Function" on page 773	Positions the file pointer to the start of the file.
	"FRLEN Function" on page 774	Returns the size of the last record that was read, or, if the file is opened for output, returns the current record size.
	"FSEP Function" on page 775	Sets the token delimiters for the FGET function.
	"FWRITE Function" on page 778	Writes a record to an external file.
	"MOPEN Function" on page 936	Opens a file by directory ID and member name, and returns either the file identifier or a 0.
	"PATHNAME Function" on page 983	Returns the physical name of an external file or a SAS library, or returns a blank.
	"RENAME Function" on page 1091	Renames a member of a SAS library, an entry in a SAS catalog, an external file, or a directory.
	"SYSMSG Function" on page 1154	Returns error or warning message text from processing the last data set or external file function.
	"SYSRC Function" on page 1158	Returns a system error number.
External Routines	"CALL MODULE Routine" on page 475	Calls an external routine without any return code.
	"MODULEC Function" on page 933	Calls an external routine and returns a character value.
	"MODULEN Function" on page 933	Calls an external routine and returns a numeric value.
Financial	"BLACKCLPRC Function" on page 418	Calculates call prices for European options on futures, based on the Black model.
	"BLACKPTPRC Function" on page 421	Calculates put prices for European options on futures, based on the Black model.
	"BLKSHCLPRC Function" on page 423	Calculates call prices for European options on stocks, based on the Black-Scholes model.
	"BLKSHPTPRC Function" on page 424	Calculates put prices for European options on stocks, based on the Black-Scholes model.
	"COMPOUND Function" on page 603	Returns compound interest parameters.
	"CONVX Function" on page 612	Returns the convexity for an enumerated cash flow.
	"CONVXP Function" on page 613	Returns the convexity for a periodic cash flow stream, such as a bond.
	"DACCDB Function" on page 626	Returns the accumulated declining balance depreciation.

Category	Functions and CALL Routines	Description
	"DACCDBSL Function" on page 627	Returns the accumulated declining balance with conversion to a straight-line depreciation.
	"DACCSL Function" on page 628	Returns the accumulated straight-line depreciation.
	"DACCSYD Function" on page 629	Returns the accumulated sum-of-years-digits depreciation.
	"DACCTAB Function" on page 630	Returns the accumulated depreciation from specified tables.
	"DEPDB Function" on page 641	Returns the declining balance depreciation.
	"DEPDBSL Function" on page 642	Returns the declining balance with conversion to a straight-line depreciation.
	"DEPSL Function" on page 643	Returns the straight-line depreciation.
	"DEPSYD Function" on page 644	Returns the sum-of-years-digits depreciation.
	"DEPTAB Function" on page 645	Returns the depreciation from specified tables.
	"DUR Function" on page 669	Returns the modified duration for an enumerated cash flow.
	"DURP Function" on page 670	Returns the modified duration for a periodic cash flow stream, such as a bond.
	"FINANCE Function" on page 693	Computes financial calculations such as depreciation, maturation, accrued interest, net present value, periodic savings, and internal rates of return.
	"GARKHCLPRC Function" on page 781	Calculates call prices for European options on stocks, based on the Garman-Kohlhagen model.
	"GARKHPTPRC Function" on page 783	Calculates put prices for European options on stocks, based on the Garman-Kohlhagen model.
	"INTRR Function" on page 853	Returns the internal rate of return as a fraction.
	"IRR Function" on page 865	Returns the internal rate of return as a percentage.
	"MARGRCLPRC Function" on page 917	Calculates call prices for European options on stocks, based on the Margrabe model.
	"MARGRPTPRC Function" on page 919	Calculates put prices for European options on stocks, based on the Margrabe model.
	"MORT Function" on page 939	Returns amortization parameters.
	"NETPV Function" on page 944	Returns the net present value as a fraction.
	"NPV Function" on page 975	Returns the net present value with the rate expressed as a percentage.

Category	Functions and CALL Routines	Description
	"PVP Function" on page 1061	Returns the present value for a periodic cash flow stream (such as a bond), with repayment of principal at maturity.
	"SAVING Function" on page 1110	Returns the future value of a periodic saving.
	"YIELDP Function" on page 1234	Returns the yield-to-maturity for a periodic cash flow stream, such as a bond.
Hyperbolic	"ARCOSH Function" on page 403	Returns the inverse hyperbolic cosine.
	"ARSINH Function" on page 405	Returns the inverse hyperbolic sine.
	"ARTANH Function" on page 406	Returns the inverse hyperbolic tangent.
	"COSH Function" on page 615	Returns the hyperbolic cosine.
	"SINH Function" on page 1125	Returns the hyperbolic sine.
	"TANH Function" on page 1161	Returns the hyperbolic tangent.
Macro	"CALL EXECUTE Routine" on page 449	Resolves the argument, and issues the resolved value for execution at the next step boundary.
	"CALL SYMPUT Routine" on page 536	Assigns DATA step information to a macro variable.
	"CALL SYMPUTX Routine" on page 537	Assigns a value to a macro variable, and removes both leading and trailing blanks.
	"RESOLVE Function" on page 1094	Returns the resolved value of the argument after it has been processed by the macro facility.
	"SYMEXIST Function" on page 1150	Returns an indication of the existence of a macro variable.
	"SYMGET Function" on page 1151	Returns the value of a macro variable during DATA step execution.
	"SYMGLOBL Function" on page 1151	Returns an indication of whether a macro variable is in global scope to the DATA step during DATA step execution.
	"SYMLOCAL Function" on page 1152	Returns an indication of whether a macro variable is in local scope to the DATA step during DATA step execution.
Mathematical	"ABS Function" on page 370	Returns the absolute value.
	"AIRY Function" on page 373	Returns the value of the Airy function.
	"BETA Function" on page 416	Returns the value of the beta function.
	"CALL LOGISTIC Routine" on page 472	Applies the logistic function to each argument.

Category	Functions and CALL Routines	Description
	"CALL SOFTMAX Routine" on page 527	Returns the softmax value.
	"CALL STDIZE Routine" on page 531	Standardizes the values of one or more variables.
	"CALL TANH Routine" on page 539	Returns the hyperbolic tangent.
	"CNONCT Function" on page 585	Returns the noncentrality parameter from a chi-square distribution.
	"COALESCE Function" on page 587	Returns the first non-missing value from a list of numeric arguments.
	"CONSTANT Function" on page 608	Computes machine and mathematical constants.
	"DAIRY Function" on page 631	Returns the derivative of the AIRY function.
	"DEVIANCE Function" on page 648	Returns the deviance based on a probability distribution.
	"DIGAMMA Function" on page 654	Returns the value of the digamma function.
	"ERF Function" on page 672	Returns the value of the (normal) error function.
	"ERFC Function" on page 673	Returns the value of the complementary (normal) error function.
	"EXP Function" on page 677	Returns the value of the exponential function.
	"FACT Function" on page 678	Computes a factorial.
	"FNONCT Function" on page 759	Returns the value of the noncentrality parameter of an F distribution.
	"GAMMA Function" on page 780	Returns the value of the gamma function.
	"GCD Function" on page 785	Returns the greatest common divisor for one or more integers.
	"IBESSEL Function" on page 811	Returns the value of the modified Bessel function.
	"JBESSEL Function" on page 866	Returns the value of the Bessel function.
	"LCM Function" on page 879	Returns the least common multiple.
	"LGAMMA Function" on page 900	Returns the natural logarithm of the Gamma function.
	"LOG Function" on page 903	Returns the natural (base e) logarithm.

Category	Functions and CALL Routines	Description
	"LOG1PX Function" on page 904	Returns the log of 1 plus the argument.
	"LOG10 Function" on page 905	Returns the logarithm to the base 10.
	"LOG2 Function" on page 906	Returns the logarithm to the base 2.
	"LOGBETA Function" on page 906	Returns the logarithm of the beta function.
	"MOD Function" on page 930	Returns the remainder from the division of the first argument by the second argument, fuzzed to avoid most unexpected floating-point results.
	"MODZ Function" on page 934	Returns the remainder from the division of the first argument by the second argument, using zero fuzzing.
	"MSPLINT Function" on page 940	Returns the ordinate of a monotonicity-preserving interpolating spline.
	"SIGN Function" on page 1123	Returns the sign of a value.
	"SQRT Function" on page 1133	Returns the square root of a value.
	"TNONCT Function" on page 1164	Returns the value of the noncentrality parameter from the Student's *t* distribution.
	"TRIGAMMA Function" on page 1172	Returns the value of the trigamma function.
Numeric	"IFN Function" on page 814	Returns a numeric value based on whether an expression is true, false, or missing.
Probability	"CDF Function" on page 558	Returns a value from a cumulative probability distribution.
	"LOGCDF Function" on page 907	Returns the logarithm of a left cumulative distribution function.
	"LOGPDF Function" on page 909	Returns the logarithm of a probability density (mass) function.
	"LOGSDF Function" on page 910	Returns the logarithm of a survival function.
	"PDF Function" on page 986	Returns a value from a probability density (mass) distribution.
	"POISSON Function" on page 1011	Returns the probability from a Poisson distribution.
	"PROBBETA Function" on page 1012	Returns the probability from a beta distribution.
	"PROBBNML Function" on page 1013	Returns the probability from a binomial distribution.
	"PROBBNRM Function" on page 1014	Returns a probability from a bivariate normal distribution.

Category	Functions and CALL Routines	Description
	"PROBCHI Function" on page 1015	Returns the probability from a chi-square distribution.
	"PROBF Function" on page 1016	Returns the probability from an *F* distribution.
	"PROBGAM Function" on page 1017	Returns the probability from a gamma distribution.
	"PROBHYPR Function" on page 1018	Returns the probability from a hypergeometric distribution.
	"PROBMC Function" on page 1020	Returns a probability or a quantile from various distributions for multiple comparisons of means.
	"PROBNEGB Function" on page 1034	Returns the probability from a negative binomial distribution.
	"PROBNORM Function" on page 1035	Returns the probability from the standard normal distribution.
	"PROBT Function" on page 1036	Returns the probability from a *t* distribution.
	"SDF Function" on page 1120	Returns a survival function.
Quantile	"BETAINV Function" on page 418	Returns a quantile from the beta distribution.
	"CINV Function" on page 582	Returns a quantile from the chi-square distribution.
	"FINV Function" on page 750	Returns a quantile from the *F* distribution.
	"GAMINV Function" on page 779	Returns a quantile from the gamma distribution.
	"PROBIT Function" on page 1019	Returns a quantile from the standard normal distribution.
	"QUANTILE Function" on page 1064	Returns the quantile from a distribution that you specify.
	"TINV Function" on page 1162	Returns a quantile from the *t* distribution.
Random Number	"CALL RANBIN Routine" on page 492	Returns a random variate from a binomial distribution.
	"CALL RANCAU Routine" on page 494	Returns a random variate from a Cauchy distribution.
	"CALL RANEXP Routine" on page 497	Returns a random variate from an exponential distribution.
	"CALL RANGAM Routine" on page 499	Returns a random variate from a gamma distribution.
	"CALL RANNOR Routine" on page 501	Returns a random variate from a normal distribution.

Category	Functions and CALL Routines	Description
	"CALL RANPOI Routine" on page 508	Returns a random variate from a Poisson distribution.
	"CALL RANTBL Routine" on page 510	Returns a random variate from a tabled probability distribution.
	"CALL RANTRI Routine" on page 513	Returns a random variate from a triangular distribution.
	"CALL RANUNI Routine" on page 515	Returns a random variate from a uniform distribution.
	"CALL STREAMINIT Routine" on page 535	Specifies a seed value to use for subsequent random number generation by the RAND function.
	"NORMAL Function" on page 948	Returns a random variate from a normal, or Gaussian, distribution.
	"RANBIN Function" on page 1067	Returns a random variate from a binomial distribution.
	"RANCAU Function" on page 1068	Returns a random variate from a Cauchy distribution.
	"RAND Function" on page 1069	Generates random numbers from a distribution that you specify.
	"RANEXP Function" on page 1082	Returns a random variate from an exponential distribution.
	"RANGAM Function" on page 1083	Returns a random variate from a gamma distribution.
	"RANNOR Function" on page 1086	Returns a random variate from a normal distribution.
	"RANPOI Function" on page 1087	Returns a random variate from a Poisson distribution.
	"RANTBL Function" on page 1088	Returns a random variate from a tabled probability distribution.
	"RANTRI Function" on page 1089	Returns a random variate from a triangular distribution.
	"RANUNI Function" on page 1090	Returns a random variate from a uniform distribution.
	"UNIFORM Function" on page 1177	Returns a random variate from a uniform distribution.
SAS File I/O	"ATTRC Function" on page 409	Returns the value of a character attribute for a SAS data set.
	"ATTRN Function" on page 411	Returns the value of a numeric attribute for a SAS data set.
	"CEXIST Function" on page 577	Verifies the existence of a SAS catalog or SAS catalog entry.
	"CLOSE Function" on page 583	Closes a SAS data set.

Category	Functions and CALL Routines	Description
	"CUROBS Function" on page 625	Returns the observation number of the current observation.
	"DROPNOTE Function" on page 667	Deletes a note marker from a SAS data set or an external file.
	"DSNAME Function" on page 668	Returns the SAS data set name that is associated with a data set identifier.
	"ENVLEN Function" on page 671	Returns the length of an environment variable.
	"EXIST Function" on page 675	Verifies the existence of a SAS library member.
	"FETCH Function" on page 684	Reads the next non-deleted observation from a SAS data set into the Data Set Data Vector (DDV).
	"FETCHOBS Function" on page 685	Reads a specified observation from a SAS data set into the Data Set Data Vector (DDV).
	"GETVARC Function" on page 794	Returns the value of a SAS data set character variable.
	"GETVARN Function" on page 795	Returns the value of a SAS data set numeric variable.
	"IORCMSG Function" on page 863	Returns a formatted error message for _IORC_.
	"LIBNAME Function" on page 900	Assigns or deassigns a libref for a SAS library.
	"LIBREF Function" on page 903	Verifies that a libref has been assigned.
	"NOTE Function" on page 956	Returns an observation ID for the current observation of a SAS data set.
	"OPEN Function" on page 980	Opens a SAS data set.
	"PATHNAME Function" on page 983	Returns the physical name of an external file or a SAS library, or returns a blank.
	"POINT Function" on page 1010	Locates an observation that is identified by the NOTE function.
	"RENAME Function" on page 1091	Renames a member of a SAS library, an entry in a SAS catalog, an external file, or a directory.
	"REWIND Function" on page 1096	Positions the data set pointer at the beginning of a SAS data set.
	"SYSMSG Function" on page 1154	Returns error or warning message text from processing the last data set or external file function.
	"SYSRC Function" on page 1158	Returns a system error number.
	"VARFMT Function" on page 1182	Returns the format that is assigned to a SAS data set variable.

Category	Functions and CALL Routines	Description
	"VARINFMT Function" on page 1184	Returns the informat that is assigned to a SAS data set variable.
	"VARLABEL Function" on page 1185	Returns the label that is assigned to a SAS data set variable.
	"VARLEN Function" on page 1186	Returns the length of a SAS data set variable.
	"VARNAME Function" on page 1187	Returns the name of a SAS data set variable.
	"VARNUM Function" on page 1188	Returns the number of a variable's position in a SAS data set.
	"VARTYPE Function" on page 1191	Returns the data type of a SAS data set variable.
Search	"WHICHC Function" on page 1230	Searches for a character value that is equal to the first argument, and returns the index of the first matching value.
	"WHICHN Function" on page 1231	Searches for a numeric value that is equal to the first argument, and returns the index of the first matching value.
Sort	"CALL SORTC Routine" on page 528	Sorts the values of character arguments.
	"CALL SORTN Routine" on page 530	Sorts the values of numeric arguments.
Special	"ADDR Function" on page 371	Returns the memory address of a variable on a 32–bit platform.
	"ADDRLONG Function" on page 372	Returns the memory address of a variable on 32-bit and 64-bit platforms.
	"CALL POKE Routine" on page 477	Writes a value directly into memory on a 32-bit platform.
	"CALL POKELONG Routine" on page 478	Writes a value directly into memory on 32-bit and 64-bit platforms.
	"CALL SLEEP Routine" on page 526	For a specified period of time, suspends the execution of a program that invokes this CALL routine.
	"CALL SYSTEM Routine" on page 538	Submits an operating environment command for execution.
	"DIF Function" on page 653	Returns differences between an argument and its *n*th lag.
	"GETOPTION Function" on page 791	Returns the value of a SAS system or graphics option.
	"INPUT Function" on page 823	Returns the value that is produced when SAS converts an expression using the specified informat.
	"INPUTC Function" on page 826	Enables you to specify a character informat at run time.

Category	Functions and CALL Routines	Description
	“INPUTN Function” on page 827	Enables you to specify a numeric informat at run time.
	“LAG Function” on page 870	Returns values from a queue.
	“PEEK Function” on page 1001	Stores the contents of a memory address in a numeric variable on a 32–bit platform.
	“PEEKC Function” on page 1002	Stores the contents of a memory address in a character variable on a 32–bit platform.
	“PEEKCLONG Function” on page 1005	Stores the contents of a memory address in a character variable on 32-bit and 64-bit platforms.
	“PEEKLONG Function” on page 1007	Stores the contents of a memory address in a numeric variable on 32-bit and 64-bit platforms.
	“PTRLONGADD Function” on page 1056	Returns the pointer address as a character variable on 32-bit and 64-bit platforms.
	“PUT Function” on page 1056	Returns a value using a specified format.
	“PUTC Function” on page 1058	Enables you to specify a character format at run time.
	“PUTN Function” on page 1060	Enables you to specify a numeric format at run time.
	“SLEEP Function” on page 1127	For a specified period of time, suspends the execution of a program that invokes this function.
	“SYSGET Function” on page 1153	Returns the value of the specified operating environment variable.
	“SYSPARM Function” on page 1155	Returns the system parameter string.
	“SYSPROCESSID Function” on page 1155	Returns the process ID of the current process.
	“SYSPROCESSNAME Function” on page 1156	Returns the process name that is associated with a given process ID, or returns the name of the current process.
	“SYSPROD Function” on page 1157	Determines whether a product is licensed.
	“SYSTEM Function” on page 1159	Issues an operating environment command during a SAS session, and returns the system return code.
	“UUIDGEN Function” on page 1181	Returns the short or binary form of a Universal Unique Identifier (UUID).
State and ZIP Code	“FIPNAME Function” on page 752	Converts two-digit FIPS codes to uppercase state names.
	“FIPNAMEL Function” on page 753	Converts two-digit FIPS codes to mixed case state names.
	“FIPSTATE Function” on page 754	Converts two-digit FIPS codes to two-character state postal codes.

Category	Functions and CALL Routines	Description
	"STFIPS Function" on page 1134	Converts state postal codes to FIPS state codes.
	"STNAME Function" on page 1136	Converts state postal codes to uppercase state names.
	"STNAMEL Function" on page 1137	Converts state postal codes to mixed case state names.
	"ZIPCITY Function" on page 1238	Returns a city name and the two-character postal code that corresponds to a ZIP code.
	"ZIPCITYDISTANCE Function" on page 1240	Returns the geodetic distance between two ZIP code locations.
	"ZIPFIPS Function" on page 1241	Converts ZIP codes to two-digit FIPS codes.
	"ZIPNAME Function" on page 1243	Converts ZIP codes to uppercase state names.
	"ZIPNAMEL Function" on page 1245	Converts ZIP codes to mixed case state names.
	"ZIPSTATE Function" on page 1246	Converts ZIP codes to two-character state postal codes.
Trigonometric	"ARCOS Function" on page 402	Returns the arccosine.
	"ARSIN Function" on page 404	Returns the arcsine.
	"ATAN Function" on page 407	Returns the arc tangent.
	"ATAN2 Function" on page 408	Returns the arc tangent of the ratio of two numeric variables.
	"COS Function" on page 615	Returns the cosine.
	"SIN Function" on page 1124	Returns the sine.
	"TAN Function" on page 1160	Returns the tangent.
Truncation	"CEIL Function" on page 573	Returns the smallest integer that is greater than or equal to the argument, fuzzed to avoid unexpected floating-point results.
	"CEILZ Function" on page 575	Returns the smallest integer that is greater than or equal to the argument, using zero fuzzing.
	"FLOOR Function" on page 757	Returns the largest integer that is less than or equal to the argument, fuzzed to avoid unexpected floating-point results.
	"FLOORZ Function" on page 758	Returns the largest integer that is less than or equal to the argument, using zero fuzzing.

Category	Functions and CALL Routines	Description
	"FUZZ Function" on page 777	Returns the nearest integer if the argument is within 1E–12 of that integer.
	"INT Function" on page 829	Returns the integer value, fuzzed to avoid unexpected floating-point results.
	"INTZ Function" on page 861	Returns the integer portion of the argument, using zero fuzzing.
	"ROUND Function" on page 1099	Rounds the first argument to the nearest multiple of the second argument, or to the nearest integer when the second argument is omitted.
	"ROUNDE Function" on page 1106	Rounds the first argument to the nearest multiple of the second argument, and returns an even multiple when the first argument is halfway between the two nearest multiples.
	"ROUNDZ Function" on page 1108	Rounds the first argument to the nearest multiple of the second argument, using zero fuzzing.
	"TRUNC Function" on page 1176	Truncates a numeric value to a specified number of bytes.
Variable Control	"CALL LABEL Routine" on page 457	Assigns a variable label to a specified character variable.
	"CALL SET Routine" on page 525	Links SAS data set variables to DATA step or macro variables that have the same name and data type.
	"CALL VNAME Routine" on page 540	Assigns a variable name as the value of a specified variable.
Variable Information	"CALL VNEXT Routine" on page 542	Returns the name, type, and length of a variable that is used in a DATA step.
	"VARRAY Function" on page 1189	Returns a value that indicates whether the specified name is an array.
	"VARRAYX Function" on page 1190	Returns a value that indicates whether the value of the specified argument is an array.
	"VFORMAT Function" on page 1194	Returns the format that is associated with the specified variable.
	"VFORMATD Function" on page 1195	Returns the decimal value of the format that is associated with the specified variable.
	"VFORMATDX Function" on page 1196	Returns the decimal value of the format that is associated with the value of the specified argument.
	"VFORMATN Function" on page 1197	Returns the format name that is associated with the specified variable.
	"VFORMATNX Function" on page 1198	Returns the format name that is associated with the value of the specified argument.
	"VFORMATW Function" on page 1200	Returns the format width that is associated with the specified variable.
	"VFORMATWX Function" on page 1201	Returns the format width that is associated with the value of the specified argument.

Category	Functions and CALL Routines	Description
	"VFORMATX Function" on page 1202	Returns the format that is associated with the value of the specified argument.
	"VINARRAY Function" on page 1203	Returns a value that indicates whether the specified variable is a member of an array.
	"VINARRAYX Function" on page 1204	Returns a value that indicates whether the value of the specified argument is a member of an array.
	"VINFORMAT Function" on page 1205	Returns the informat that is associated with the specified variable.
	"VINFORMATD Function" on page 1206	Returns the decimal value of the informat that is associated with the specified variable.
	"VINFORMATDX Function" on page 1207	Returns the decimal value of the informat that is associated with the value of the specified variable.
	"VINFORMATN Function" on page 1208	Returns the informat name that is associated with the specified variable.
	"VINFORMATNX Function" on page 1209	Returns the informat name that is associated with the value of the specified argument.
	"VINFORMATW Function" on page 1211	Returns the informat width that is associated with the specified variable.
	"VINFORMATWX Function" on page 1212	Returns the informat width that is associated with the value of the specified argument.
	"VINFORMATX Function" on page 1213	Returns the informat that is associated with the value of the specified argument.
	"VLABEL Function" on page 1214	Returns the label that is associated with the specified variable.
	"VLABELX Function" on page 1215	Returns the label that is associated with the value of the specified argument.
	"VLENGTH Function" on page 1217	Returns the compile-time (allocated) size of the specified variable.
	"VLENGTHX Function" on page 1218	Returns the compile-time (allocated) size for the variable that has a name that is the same as the value of the argument.
	"VNAME Function" on page 1219	Returns the name of the specified variable.
	"VNAMEX Function" on page 1220	Validates the value of the specified argument as a variable name.
	"VTYPE Function" on page 1221	Returns the type (character or numeric) of the specified variable.
	"VTYPEX Function" on page 1222	Returns the type (character or numeric) for the value of the specified argument.
	"VVALUE Function" on page 1224	Returns the formatted value that is associated with the variable that you specify.
	"VVALUEX Function" on page 1225	Returns the formatted value that is associated with the argument that you specify.

Category	Functions and CALL Routines	Description
Web Tools	"HTMLDECODE Function" on page 808	Decodes a string that contains HTML numeric character references or HTML character entity references, and returns the decoded string.
	"HTMLENCODE Function" on page 809	Encodes characters using HTML character entity references, and returns the encoded string.
	"URLDECODE Function" on page 1178	Returns a string that was decoded using the URL escape syntax.
	"URLENCODE Function" on page 1179	Returns a string that was encoded using the URL escape syntax.

Dictionary

ABS Function

Returns the absolute value.

Category: Mathematical

Syntax

ABS (*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The ABS function returns a nonnegative number that is equal in magnitude to the magnitude of the argument.

Examples

SAS Statements	Results
`x=abs(2.4);`	`2.4`
`x=abs(-3);`	`3`

ADDR Function

Returns the memory address of a variable on a 32–bit platform.

Category: Special

Restriction: Use on 32–bit platforms only.

Syntax

ADDR(*variable*)

Arguments

variable
specifies a variable name.

Details

The value that is returned is numeric. Because the storage location of a variable can vary from one execution to the next, the value that is returned by ADDR can vary. The ADDR function is used mostly in combination with the PEEK and PEEKC functions and the CALL POKE routine.

You cannot use the ADDR function on 64-bit platforms. If you attempt to use it, SAS writes a message to the log stating that this restriction applies. If you have legacy applications that use ADDR, change the applications and use ADDRLONG instead. You can use ADDRLONG on both 32–bit and 64–bit platforms.

Comparisons

The ADDR function returns the memory address of a variable on a 32–bit platform. ADDRLONG returns the memory address of a variable on 32–bit and 64–bit platforms.

Note: SAS recommends that you use ADDRLONG instead of ADDR because ADDRLONG can be used on both 32–bit and 64–bit platforms. △

Examples

The following example returns the address at which the variable FIRST is stored:

```
data numlist;
   first=3;
   x=addr(first);
run;
```

See Also

CALL Routine:

"CALL POKE Routine" on page 477

Functions:

"PEEK Function" on page 1001
"PEEKC Function" on page 1002
"ADDRLONG Function" on page 372

ADDRLONG Function

Returns the memory address of a variable on 32-bit and 64-bit platforms.

Category: Special

Syntax

ADDRLONG(*variable*)

Arguments

variable
specifies a variable.

Details

The return value is a character string that contains the binary representation of the address. To display this value, use the $HEX*w*. format to convert the binary value to its hexadecimal equivalent. If you store the result in a variable, that variable should be a character variable with a length of at least eight characters for portability. If you assign the result to a variable that does not yet have a length defined, that variable is given a length of 20 characters.

Examples

The following example returns the pointer address for the variable ITEM, and formats the value.

```
data characterlist;
   item=6345;
   x=addrlong(item);
   put x $hex16.;
run;
```

The following line is written to the SAS log:

```
480063B020202020
```

AIRY Function

Returns the value of the Airy function.

Category: Mathematical

Syntax

AIRY(*x*)

Arguments

x
specifies a numeric constant, variable, or expression.

Details

The AIRY function returns the value of the Airy function (Abramowitz and Stegun 1964; Amos, Daniel and Weston 1977) (See "References" on page 1255). It is the solution of the differential equation

$$w^{(2)} - xw = 0$$

with the conditions

$$w(0) = \frac{1}{3^{\frac{2}{3}}\Gamma\left(\frac{2}{3}\right)}$$

and

$$w'(0) = -\frac{1}{3^{\frac{1}{3}}\Gamma\left(\frac{1}{3}\right)}$$

Examples

SAS Statements	Results
`x=airy(2.0);`	`0.0349241304`
`x=airy(-2.0);`	`0.2274074282`

ALLCOMB Function

Generates all combinations of the values of *n* variables taken *k* at a time in a minimal change order.

Category: Combinatorial

Restriction: The ALLCOMB function cannot be executed when you use the %SYSFUNC macro.

Syntax

ALLCOMB(*count*, *k*, *variable-1, … , variable-n*)

Arguments

count
specifies an integer variable that is assigned values from 1 to the number of combinations in a loop.

k
specifies an integer constant, variable, or expression between 1 and *n*, inclusive, that specifies the number of items in each combination.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before executing the ALLCOMB function.

Restriction: Specify no more than 33 items. If you need to find combinations of more than 33 items, use the CALL ALLCOMBI routine.

Tip: After executing ALLCOMB, the first *k* variables contain the values in one combination.

Details

Use the ALLCOMB function in a loop where the first argument to ALLCOMB accepts each integral value from 1 to the number of combinations, and where *k* is constant. The number of combinations can be computed by using the COMB function. On the first execution, the argument types and lengths are checked for consistency. On each subsequent execution, the values of two variables are interchanged.

For the ALLCOMB function, the following actions occur:

- On the first execution, ALLCOMB returns 0.
- If the values of *variable-i* and *variable-j* were interchanged, where$i<j$, then ALLCOMB returns *i*.
- If no values were interchanged because all combinations were already generated, then ALLCOMB returns –1.

If you execute the ALLCOMB function with the first argument out of sequence, the results are not useful. In particular, if you initialize the variables and then immediately execute the ALLCOMB function with a first argument of *j*, then you will not get the j^{th} combination (except when *j* is 1). To get the j^{th} combination, you must execute ALLCOMB *j* times, with the first argument taking values from 1 through *j* in that exact order.

Comparisons

SAS provides four functions or CALL routines for generating combinations:

- ALLCOMB generates all *possible* combinations of the *values*, *missing or nonmissing*, of N variables. The values can be any numeric or character values. Each combination is formed from the previous combination by removing one value and inserting another value.
- LEXCOMB generates all *distinct* combinations of the *nonmissing values* of several variables. The values can be any numeric or character values. The combinations are generated in lexicographic order.
- ALLCOMBI generates all combinations of the *indices* of N items, where *indices* are integers from 1 to N. Each combination is formed from the previous combination by removing one index and inserting another index.
- LEXCOMBI generates all combinations of the *indices* of N items, where *indices* are integers from 1 to N. The combinations are generated in lexicographic order.

ALLCOMBI is the fastest of these functions and CALL routines. LEXCOMB is the slowest.

Examples

The following is an example of the ALLCOMB function.

```
data _null_;
   array x[5] $3 ('ant' 'bee' 'cat' 'dog' 'ewe');
   n=dim(x);
   k=3;
   ncomb=comb(n,k);
   do j=1 to ncomb+1;
      rc=allcomb(j, k, of x[*]);
      put j 5. +3 x1-x3 +3 rc=;
   end;
run;
```

SAS writes the following output to the log:

```
    1   ant bee cat    rc=0
    2   ant bee ewe    rc=3
    3   ant bee dog    rc=3
    4   ant cat dog    rc=2
    5   ant cat ewe    rc=3
    6   ant dog ewe    rc=2
    7   bee dog ewe    rc=1
    8   bee dog cat    rc=3
    9   bee ewe cat    rc=2
   10   dog ewe cat    rc=1
   11   dog ewe cat    rc=-1
```

See Also

Functions and CALL Routines:

"CALL ALLCOMB Routine" on page 432

ALLPERM Function

Generates all permutations of the values of several variables in a minimal change order.

Category: Combinatorial

Syntax

ALLPERM(*count*, *variable-1* <,*variable-2* ...>)

Arguments

count
specifies a variable with an integer value that ranges from 1 to the number of permutations.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before you execute the ALLPERM function.

Restriction: Specify no more than 18 variables.

Details

The Basics Use the ALLPERM function in a loop where the first argument to ALLPERM accepts each integral value from 1 to the number of permutations. On the first execution, the argument types and lengths are checked for consistency. On each subsequent execution, the values of two consecutive variables are interchanged.

Note: You can compute the number of permutations by using the PERM function. See "PERM Function" on page 1008 for more information.

For the ALLPERM function, the following values are returned:

- 0 if *count*=1
- J if the values of *variable-J* and *variable-K* are interchanged, where K=J+1
- –1 if *count*>N!

△

If you use the ALLPERM function and the first argument is out of sequence, the results are not useful. For example, if you initialize the variables and then immediately execute the ALLPERM function with a first argument of K, your result will not be the K*th* permutation (except when K is 1). To get the K*th* permutation, you must execute the ALLPERM function K times, with the first argument taking values from 1 through K in that exact order.

ALLPERM always produces N! permutations even if some of the variables have equal values or missing values. If you want to generate only the distinct permutations when there are equal values, or if you want to omit missing values from the permutations, use the LEXPERM function instead.

Note: The ALLPERM function cannot be executed when you use the %SYSFUNC macro. △

Comparisons

SAS provides three functions or CALL routines for generating all permutations:

- ALLPERM generates all *possible* permutations of the values, *missing or non-missing*, of several variables. Each permutation is formed from the previous permutation by interchanging two consecutive values.
- LEXPERM generates all *distinct* permutations of the *non-missing* values of several variables. The permutations are generated in lexicographic order.
- LEXPERK generates all *distinct* permutations of K of the *non-missing* values of N variables. The permutations are generated in lexicographic order.

ALLPERM is the fastest of these functions and CALL routines. LEXPERK is the slowest.

Examples

The following example generates permutations of given values by using the ALLPERM function.

```
data _null_;
   array x [4] $3 ('ant' 'bee' 'cat' 'dog');
   n=dim(x);
   nfact=fact(n);
   do i=1 to nfact+1;
      change=allperm(i, of x[*]);
      put i 5. +2 change +2 x[*];
   end;
run;
```

SAS writes the following output to the log:

```
    1  0   ant bee cat dog
    2  3   ant bee dog cat
    3  2   ant dog bee cat
    4  1   dog ant bee cat
    5  3   dog ant cat bee
    6  1   ant dog cat bee
    7  2   ant cat dog bee
    8  3   ant cat bee dog
    9  1   cat ant bee dog
   10  3   cat ant dog bee
   11  2   cat dog ant bee
   12  1   dog cat ant bee
   13  3   dog cat bee ant
   14  1   cat dog bee ant
   15  2   cat bee dog ant
   16  3   cat bee ant dog
   17  1   bee cat ant dog
   18  3   bee cat dog ant
   19  2   bee dog cat ant
   20  1   dog bee cat ant
   21  3   dog bee ant cat
   22  1   bee dog ant cat
   23  2   bee ant dog cat
   24  3   bee ant cat dog
   25  -1   bee ant cat dog
```

See Also

Functions and CALL Routines:

"CALL ALLPERM Routine" on page 437
"LEXPERM Function" on page 896
"CALL RANPERK Routine" on page 503
"CALL RANPERM Routine" on page 505

ANYALNUM Function

Searches a character string for an alphanumeric character, and returns the first position at which the character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYALNUM(*string* <,*start*>)

Arguments

string
specifies a character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYALNUM function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYALNUM function searches a string for the first occurrence of any character that is a digit or an uppercase or lowercase letter. If such a character is found, ANYALNUM returns the position in the string of that character. If no such character is found, ANYALNUM returns a value of 0.

If you use only one argument, ANYALNUM begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYALNUM returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYALNUM function searches a character string for an alphanumeric character. The NOTALNUM function searches a character string for a non-alphanumeric character.

Examples

Example 1: Scanning a String from Left to Right The following example uses the ANYALNUM function to search a string from left to right for alphanumeric characters.

```
data _null_;
   string='Next = Last + 1;';
   j=0;
   do until(j=0);
      j=anyalnum(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=8 c=L
   j=9 c=a
   j=10 c=s
   j=11 c=t
   j=15 c=1
   That's all
```

Example 2: Scanning a String from Right to Left The following example uses the ANYALNUM function to search a string from right to left for alphanumeric characters.

```
data _null_;
   string='Next = Last + 1;';
   j=999999;
   do until(j=0);
      j=anyalnum(string,1-j);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=15 c=1
   j=11 c=t
   j=10 c=s
   j=9 c=a
   j=8 c=L
   j=4 c=t
   j=3 c=x
```

```
j=2 c=e
 j=1 c=N
 That's all
```

See Also

Function:

"NOTALNUM Function" on page 948

ANYALPHA Function

Searches a character string for an alphabetic character, and returns the first position at which the character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYALPHA(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYALPHA function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option"and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYALPHA function searches a string for the first occurrence of any character that is an uppercase or lowercase letter. If such a character is found, ANYALPHA returns the position in the string of that character. If no such character is found, ANYALPHA returns a value of 0.

If you use only one argument, ANYALPHA begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYALPHA returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYALPHA function searches a character string for an alphabetic character. The NOTALPHA function searches a character string for a non-alphabetic character.

Examples

Example 1: Searching a String for Alphabetic Characters The following example uses the ANYALPHA function to search a string from left to right for alphabetic characters.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anyalpha(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=9 c=n
   j=16 c=E
   That's all
```

Example 2: Identifying Control Characters by Using the ANYALPHA Function You can execute the following program to show the control characters that are identified by the ANYALPHA function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   anyalpha=anyalpha(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"NOTALPHA Function" on page 950

ANYCNTRL Function

Searches a character string for a control character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYCNTRL(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYCNTRL function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in the *SAS National Language Support (NLS): Reference Guide*.

The ANYCNTRL function searches a string for the first occurrence of a control character. If such a character is found, ANYCNTRL returns the position in the string of that character. If no such character is found, ANYCNTRL returns a value of 0.

If you use only one argument, ANYCNTRL begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYCNTRL returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYCNTRL function searches a character string for a control character. The NOTCNTRL function searches a character string for a character that is not a control character.

Examples

You can execute the following program to show the control characters that are identified by the ANYCNTRL function.

```
data test;
do dec=0 to 255;
   drop byte;
   byte=byte(dec);
   hex=put(dec,hex2.);
   anycntrl=anycntrl(byte);
   if anycntrl then output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"NOTCNTRL Function" on page 952

ANYDIGIT Function

Searches a character string for a digit, and returns the first position at which the digit is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYDIGIT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The ANYDIGIT function does not depend on the TRANTAB, ENCODING, or LOCALE options.

The ANYDIGIT function searches a string for the first occurrence of any character that is a digit. If such a character is found, ANYDIGIT returns the position in the string of that character. If no such character is found, ANYDIGIT returns a value of 0.

If you use only one argument, ANYDIGIT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYDIGIT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYDIGIT function searches a character string for a digit. The NOTDIGIT function searches a character string for any character that is not a digit.

Examples

The following example uses the ANYDIGIT function to search for a character that is a digit.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anydigit(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=14 c=1
   j=15 c=2
   j=17 c=3
   That's all
```

See Also

Function:

"NOTDIGIT Function" on page 954

ANYFIRST Function

Searches a character string for a character that is valid as the first character in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYFIRST(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The ANYFIRST function does not depend on the TRANTAB, ENCODING, or LOCALE options.

The ANYFIRST function searches a string for the first occurrence of any character that is valid as the first character in a SAS variable name under VALIDVARNAME=V7. These characters are the underscore (_) and uppercase or lowercase English letters. If such a character is found, ANYFIRST returns the position in the string of that character. If no such character is found, ANYFIRST returns a value of 0.

If you use only one argument, ANYFIRST begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYFIRST returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYFIRST function searches a string for the first occurrence of any character that is valid as the first character in a SAS variable name under VALIDVARNAME=V7. The NOTFIRST function searches a string for the first occurrence of any character that is not valid as the first character in a SAS variable name under VALIDVARNAME=V7.

Examples

The following example uses the ANYFIRST function to search a string for any character that is valid as the first character in a SAS variable name under VALIDVARNAME=V7.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anyfirst(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=8 c=_
   j=9 c=n
   j=10 c=_
   j=16 c=E
   That's all
```

See Also

Function:

"NOTFIRST Function" on page 957

ANYGRAPH Function

Searches a character string for a graphical character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYGRAPH(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYGRAPH function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYGRAPH function searches a string for the first occurrence of a graphical character. A graphical character is defined as any printable character other than white space. If such a character is found, ANYGRAPH returns the position in the string of that character. If no such character is found, ANYGRAPH returns a value of 0.

If you use only one argument, ANYGRAPH begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYGRAPH returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYGRAPH function searches a character string for a graphical character. The NOTGRAPH function searches a character string for a non-graphical character.

Examples

Example 1: Searching a String for Graphical Characters The following example uses the ANYGRAPH function to search a string for graphical characters.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anygraph(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=6 c==
   j=8 c=_
   j=9 c=n
   j=10 c=_
   j=12 c=+
   j=14 c=1
   j=15 c=2
   j=16 c=E
   j=17 c=3
   j=18 c=;
   That's all
```

Example 2: Identifying Control Characters by Using the ANYGRAPH Function You can execute the following program to show the control characters that are identified by the ANYGRAPH function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   anygraph=anygraph(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"NOTGRAPH Function" on page 959

ANYLOWER Function

Searches a character string for a lowercase letter, and returns the first position at which the letter is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYLOWER(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYLOWER function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYLOWER function searches a string for the first occurrence of a lowercase letter. If such a character is found, ANYLOWER returns the position in the string of that character. If no such character is found, ANYLOWER returns a value of 0.

If you use only one argument, ANYLOWER begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYLOWER returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYLOWER function searches a character string for a lowercase letter. The NOTLOWER function searches a character string for a character that is not a lowercase letter.

Examples

The following example uses the ANYLOWER function to search a string for any character that is a lowercase letter.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anylower(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=9 c=n
   That's all
```

See Also

Function:

"NOTLOWER Function" on page 961

ANYNAME Function

Searches a character string for a character that is valid in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYNAME(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The ANYNAME function does not depend on the TRANTAB, ENCODING, or LOCALE system options.

The ANYNAME function searches a string for the first occurrence of any character that is valid in a SAS variable name under VALIDVARNAME=V7. These characters are the underscore (_), digits, and uppercase or lowercase English letters. If such a character is found, ANYNAME returns the position in the string of that character. If no such character is found, ANYNAME returns a value of 0.

If you use only one argument, ANYNAME begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYNAME returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYNAME function searches a string for the first occurrence of any character that is valid in a SAS variable name under VALIDVARNAME=V7. The NOTNAME function searches a string for the first occurrence of any character that is not valid in a SAS variable name under VALIDVARNAME=V7.

Examples

The following example uses the ANYNAME function to search a string for any character that is valid in a SAS variable name under VALIDVARNAME=V7.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anyname(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
j=1 c=N
j=2 c=e
j=3 c=x
j=4 c=t
j=8 c=_
j=9 c=n
j=10 c=_
j=14 c=1
j=15 c=2
j=16 c=E
j=17 c=3
That's all
```

See Also

Function:

"NOTNAME Function" on page 963

ANYPRINT Function

Searches a character string for a printable character, and returns the first position at which that character is found.

Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

ANYPRINT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYPRINT function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYPRINT function searches a string for the first occurrence of a printable character. If such a character is found, ANYPRINT returns the position in the string of that character. If no such character is found, ANYPRINT returns a value of 0.

If you use only one argument, ANYPRINT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYPRINT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYPRINT function searches a character string for a printable character. The NOTPRINT function searches a character string for a non-printable character.

Examples

Example 1: Searching a String for a Printable Character The following example uses the ANYPRINT function to search a string for printable characters.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anyprint(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=5 c=
   j=6 c==
   j=7 c=
   j=8 c=_
   j=9 c=n
   j=10 c=_
   j=11 c=
   j=12 c=+
   j=13 c=
   j=14 c=1
   j=15 c=2
   j=16 c=E
   j=17 c=3
   j=18 c=;
   That's all
```

Example 2: Identifying Control Characters by Using the ANYPRINT Function You can execute the following program to show the control characters that are identified by the ANYPRINT function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   anyprint=anyprint(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"NOTPRINT Function" on page 965

ANYPUNCT Function

Searches a character string for a punctuation character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYPUNCT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYPUNCT function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYPUNCT function searches a string for the first occurrence of a punctuation character. If such a character is found, ANYPUNCT returns the position in the string of that character. If no such character is found, ANYPUNCT returns a value of 0.

If you use only one argument, ANYPUNCT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYPUNCT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYPUNCT function searches a character string for a punctuation character. The NOTPUNCT function searches a character string for a character that is not a punctuation character.

Examples

Example 1: Searching a String for Punctuation Characters The following example uses the ANYPUNCT function to search a string for punctuation characters.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anypunct(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=6 c==
   j=8 c=_
   j=10 c=_
   j=12 c=+
   j=18 c=;
   That's all
```

Example 2: Identifying Control Characters by Using the ANYPUNCT Function You can execute the following program to show the control characters that are identified by the ANYPUNCT function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   anypunct=anypunct(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"NOTPUNCT Function" on page 967

ANYSPACE Function

Searches a character string for a white-space character (blank, horizontal and vertical tab, carriage return, line feed, and form feed), and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYSPACE(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYSPACE function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYSPACE function searches a string for the first occurrence of any character that is a blank, horizontal tab, vertical tab, carriage return, line feed, or form feed. If such a character is found, ANYSPACE returns the position in the string of that character. If no such character is found, ANYSPACE returns a value of 0.

If you use only one argument, ANYSPACE begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYSPACE returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYSPACE function searches a character string for the first occurrence of a character that is a blank, horizontal tab, vertical tab, carriage return, line feed, or form feed. The NOTSPACE function searches a character string for the first occurrence of a character that is not a blank, horizontal tab, vertical tab, carriage return, line feed, or form feed.

Examples

Example 1: Searching a String for a White-Space Character The following example uses the ANYSPACE function to search a string for a character that is a white-space character.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anyspace(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=5 c=
   j=7 c=
   j=11 c=
   j=13 c=
   That's all
```

Example 2: Identifying Control Characters by Using the ANYSPACE Function You can execute the following program to show the control characters that are identified by the ANYSPACE function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   anyspace=anyspace(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"NOTSPACE Function" on page 969

ANYUPPER Function

Searches a character string for an uppercase letter, and returns the first position at which the letter is found.

Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

ANYUPPER(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The results of the ANYUPPER function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The ANYUPPER function searches a string for the first occurrence of an uppercase letter. If such a character is found, ANYUPPER returns the position in the string of that character. If no such character is found, ANYUPPER returns a value of 0.

If you use only one argument, ANYUPPER begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYUPPER returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYUPPER function searches a character string for an uppercase letter. The NOTUPPER function searches a character string for a character that is not an uppercase letter.

Examples

The following example uses the ANYUPPER function to search a string for an uppercase letter.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anyupper(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
```

```
  end;
run;
```

The following lines are written to the SAS log:

```
j=1 c=N
j=16 c=E
That's all
```

See Also

Function:

"NOTUPPER Function" on page 971

ANYXDIGIT Function

Searches a character string for a hexadecimal character that represents a digit, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

ANYXDIGIT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional integer that specifies the position at which the search should start and the direction in which to search.

Details

The ANYXDIGIT function does not depend on the TRANTAB, ENCODING, or LOCALE options.

The ANYXDIGIT function searches a string for the first occurrence of any character that is a digit or an uppercase or lowercase A, B, C, D, E, or F. If such a character is found, ANYXDIGIT returns the position in the string of that character. If no such character is found, ANYXDIGIT returns a value of 0.

If you use only one argument, ANYXDIGIT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.

- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

ANYXDIGIT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The ANYXDIGIT function searches a character string for a character that is a hexadecimal character. The NOTXDIGIT function searches a character string for a character that is not a hexadecimal character.

Examples

The following example uses the ANYXDIGIT function to search a string for a hexadecimal character that represents a digit.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=anyxdigit(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=2 c=e
   j=14 c=1
   j=15 c=2
   j=16 c=E
   j=17 c=3
   That's all
```

See Also

Function:

"NOTXDIGIT Function" on page 973

ARCOS Function

Returns the arccosine.

Category: Trigonometric

Syntax

ARCOS (*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Range: between –1 and 1

Details

The ARCOS function returns the arccosine (inverse cosine) of the argument. The value that is returned is specified in radians.

Examples

SAS Statements	Results
`x=arcos(1);`	`0`
`x=arcos(0);`	`1.5707963268`
`x=arcos(-0.5);`	`2.0943951024`

ARCOSH Function

Returns the inverse hyperbolic cosine.

Category: Hyperbolic

Syntax

ARCOSH(*x*)

Arguments

x
specifies a numeric constant, variable, or expression.

Range: $x >= 1$

Details

The ARCOSH function computes the inverse hyperbolic cosine. The ARCOSH function is mathematically defined by the following equation, where $x >= 1$:

$$ARCOSH(x) = log\left(x + \sqrt{x^2 - 1}\right)$$

Examples

The following example computes the inverse hyperbolic cosine.

```
data _null_;
   x=arcosh(5);
   x1=arcosh(13);
   put x=;
   put x1=;
run;
```

SAS writes the following output to the log:

```
x=2.2924316696
x1=3.2566139548
```

See Also

Functions:

"COSH Function" on page 615
"SINH Function" on page 1125
"TANH Function" on page 1161
"ARSINH Function" on page 405
"ARTANH Function" on page 406

ARSIN Function

Returns the arcsine.

Category: Trigonometric

Syntax

ARSIN (*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Range: between –1 and 1

Details

The ARSIN function returns the arcsine (inverse sine) of the argument. The value that is returned is specified in radians.

Examples

SAS Statements	Results
`x=arsin(0);`	`0`
`x=arsin(1);`	`1.5707963268`
`x=arsin(-0.5);`	`-0.523598776`

ARSINH Function

Returns the inverse hyperbolic sine.

Category: Hyperbolic

Syntax

ARSINH(*x*)

Arguments

x
specifies a numeric constant, variable, or expression.
Range: $-\infty < x < \infty$

Details

The ARSINH function computes the inverse hyperbolic sine. The ARSINH function is mathematically defined by the following equation, where $-\infty < x < \infty$:

$$ARSINH\,(x) = log\left(x + \sqrt{x^2 + 1}\right)$$

Replace the infinity symbol with the largest double precision number that is available on your machine.

Examples

The following example computes the inverse hyperbolic sine.

```
data _null_;
   x=arsinh(5);
   x1=arsinh(-5);
   put x=;
   put x1=;
run;
```

SAS writes the following output to the log:

```
x=2.3124383413
x1=-2.312438341
```

See Also

Functions:

"COSH Function" on page 615
"SINH Function" on page 1125
"TANH Function" on page 1161
"ARCOSH Function" on page 403
"ARTANH Function" on page 406

ARTANH Function

Returns the inverse hyperbolic tangent.

Category: Hyperbolic

Syntax

ARTANH(*x*)

Arguments

x
specifies a numeric constant, variable, or expression.

Range: –1 < *x* < 1

Details

The ARTANH function computes the inverse hyperbolic tangent. The ARTANH function is mathematically defined by the following equation, where –1 < x< 1:

$$ARTANH\,(x) = \frac{1}{2} log\left(\frac{1+x}{1-x}\right)$$

Examples

The following example computes the inverse hyperbolic tangent.

```
data _null_;
   x=artanh(0.5);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=0.5493061443
```

See Also

Functions:

"COSH Function" on page 615
"SINH Function" on page 1125
"TANH Function" on page 1161
"ARCOSH Function" on page 403
"ARSINH Function" on page 405

ATAN Function

Returns the arc tangent.

Category: Trigonometric

Syntax

ATAN (*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The ATAN function returns the 2-quadrant arc tangent (inverse tangent) of the argument. The value that is returned is the angle (in radians) whose tangent is x and whose value ranges from $-\pi/2$ to $\pi/2$. If the argument is missing, then ATAN returns a missing value.

Comparisons

The ATAN function is similar to the ATAN2 function except that ATAN2 calculates the arc tangent of the angle from the ratio of two arguments rather than from one argument.

Examples

SAS Statements	Results
`x=atan(0);`	`0`
`x=atan(1);`	`0.7853981634`
`x=atan(-9.0);`	`-1.460139106`

See Also

"ATAN2 Function" on page 408

ATAN2 Function

Returns the arc tangent of the ratio of two numeric variables.

Category: Trigonometric

Syntax

ATAN2(*argument-1*, *argument-2*)

Arguments

argument-1
specifies a numeric constant, variable, or expression.

argument-2
specifies a numeric constant, variable, or expression.

Details

The ATAN2 function returns the arc tangent (inverse tangent) of two numeric variables. The result of this function is similar to the result of calculating the arc tangent of *argument-1* / *argument-2*, except that the signs of both arguments are used to determine the quadrant of the result. ATAN2 returns the result in radians, which is a value between $-\pi$ and π. If either of the arguments in ATAN2 is missing, then ATAN2 returns a missing value.

Comparisons

The ATAN2 function is similar to the ATAN function except that ATAN calculates the arc tangent of the angle from the value of one argument rather than from two arguments.

Examples

SAS statements	Results
`a=atan2(-1, 0.5);`	`-1.107148718`
`b=atan2(6,8);`	`0.6435011088`
`c=atan2(5,-3);`	`2.1112158271`

See Also

Functions:

"ATAN Function" on page 407

ATTRC Function

Returns the value of a character attribute for a SAS data set.

Category: SAS File I/O

Syntax

ATTRC(*data-set-id*,*attr-name*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

attr-name
is an attribute name. If *attr-name* is invalid, a missing value is returned.

Valid values for use with *attr-name* are:

CHARSET
returns a value for the character set of the computer that created the data set.

empty string	Data set not sorted
ASCII	ASCII character set
EBCDIC	EBCDIC character set
ANSI	OS/2 ANSI standard ASCII character set
OEM	OS/2 OEM code format

ENCRYPT
returns 'YES' or 'NO' depending on whether the SAS data set is encrypted.

ENGINE
returns the name of the engine that is used to access the data set.

LABEL
returns the label assigned to the data set.

LIB
returns the libref of the SAS library in which the data set resides.

MEM
returns the SAS data set name.

MODE
returns the mode in which the SAS data set was opened, such as:

I	INPUT mode allows random access if the engine supports it. Otherwise, it defaults to IN mode.
IN	INPUT mode reads sequentially and allows revisiting observations.
IS	INPUT mode reads sequentially but does not allow revisiting observations.
N	NEW mode creates a new data set.
U	UPDATE mode allows random access if the engine supports it. Otherwise, it defaults to UN mode.
UN	UPDATE mode reads sequentially and allows revisiting observations.
US	UPDATE mode reads sequentially but does not allow revisiting observations.
V	UTILITY mode allows modification of variable attributes and indexes associated with the data set.

MTYPE
returns the SAS library member type.

SORTEDBY
returns an empty string if the data set is not sorted. Otherwise, it returns the names of the BY variables in the standard BY statement format.

SORTLVL
returns a value that indicates how a data set was sorted:

Empty string	Data set is not sorted.
WEAK	Sort order of the data set was established by the user (for example, through the SORTEDBY data set option). The system cannot validate its correctness, so the order of observations cannot be depended on.
STRONG	Sort order of the data set was established by the software (for example, through PROC SORT or the OUT= option in the CONTENTS procedure).

SORTSEQ
returns an empty string if the data set is sorted on the native computer or if the sort collating sequence is the default for the operating environment. Otherwise, it returns the name of the alternate collating sequence used to sort the file.

TYPE
returns the SAS data set type.

Examples

- This example generates a message if the SAS data set has not been opened in INPUT SEQUENTIAL mode. The message is written to the SAS log as follows:

```
%let mode=%sysfunc(attrc(&dsid,MODE));
%if &mode ne IS %then
   %put Data set has not been opened in INPUT SEQUENTIAL mode.;
```

- This example tests whether a data set has been sorted and writes the result to the SAS log.

```
data _null_;
   dsid=open("sasdata.sortcars","i");
   charset=attrc(dsid,"CHARSET");
   if charset = "" then
      put "Data set has not been sorted.";
   else put "Data set sorted with " charset
            "character set.";
   rc=close(dsid);
run;
```

See Also

Functions:

"ATTRN Function" on page 411
"OPEN Function" on page 980

ATTRN Function

Returns the value of a numeric attribute for a SAS data set.

Category: SAS File I/O

Syntax

ATTRN(*data-set-id*,*attr-name*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

attr-name
is the name of the SAS data set attribute whose numeric value is returned. If the value of *attr-name* is invalid, a missing value is returned. The following is a list of SAS data set attribute names and their values:

ALTERPW
specifies whether a password is required to alter the data set.

1	the data set is alter protected.
0	the data set is not alter protected.

ANOBS
specifies whether the engine knows the number of observations.

1	the engine knows the number of observations.
0	the engine does not know the number of observations.

ANY
specifies whether the data set has observations or variables.

–1	the data set has no observations or variables.
0	the data set has no observations.
1	the data set has observations and variables.

Alias: VAROBS

ARAND
specifies whether the engine supports random access.

1	the engine supports random access.
0	the engine does not support random access.

Alias: RANDOM

ARWU
specifies whether the engine can manipulate files.

1	the engine is not read-only. It can create or update SAS files.
0	the engine is read-only.

AUDIT
specifies whether logging to an audit file is enabled.

1	logging is enabled.
0	logging is suspended.

AUDIT_DATA
specifies whether after-update record images are stored.

1	after-update record images are stored.
0	after-update record images are not stored.

AUDIT_BEFORE
specifies whether before-update record images are stored.

1	before-update record images are stored.
0	before-update record images are not stored.

AUDIT_ERROR
specifies whether unsuccessful after-update record images are stored.

1	unsuccessful after-update record images are stored.
0	unsuccessful after-update record images are not stored.

CRDTE
specifies the date that the data set was created. The value that is returned is the internal SAS datetime value for the creation date.

Tip: Use the DATETIME. format to display this value.

ICONST
returns information about the existence of integrity constraints for a SAS data set.

0	no integrity constraints.
1	one or more general integrity constraints.
2	one or more referential integrity constraints.
3	both one or more general integrity constraints and one or more referential integrity constraints.

INDEX
specifies whether the data set supports indexing.

1	indexing is supported.
0	indexing is not supported.

ISINDEX
specifies whether the data set is indexed.

1	at least one index exists for the data set.
0	the data set is not indexed.

ISSUBSET
specifies whether the data set is a subset.

1	at least one WHERE clause is active.
0	no WHERE clause is active.

LRECL
specifies the logical record length.

LRID
specifies the length of the record ID.

MAXGEN
specifies the maximum number of generations.

MAXRC
specifies whether an application checks return codes.

1	an application checks return codes.
0	an application does not check return codes.

MODTE
specifies the last date and time that the data set was modified. The value returned is the internal SAS datetime value.

Tip: Use the DATETIME. format to display this value.

NDEL
specifies the number of observations in the data set that are marked for deletion.

NEXTGEN
specifies the next generation number to generate.

NLOBS
specifies the number of logical observations (the observations that are not marked for deletion). An active WHERE clause does not affect this number.

–1 the number of observations is not available.

NLOBSF
specifies the number of logical observations (the observations that are not marked for deletion) by forcing each observation to be read and by taking the FIRSTOBS system option, the OBS system option, and the WHERE clauses into account.

Tip: Passing NLOBSF to ATTRN requires the engine to read every observation from the data set that matches the WHERE clause. Based on the file type and file size, reading these observations can be a time-consuming process.

NOBS
specifies the number of physical observations (including the observations that are marked for deletion). An active WHERE clause does not affect this number.

–1 the number of observations is not available.

NVARS
specifies the number of variables in the data set.

PW
specifies whether a password is required to access the data set.

1 the data set is protected.

0 the data set is not protected.

RADIX
specifies whether access by observation number (radix addressability) is allowed.

1 access by observation number is allowed.

0 access by observation number is not allowed.

Note: A data set that is accessed by a tape engine is index addressable although it cannot be accessed by an observation number.

READPW
specifies whether a password is required to read the data set.

1 the data set is read protected.

0 the data set is not read protected.

TAPE
specifies the status of the data set tape.

1 the data set is a sequential file.

0 the data set is not a sequential file.

WHSTMT
specifies the active WHERE clauses.

0	no WHERE clause is active.
1	a permanent WHERE clause is active.
2	a temporary WHERE clause is active.
3	both permanent and temporary WHERE clauses are active.

WRITEPW
specifies whether a password is required to write to the data set.

1	the data set is write protected.
0	the data set is not write protected.

Examples

□ This example checks whether a WHERE clause is currently active for a data set.

```
%let iswhere=%sysfunc(attrn(&dsid,whstmt));
%if &iswhere %then
   %put A WHERE clause is currently active.;
```

□ This example checks whether a data set is indexed.

```
data _null_;
   dsid=open("mydata");
   isindex=attrn(dsid,"isindex");
   if isindex then put "data set is indexed";
   else put "data set is not indexed";
run;
```

□ This example checks whether a data set is protected with a password.

```
data _null_;
   dsid=open("mydata");
   pw=attrn(dsid,"pw");
   if pw then put "data set is protected";
run;
```

See Also

Functions:

"ATTRC Function" on page 409
"OPEN Function" on page 980

BAND Function

Returns the bitwise logical AND of two arguments.

Category: Bitwise Logical Operations

Syntax

band(*argument-1,argument-2*)

Arguments

argument-1, argument-2
specifies a numeric constant, variable, or expression.

Range: between 0 and (2^{32})-1 inclusive

Details

If either argument contains a missing value, then the function returns a missing value and sets _ERROR_ equal to 1.

Examples

SAS Statements	Results
`x=band(0Fx,05x);` `put x=hex.;`	`x=00000005`

BETA Function

Returns the value of the beta function.

Category: Mathematical

Syntax

BETA(*a*,*b*)

Arguments

a
is the first shape parameter, where $a>0$.

b
is the second shape parameter, where $b>0$.

Details

The BETA function is mathematically given by the equation

$$\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$$

with $a>0$, $b>0$. It should be noted that

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

where $\Gamma(.)$ is the gamma function.
If the expression cannot be computed, BETA returns a missing value.

Examples

SAS Statements	Results
`x=beta(5,3);`	`0.9523809524e-2`

See Also

Function:
"LOGBETA Function" on page 906

BETAINV Function

Returns a quantile from the beta distribution.

Category: Quantile

Syntax

BETAINV (*p,a,b*)

Arguments

p
is a numeric probability.

Range: $0 \leq p \leq 1$

a
is a numeric shape parameter.

Range: $a > 0$

b
is a numeric shape parameter.

Range: $b > 0$

Details

The BETAINV function returns the *p*th quantile from the beta distribution with shape parameters *a* and *b*. The probability that an observation from a beta distribution is less than or equal to the returned quantile is *p*.

Note: BETAINV is the inverse of the PROBBETA function. △

Examples

SAS Statements	Results
`x=betainv(0.001,2,4);`	`0.0101017879`

See Also

Functions:

"QUANTILE Function" on page 1064

BLACKCLPRC Function

Calculates call prices for European options on futures, based on the Black model.

Category: Financial

Syntax

BLACKCLPRC(*E*, *t*, *F*, *r*, *sigma*)

Arguments

E
is a non-missing, positive value that specifies exercise price.
Requirement: Specify *E* and *F* in the same units.

t
is a non-missing value that specifies time to maturity.

F
is a non-missing, positive value that specifies future price.
Requirement: Specify *F* and *E* in the same units.

r
is a non-missing, positive fraction that specifies the risk-free interest rate between the present time and *t*.
Requirement: Specify a value for *r* for the same time period as the unit of *t*.

sigma
is a non-missing, positive fraction that specifies the volatility (the square root of the variance of *r*).
Requirement: Specify a value for *sigma* for the same time period as the unit of *t*.

Details

The BLACKCLPRC function calculates call prices for European options on futures, based on the Black model. The function is based on the following relationship:

$$\texttt{CALL} = e^{-rt}\left(FN\left(d_1\right) - EN\left(d_2\right)\right)$$

where

F	specifies future price.
N	specifies the cumulative normal density function.
E	specifies the exercise price of the option.
r	specifies the risk-free interest rate for period *t*.
t	specifies the time to expiration.

$$d_1 = \frac{\left(\ln\left(\frac{F}{E}\right) + \left(\frac{\sigma^2}{2}\right)t\right)}{\sigma\sqrt{t}}$$

$$d_2 = d_1 - \sigma\sqrt{t}$$

where

σ specifies the volatility of the underlying asset.

σ^2 specifies the variance of the rate of return.

For the special case of t=0, the following equation is true:

$$\texttt{CALL} = \max\left(\left(F - E\right), 0\right)$$

For information about the basics of pricing, see “Using Pricing Functions” on page 311.

Comparisons

The BLACKCLPRC function calculates call prices for European options on futures, based on the Black model. The BLACKPTPRC function calculates put prices for European options on futures, based on the Black model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=blackclprc(1000, .5, 950, 4, 2);` `put a;`	`65.335687119`
`b=blackclprc(850, 2.5, 125, 3, 1);` `put b;`	`0.012649067`
`c=blackclprc(7500, .9, 950, 3, 2);` `put c;`	`17.880939441`
`d=blackclprc(5000, -.5, 237, 3, 2);` `put d;`	`0`

See Also

Function:

“BLACKPTPRC Function” on page 421

BLACKPTPRC Function

Calculates put prices for European options on futures, based on the Black model.

Category: Financial

Syntax

BLACKPTPRC(*E*, *t*, *F*, *r*, *sigma*)

Arguments

E
is a non-missing, positive value that specifies exercise price.

Requirement: Specify *E* and *F* in the same units.

t
is a non-missing value that specifies time to maturity.

F
is a non-missing, positive value that specifies future price.

Requirement: Specify *F* and *E* in the same units.

r
is a non-missing, positive fraction that specifies the risk-free interest rate between the present time and *t*.

Requirement: Specify a value for *r* for the same time period as the unit of *t*.

sigma
is a non-missing, positive fraction that specifies the volatility (the square root of the variance of *r*).

Requirement: Specify a value for *sigma* for the same time period as the unit of *t*.

Details

The BLACKPTPRC function calculates put prices for European options on futures, based on the Black model. The function is based on the following relationship:

$$\texttt{PUT} = \texttt{CALL} + e^{-rt}\left(E - F\right)$$

where

E	specifies the exercise price of the option.
r	specifies the risk-free interest rate for period *t*.
t	specifies the time to expiration.
F	specifies future price.

$$d_1 = \frac{\left(\ln\left(\frac{F}{E}\right) + \left(\frac{\sigma^2}{2}\right)t\right)}{\sigma\sqrt{t}}$$
$$d_2 = d_1 - \sigma\sqrt{t}$$

where

σ specifies the volatility of the underlying asset.

σ^2 specifies the variance of the rate of return.

For the special case of t=0, the following equation is true:

$$\texttt{PUT} = \max\left(\left(E - F\right), 0\right)$$

For information about the basics of pricing, see “Using Pricing Functions” on page 311.

Comparisons

The BLACKPTPRC function calculates put prices for European options on futures, based on the Black model. The BLACKCLPRC function calculates call prices for European options on futures, based on the Black model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=blackptprc(1000, .5, 950, 4, 2);` `put a;`	`72.102451281`
`b=blackptprc(850, 2.5, 125, 3, 1);` `put b;`	`0.4136352354`
`c=blackptprc(7500, .9, 950, 3, 2);` `put c;`	`458.07704789`
`d=blackptprc(5000, -.5, 237, 3, 2);` `put d;`	`0`

See Also

Function:

“BLACKCLPRC Function” on page 418

BLKSHCLPRC Function

Calculates call prices for European options on stocks, based on the Black-Scholes model.

Category: Financial

Syntax

BLKSHCLPRC(*E*, *t*, *S*, *r*, *sigma*)

Arguments

E
is a non-missing, positive value that specifies the exercise price.
Requirement: Specify *E* and *S* in the same units.

t
is a non-missing value that specifies the time to maturity.

S
is a non-missing, positive value that specifies the share price.
Requirement: Specify *S* and *E* in the same units.

r
is a non-missing, positive fraction that specifies the risk-free interest rate for period *t*.
Requirement: Specify a value for *r* for the same time period as the unit of *t*.

sigma
is a non-missing, positive fraction that specifies the volatility of the underlying asset.
Requirement: Specify a value for *sigma* for the same time period as the unit of *t*.

Details

The BLKSHCLPRC function calculates the call prices for European options on stocks, based on the Black-Scholes model. The function is based on the following relationship:

$$\texttt{CALL} = SN\left(d_1\right) - EN\left(d_2\right)e^{-rt}$$

where

S	is a non-missing, positive value that specifies the share price.
N	specifies the cumulative normal density function.
E	is a non-missing, positive value that specifies the exercise price of the option.

$$d_1 = \frac{\left(\ln\left(\frac{S}{E}\right) + \left(r + \frac{\sigma^2}{2}\right)t\right)}{\sigma\sqrt{t}}$$

$$d_2 = d_1 - \sigma\sqrt{t}$$

where

t	specifies the time to expiration.
r	specifies the risk-free interest rate for period t.
σ	specifies the volatility (the square root of the variance).
σ^2	specifies the variance of the rate of return.

For the special case of t=0, the following equation is true:

$$\texttt{CALL} = \max\left(\left(S - E\right), 0\right)$$

For information about the basics of pricing, see "Using Pricing Functions" on page 311.

Comparisons

The BLKSHCLPRC function calculates the call prices for European options on stocks, based on the Black-Scholes model. The BLKSHPTPRC function calculates the put prices for European options on stocks, based on the Black-Scholes model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=blkshclprc(1000, .5, 950, 4, 2);` `put a;`	`831.05008469`
`b=blkshclprc(850, 2.5, 125, 3, 1);` `put b;`	`124.53035232`
`c=blkshclprc(7500, .9, 950, 3, 2);` `put c;`	`719.40891129`
`d=blkshclprc(5000, -.5, 237, 3, 2);` `put d;`	`0`

See Also

Function:

"BLKSHPTPRC Function" on page 424

BLKSHPTPRC Function

Calculates put prices for European options on stocks, based on the Black-Scholes model.

Category: Financial

Syntax

BLKSHPTPRC(*E*, *t*, *S*, *r*, *sigma*)

Arguments

E
is a non-missing, positive value that specifies the exercise price.
Requirement: Specify *E* and *S* in the same units.

t
is a non-missing value that specifies the time to maturity.

S
is a non-missing, positive value that specifies the share price.
Requirement: Specify *S* and *E* in the same units.

r
is a non-missing, positive fraction that specifies the risk-free interest rate for period *t*.
Requirement: Specify a value for *r* for the same time period as the unit of *t*.

sigma
is a non-missing, positive fraction that specifies the volatility of the underlying asset.
Requirement: Specify a value for *sigma* for the same time period as the unit of *t*.

Details

The BLKSHPTPRC function calculates the put prices for European options on stocks, based on the Black-Scholes model. The function is based on the following relationship:

$$\texttt{PUT} = \texttt{CALL} - S + Ee^{-rt}$$

where

S is a non-missing, positive value that specifies the share price.

E is a non-missing, positive value that specifies the exercise price of the option.

$$d_1 = \frac{\left(\ln\left(\frac{S}{E}\right) + \left(r + \frac{\sigma^2}{2}\right)t\right)}{\sigma\sqrt{t}}$$
$$d_2 = d_1 - \sigma\sqrt{t}$$

where

t specifies the time to expiration.

r specifies the risk-free interest rate for period *t*.

σ specifies the volatility (the square root of the variance).

σ^2 specifies the variance of the rate of return.

For the special case of t=0, the following equation is true:

$$\texttt{PUT} = \max\left(\left(E - S\right), 0\right)$$

For information about the basics of pricing, see "Using Pricing Functions" on page 311.

Comparisons

The BLKSHPTPRC function calculates the put prices for European options on stocks, based on the Black-Scholes model. The BLKSHCLPRC function calculates the call prices for European options on stocks, based on the Black-Scholes model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=blkshptprc(1000, .5, 950, 4, 2);` `put a;`	`16.385367922`
`b=blkshptprc(850, 1.2, 125, 3, 1);` `put b;`	`1.426971358`
`c=blkshptprc(7500, .9, 950, 3, 2);` `put c;`	`273.45025684`
`d=blkshptprc(5000, -.5, 237, 3, 2);` `put d;`	`0`

See Also

Function:

"BLKSHCLPRC Function" on page 423

BLSHIFT Function

Returns the bitwise logical left shift of two arguments.

Category: Bitwise Logical Operations

Syntax

BLSHIFT(*argument-1,argument-2*)

Arguments

argument-1
specifies a numeric constant, variable, or expression.

Range: between 0 and (2^{32})-1 inclusive

argument-2
specifies a numeric constant, variable, or expression.

Range: 0 to 31, inclusive

Details

If either argument contains a missing value, then the function returns a missing value and sets _ERROR_ equal to 1.

Examples

SAS Statements	Results
`x=blshift(07x,2);` `put x=hex.;`	`x=0000001C`

BNOT Function

Returns the bitwise logical NOT of an argument.

Category: Bitwise Logical Operations

Syntax

BNOT(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Range: between 0 and (2^{32})-1 inclusive

Details

If the argument contains a missing value, then the function returns a missing value and sets _ERROR_ equal to 1.

Examples

SAS Statements	Results
`x=bnot(0F000000Fx);` `put x=hex.;`	`x=0FFFFFF0`

BOR Function

Returns the bitwise logical OR of two arguments.

Category: Bitwise Logical Operations

Syntax

BOR(*argument-1,argument-2*)

Arguments

argument-1, argument-2
specifies a numeric constant, variable, or expression.

Range: between 0 and (2^{32})-1 inclusive

Details

If either argument contains a missing value, then the function returns a missing value and sets _ERROR_ equal to 1.

Examples

SAS Statements	Results
`x=bor(01x,0F4x);` `put x=hex.;`	`x=000000F5`

BRSHIFT Function

Returns the bitwise logical right shift of two arguments.

Category: Bitwise Logical Operations

Syntax

BRSHIFT(*argument-1, argument-2*)

Arguments

argument-1
specifies a numeric constant, variable, or expression.

Range: between 0 and (2^{32})-1 inclusive

argument-2
specifies a numeric constant, variable, or expression.

Range: 0 to 31, inclusive

Details

If either argument contains a missing value, then the function returns a missing value and sets _ERROR_ equal to 1.

Examples

SAS Statements	Results
`x=brshift(01Cx,2);` `put x=hex.;`	`x=00000007`

BXOR Function

Returns the bitwise logical EXCLUSIVE OR of two arguments.

Category: Bitwise Logical Operations

Syntax

BXOR(*argument-1, argument-2*)

Arguments

argument-1, argument-2
specifies a numeric constant, variable, or expression.

Range: between 0 and (2^{32})-1 inclusive

Details

If either argument contains a missing value, then the function returns a missing value and sets _ERROR_ equal to 1.

Examples

SAS Statements	Results
`x=bxor(03x,01x);` `put x=hex.;`	`x=00000002`

BYTE Function

Returns one character in the ASCII or the EBCDIC collating sequence.

Category: Character
Restriction: "I18N Level 0" on page 313
See: BYTE Function in the documentation for your operating environment.

Syntax

BYTE (*n*)

Arguments

n
specifies an integer that represents a specific ASCII or EBCDIC character.
Range: 0–255

Details

Length of Returned Variable In a DATA step, if the BYTE function returns a value to a variable that has not previously been assigned a length, then that variable is assigned a length of 1.

ASCII and EBCDIC Collating Sequences For EBCDIC collating sequences, *n* is between 0 and 255. For ASCII collating sequences, the characters that correspond to values between 0 and 127 represent the standard character set. Other ASCII characters that correspond to values between 128 and 255 are available on certain ASCII operating environments, but the information those characters represent varies with the operating environment.

Examples

SAS Statements	Results	
	ASCII	EBCDIC
	`----+----1----+----2`	`----+----1----+----2`
`x=byte(80);` `put x;`	`P`	`&`

See Also

Functions:
"COLLATE Function" on page 589
"RANK Function" on page 1085

CALL ALLCOMB Routine

Generates all combinations of the values of *n* variables taken *k* at a time in a minimal change order.

Category: Combinatorial

Syntax

CALL ALLCOMB(*count*, *k*, *variable-1*, …, *variable-n*);

Arguments

count
specifies an integer variable that is assigned from 1 to the number of combinations in a loop.

k
specifies an integer constant, variable, or expression between 1 and *n*, inclusive, that specifies the number of items in each combination.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before calling the ALLCOMB routine.

Restriction: Specify no more than 33 items. If you need to find combinations of more than 33 items, use the CALL ALLCOMBI routine.

Tip: After calling the ALLCOMB routine, the first *k* variables contain the values in one combination.

Details

CALL ALLCOMB Processing Use the CALL ALLCOMB routine in a loop where the first argument to CALL ALLCOMB accepts each integral value from 1 to the number of combinations, and where *k* is constant. The number of combinations can be computed by using the COMB function. On the first call, the argument types and lengths are checked for consistency. On each subsequent call, the values of two variables are interchanged.

If you call the ALLCOMB routine with the first argument out of sequence, the results are not useful. In particular, if you initialize the variables and then immediately call ALLCOMB with a first argument of *j*, then you will not get the j^{th} combination (except when *j* is 1). To get the j^{th} combination, you must call ALLCOMB *j* times, with the first argument taking values from 1 through *j* in that exact order.

Using the CALL ALLCOMB Routine with Macros You can call the ALLCOMB routine when you use the %SYSCALL macro. In this case, the *variable* arguments are not required to be the same type or length. If %SYSCALL identifies an argument as numeric, then %SYSCALL reformats the returned value.

If an error occurs during the execution of the CALL ALLCOMB routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.

- &SYSINFO is assigned a value that is less than –100.

If there are no errors, then &SYSERR is set to zero, and &SYSINFO is set to one of the following values:

- 0 if *count*=1
- *j* if the values of *variable-j* and *variable-k* were interchanged, where *j*<*k*
- –1 if no values were interchanged because all distinct combinations were already generated

Comparisons

SAS provides four functions or CALL routines for generating combinations:

- ALLCOMB generates all *possible* combinations of the *values*, *missing or non-missing*, of *n* variables. The values can be any numeric or character values. Each combination is formed from the previous combination by removing one value and inserting another value.
- LEXCOMB generates all *distinct* combinations of the *non-missing values* of several variables. The values can be any numeric or character values. The combinations are generated in lexicographic order.
- ALLCOMBI generates all combinations of the *indices* of *n* items, where *indices* are integers from 1 to *n*. Each combination is formed from the previous combination by removing one index and inserting another index.
- LEXCOMBI generates all combinations of the *indices* of *n* items, where *indices* are integers from 1 to *n*. The combinations are generated in lexicographic order.

ALLCOMBI is the fastest of these functions and CALL routines. LEXCOMB is the slowest.

Examples

Example 1: Using CALL ALLCOMB in a DATA Step The following is an example of the CALL ALLCOMB routine that is used with the DATA step.

```
data _null_;
   array x[5] $3 ('ant' 'bee' 'cat' 'dog' 'ewe');
   n=dim(x);
   k=3;
   ncomb=comb(n,k);
   do j=1 to ncomb+1;
      call allcomb(j, k, of x[*]);
      put j 5. +3 x1-x3;
   end;
run;
```

SAS writes the following output to the log:

```
    1   ant bee cat
    2   ant bee ewe
    3   ant bee dog
    4   ant cat dog
    5   ant cat ewe
    6   ant dog ewe
    7   bee dog ewe
    8   bee dog cat
```

```
9   bee ewe cat
10   dog ewe cat
11   dog ewe cat
```

Example 2: Using CALL ALLCOMB with Macros and Displaying the Return Code The following is an example of the CALL ALLCOMB routine that is used with macros. The output includes values for the %SYSINFO macro.

```
%macro test;
   %let x1=ant;
   %let x2=-.1234;
   %let x3=1e10;
   %let x4=hippopotamus;
   %let x5=zebra;
   %let k=2;
   %let ncomb=%sysfunc(comb(5,&k));
   %do j=1 %to &ncomb+1;
      %syscall allcomb(j, k, x1, x2, x3, x4, x5);
      %let jfmt=%qsysfunc(putn(&j,5.));
      %let pad=%qsysfunc(repeat(%str(),30-%length(&x1 &x2)));
      %put &jfmt:  &x1 &x2 &pad sysinfo=&sysinfo;
   %end;
%mend;

%test
```

SAS writes the following output to the log:

```
1:  ant -0.1234  sysinfo=0
 2:  ant zebra  sysinfo=2
 3:  ant hippopotamus  sysinfo=2
 4:  ant 10000000000  sysinfo=2
 5:  -0.1234 10000000000  sysinfo=1
 6:  -0.1234 zebra  sysinfo=2
 7:  -0.1234 hippopotamus  sysinfo=2
 8:  10000000000 hippopotamus  sysinfo=1
 9:  10000000000 zebra  sysinfo=2
10:  hippopotamus zebra  sysinfo=1
11:  hippopotamus zebra  sysinfo=-1
```

See Also

Functions and CALL Routines:

"ALLCOMB Function" on page 374

CALL ALLCOMBI Routine

Generates all combinations of the indices of *n* objects taken *k* at a time in a minimal change order.

Category: Combinatorial

Syntax

CALL ALLCOMBI(*N*, *K*, *index-1*, …, *index-K*, <, *index-added*, *index-removed*>);

Arguments

N
is a numeric constant, variable, or expression that specifies the total number of objects.

K
is a numeric constant, variable, or expression that specifies the number of objects in each combination.

index
is a numeric variable that contains indices of the objects in the returned combination. Indices are integers between 1 and N inclusive.

Tip: If *index-1* is missing or zero, then ALLCOMBI initializes the indices to **`index-1=1`** through **`index-K=K`**. Otherwise, ALLCOMBI creates a new combination by removing one index from the combination and adding another index.

index-added
is a numeric variable in which ALLCOMBI returns the value of the index that was added.

index-removed
is a numeric variable in which ALLCOMBI returns the value of the index that was removed.

Details

CALL ALLCOMBI Processing Before you make the first call to ALLCOMBI, complete one of the following tasks:

- Set *index-1* equal to zero or to a missing value.
- Initialize *index-1* through *index-K* to distinct integers between 1 and N inclusive.

The number of combinations of N objects taken K at a time can be computed as COMB(*N*, *K*). To generate all combinations of N objects taken K at a time, call ALLCOMBI in a loop that executes COMB(*N*, *K*) times.

Using the CALL ALLCOMBI Routine with Macros If you call ALLCOMBI from the macro processor with %SYSCALL, then you must initialize all arguments to numeric values. &SYSCALL reformats the values that are returned.

If an error occurs during the execution of the CALL ALLCOMBI routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than –100.

If there are no errors, then &SYSERR and &SYSINFO are set to zero.

Comparisons

The CALL ALLCOMBI routine generates all combinations of the indices of N objects taken K at a time in a minimal change order. The CALL ALLCOMB routine generates all combinations of the values of N variables taken K at a time in a minimal change order.

Examples

Example 1: Using CALL ALLCOMBI in a DATA Step The following is an example of the CALL ALLCOMBI routine that is used in a DATA step.

```
data _null_;
   array x[5] $3 ('ant' 'bee' 'cat' 'dog' 'ewe');
   array c[3] $3;
   array i[3];
   n=dim(x);
   k=dim(i);
   i[1]=0;
   ncomb=comb(n,k);    /* The one extra call goes back */
   do j=1 to ncomb+1;  /* to the first combination. */
      call allcombi(n, k, of i[*], add, remove);
      do h=1 to k;
         c[h]=x[i[h]];
      end;
      put @4 j= @10 'i= ' i[*] +3 'c= ' c[*] +3 add= remove=;
   end;
run;
```

SAS writes the following output to the log:

```
j=1   i= 1 2 3    c= ant bee cat    add=0 remove=0
j=2   i= 1 3 4    c= ant cat dog    add=4 remove=2
j=3   i= 2 3 4    c= bee cat dog    add=2 remove=1
j=4   i= 1 2 4    c= ant bee dog    add=1 remove=3
j=5   i= 1 4 5    c= ant dog ewe    add=5 remove=2
j=6   i= 2 4 5    c= bee dog ewe    add=2 remove=1
j=7   i= 3 4 5    c= cat dog ewe    add=3 remove=2
j=8   i= 1 3 5    c= ant cat ewe    add=1 remove=4
j=9   i= 2 3 5    c= bee cat ewe    add=2 remove=1
j=10  i= 1 2 5    c= ant bee ewe    add=1 remove=3
j=11  i= 1 2 3    c= ant bee cat    add=3 remove=5
```

Example 2: Using CALL ALLCOMBI with Macros The following is an example of the CALL ALLCOMBI routine that is used with macros.

```
%macro test;
   %let x1=0;
   %let x2=0;
   %let x3=0;
   %let add=0;
   %let remove=0;
   %let n=5;
   %let k=3;
   %let ncomb=%sysfunc(comb(&n,&k));
   %do j=1 %to &ncomb;
      %syscall allcombi(n,k,x1,x2,x3,add,remove);
```

```
      %let jfmt=%qsysfunc(putn(&j,5.));
      %put &jfmt: &x1 &x2 &x3 add=&add remove=&remove;
   %end;
%mend;

%test
```

SAS writes the following output to the log:

```
    1: 1 2 3 add=0 remove=0
    2: 1 3 4 add=4 remove=2
    3: 2 3 4 add=2 remove=1
    4: 1 2 4 add=1 remove=3
    5: 1 4 5 add=5 remove=2
    6: 2 4 5 add=2 remove=1
    7: 3 4 5 add=3 remove=2
    8: 1 3 5 add=1 remove=4
    9: 2 3 5 add=2 remove=1
   10: 1 2 5 add=1 remove=3
```

Examples

See Also

Functions and CALL Routines:

"CALL ALLCOMB Routine" on page 432

CALL ALLPERM Routine

Generates all permutations of the values of several variables in a minimal change order.

Category: Combinatorial

Syntax

CALL ALLPERM(*count*, *variable–1*<, *variable–2 ...*>);

Arguments

count
specifies an integer variable that ranges from 1 to the number of permutations.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before you call the ALLPERM routine.

Restriction: Specify no more than 18 variables.

Details

CALL ALLPERM Processing Use the CALL ALLPERM routine in a loop where the first argument to CALL ALLPERM takes each integral value from 1 to the number of permutations. On the first call, the argument types and lengths are checked for consistency. On each subsequent call, the values of two consecutive variables are interchanged.

Note: You can compute the number of permutations by using the PERM function. See "PERM Function" on page 1008 for more information. △

If you call the ALLPERM routine and the first argument is out of sequence, the results are not useful. In particular, if you initialize the variables and then immediately call the ALLPERM routine with a first argument of K, your result will not be the K*th* permutation (except when K is 1). To get the K*th* permutation, you must call the ALLPERM routine K times, with the first argument taking values from 1 through K in that exact order.

ALLPERM always produces N! permutations even if some of the variables have equal values or missing values. If you want to generate only the distinct permutations when there are equal values, or if you want to omit missing values from the permutations, use the LEXPERM function instead.

Using the CALL ALLPERM Routine with Macros You can call the ALLPERM routine when you use the %SYSCALL macro. In this case, the *variable* arguments are not required to be the same type or length. If %SYSCALL identifies an argument as numeric, then %SYSCALL reformats the returned value.

If an error occurs during the execution of the CALL ALLPERM routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than -100.

If there are no errors, then &SYSERR is set to zero, and &SYSINFO is set to one of the following values:

- 0 if *count*=1
- J if 1<*count*<=N! and the values of *variable-J* and *variable-K* were interchanged, where J+1=K
- -1 if *count*>N!

Comparisons

SAS provides three functions or CALL routines for generating all permutations:

- ALLPERM generates all *possible* permutations of the values, *missing or nonmissing*, of several variables. Each permutation is formed from the previous permutation by interchanging two consecutive values.
- LEXPERM generates all *distinct* permutations of the *nonmissing* values of several variables. The permutations are generated in lexicographic order.
- LEXPERK generates all *distinct* permutations of K of the *nonmissing* values of N variables. The permutations are generated in lexicographic order.

ALLPERM is the fastest of these functions and CALL routines. LEXPERK is the slowest.

Examples

Example 1: Using CALL ALLPERM in a DATA Step The following example generates permutations of given values by using the CALL ALLPERM routine.

```
data _null_;
   array x [4] $3 ('ant' 'bee' 'cat' 'dog');
   n=dim(x);
   nfact=fact(n);
   do i=1 to nfact;
      call allperm(i, of x[*]);
      put i 5. +2 x[*];
   end;
run;
```

SAS writes the following output to the log:

```
    1  ant bee cat dog
    2  ant bee dog cat
    3  ant dog bee cat
    4  dog ant bee cat
    5  dog ant cat bee
    6  ant dog cat bee
    7  ant cat dog bee
    8  ant cat bee dog
    9  cat ant bee dog
   10  cat ant dog bee
   11  cat dog ant bee
   12  dog cat ant bee
   13  dog cat bee ant
   14  cat dog bee ant
   15  cat bee dog ant
   16  cat bee ant dog
   17  bee cat ant dog
   18  bee cat dog ant
   19  bee dog cat ant
   20  dog bee cat ant
   21  dog bee ant cat
   22  bee dog ant cat
   23  bee ant dog cat
   24  bee ant cat dog
```

Example 2: Using CALL ALLPERM with Macros The following is an example of the CALL ALLPERM routine that is used with macros. The output includes values for the %SYSINFO macro.

```
%macro test;
   %let x1=ant;
   %let x2=-.1234;
   %let x3=1e10;
   %let x4=hippopotamus;
   %let nperm=%sysfunc(perm(4));
   %do j=1 %to &nperm+1;
      %syscall allperm(j, x1, x2, x3, x4);
```

```
      %let jfmt=%qsysfunc(putn(&j,5.));
      %put &jfmt:   &x1 &x2 &x3 &x4 sysinfo=&sysinfo;
   %end;
%mend;

%test;
```

SAS writes the following output to the log:

```
 1:   ant -0.1234 10000000000 hippopotamus sysinfo=0
 2:   ant -0.1234 hippopotamus 10000000000 sysinfo=3
 3:   ant hippopotamus -0.1234 10000000000 sysinfo=2
 4:   hippopotamus ant -0.1234 10000000000 sysinfo=1
 5:   hippopotamus ant 10000000000 -0.1234 sysinfo=3
 6:   ant hippopotamus 10000000000 -0.1234 sysinfo=1
 7:   ant 10000000000 hippopotamus -0.1234 sysinfo=2
 8:   ant 10000000000 -0.1234 hippopotamus sysinfo=3
 9:   10000000000 ant -0.1234 hippopotamus sysinfo=1
10:   10000000000 ant hippopotamus -0.1234 sysinfo=3
11:   10000000000 hippopotamus ant -0.1234 sysinfo=2
12:   hippopotamus 10000000000 ant -0.1234 sysinfo=1
13:   hippopotamus 10000000000 -0.1234 ant sysinfo=3
14:   10000000000 hippopotamus -0.1234 ant sysinfo=1
15:   10000000000 -0.1234 hippopotamus ant sysinfo=2
16:   10000000000 -0.1234 ant hippopotamus sysinfo=3
17:   -0.1234 10000000000 ant hippopotamus sysinfo=1
18:   -0.1234 10000000000 hippopotamus ant sysinfo=3
19:   -0.1234 hippopotamus 10000000000 ant sysinfo=2
20:   hippopotamus -0.1234 10000000000 ant sysinfo=1
21:   hippopotamus -0.1234 ant 10000000000 sysinfo=3
22:   -0.1234 hippopotamus ant 10000000000 sysinfo=1
23:   -0.1234 ant hippopotamus 10000000000 sysinfo=2
24:   -0.1234 ant 10000000000 hippopotamus sysinfo=3
25:   -0.1234 ant 10000000000 hippopotamus sysinfo=-1
```

See Also

Functions and CALL Routines:

“LEXPERM Function” on page 896
“ALLPERM Function” on page 376
“CALL RANPERK Routine” on page 503
“CALL RANPERM Routine” on page 505

CALL CATS Routine

Removes leading and trailing blanks, and returns a concatenated character string.

Category: Character

Syntax

CALL CATS(*result* <, *item-1*, ..., *item-n*>);

Arguments

result
specifies a character variable.

Restriction: The CALL CATS routine accepts only a character variable as a valid argument for *result*. Do not use a constant or a SAS expression because CALL CATS is unable to update these arguments.

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string using the BEST*w*. format. In this case, SAS does not write a note to the log.

Details

The CALL CATS routine returns the result in the first argument, *result*. The routine appends the values of the arguments that follow to *result*. If the length of *result* is not large enough to contain the entire result, SAS does the following:

- writes a warning message to the log stating that the result was truncated
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation, except in SQL or in a WHERE clause
- sets _ERROR_ to 1 in the DATA step, except in a WHERE clause

The CALL CATS routine removes leading and trailing blanks from numeric arguments after it formats the numeric value with the BEST*w*. format.

Comparisons

The results of the CALL CATS, CALL CATT, and CALL CATX routines are usually equivalent to statements that use the concatenation operator (||) and the TRIM and LEFT functions. However, using the CALL CATS, CALL CATT, and CALL CATX routines is faster than using TRIM and LEFT.

The following table shows statements that are equivalent to CALL CATS, CALL CATT, and CALL CATX. The variables X1 through X4 specify character variables, and SP specifies a separator, such as a blank or comma.

CALL Routine	Equivalent Statement
`CALL CATS(OF X1-X4);`	`X1=TRIM(LEFT(X1))||TRIM(LEFT(X2))||TRIM(LEFT(X3))||TRIM(LEFT(X4));`
`CALL CATT(OF X1-X4);`	`X1=TRIM(X1)||TRIM(X2)||TRIM(X3)||TRIM(X4);`
`CALL CATX(SP, OF X1-X4);` *	`X1=TRIM(LEFT(X1))||SP||TRIM(LEFT(X2))||SP||TRIM(LEFT(X3))||SP||TRIM(LEFT(X4));`

* If any of the arguments is blank, the results that are produced by CALL CATX differ slightly from the results that are produced by the concatenated code. In this case, CALL CATX omits the corresponding separator. For example, **`CALL CATX("+","X"," ", "Z"," ");`** produces **`X+Z`**.

Examples

The following example shows how the CALL CATS routine concatenates strings.

```
data _null_;
   length answer $ 36;
   x='Athens is t  ';
   y=' he Olym   ';
   z='   pic site for 2004. ';
   call cats(answer,x,y,z);
   put answer;
run;
```

The following line is written to the SAS log:

```
   ----+----1----+----2----+----3----+----4----+----5----+----6----+----7
   Athens is the Olympic site for 2004.
```

See Also

Functions and CALL Routines:

"CALL CATT Routine" on page 443
"CALL CATX Routine" on page 445
"CAT Function" on page 543
"CATQ Function" on page 546
"CATS Function" on page 550
"CATT Function" on page 552
"CATX Function" on page 554

CALL CATT Routine

Removes trailing blanks, and returns a concatenated character string.

Category: Character

Syntax

CALL CATT(*result* <, *item-1*, … *item-n*>);

Arguments

result
specifies a character variable.

Restriction: The CALL CATT routine accepts only a character variable as a valid argument for *result*. Do not use a constant or a SAS expression because CALL CATT is unable to update these arguments.

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string using the BEST*w*. format. In this case, leading blanks are removed and SAS does not write a note to the log.

Details

The CALL CATT routine returns the result in the first argument, *result*. The routine appends the values of the arguments that follow to *result*. If the length of *result* is not large enough to contain the entire result, SAS does the following:

- writes a warning message to the log stating that the result was truncated
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation, except in SQL or in a WHERE clause
- sets _ERROR_ to 1 in the DATA step, except in a WHERE clause

The CALL CATT routine removes leading and trailing blanks from numeric arguments after it formats the numeric value with the BEST*w*. format.

Comparisons

The results of the CALL CATS, CALL CATT, and CALL CATX routines are usually equivalent to statements that use the concatenation operator (||) and the TRIM and LEFT functions. However, using the CALL CATS, CALL CATT, and CALL CATX routines is faster than using TRIM and LEFT.

The following table shows statements that are equivalent to CALL CATS, CALL CATT, and CALL CATX. The variables X1 through X4 specify character variables, and SP specifies a separator, such as a blank or comma.

CALL Routine	Equivalent Statement
`CALL CATS(OF X1-X4);`	`X1=TRIM(LEFT(X1))||TRIM(LEFT(X2))||TRIM(LEFT(X3))||TRIM(LEFT(X4));`
`CALL CATT(OF X1-X4);`	`X1=TRIM(X1)||TRIM(X2)||TRIM(X3)||TRIM(X4);`
`CALL CATX(SP, OF X1-X4);` *	`X1=TRIM(LEFT(X1))||SP||TRIM(LEFT(X2))||SP||TRIM(LEFT(X3))||SP||TRIM(LEFT(X4));`

* If any of the arguments is blank, the results that are produced by CALL CATX differ slightly from the results that are produced by the concatenated code. In this case, CALL CATX omits the corresponding separator. For example, **`CALL CATX("+","X"," ", "Z"," ");`** produces **`X+Z`**.

Examples

The following example shows how the CALL CATT routine concatenates strings.

```
data _null_;
   length answer $ 36;
   x='Athens is t  ';
   y='he Olym   ';
   z='pic site for 2004. ';
   call catt(answer,x,y,z);
   put answer;
run;
```

The following line is written to the SAS log:

```
   ----+----1----+----2----+----3----+----4
   Athens is the Olympic site for 2004.
```

See Also

Functions and CALL Routines:

"CALL CATS Routine" on page 441
"CALL CATX Routine" on page 445
"CAT Function" on page 543
"CATQ Function" on page 546
"CATS Function" on page 550
"CATT Function" on page 552
"CATX Function" on page 554

CALL CATX Routine

Removes leading and trailing blanks, inserts delimiters, and returns a concatenated character string.

Category: Character

Syntax

CALL CATX(*delimiter*, *result*<, *item-1* , … *item-n*>);

Arguments

delimiter
specifies a character string that is used as a delimiter between concatenated strings.

result
specifies a character variable.

Restriction: The CALL CATX routine accepts only a character variable as a valid argument for *result*. Do not use a constant or a SAS expression because CALL CATX is unable to update these arguments.

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string using the BEST*w*. format. In this case, SAS does not write a note to the log.

Details

The CALL CATX routine returns the result in the second argument, *result*. The routine appends the values of the arguments that follow to *result*. If the length of *result* is not large enough to contain the entire result, SAS does the following:

- writes a warning message to the log stating that the result was truncated
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation, except in SQL or in a WHERE clause
- sets _ERROR_ to 1 in the DATA step, except in a WHERE clause

The CALL CATX routine removes leading and trailing blanks from numeric arguments after formatting the numeric value with the BEST*w*. format.

Comparisons

The results of the CALL CATS, CALL CATT, and CALL CATX routines are usually equivalent to statements that use the concatenation operator (||) and the TRIM and LEFT functions. However, using the CALL CATS, CALL CATT, and CALL CATX routines is faster than using TRIM and LEFT.

The following table shows statements that are equivalent to CALL CATS, CALL CATT, and CALL CATX. The variables X1 through X4 specify character variables, and SP specifies a delimiter, such as a blank or comma.

CALL Routine	Equivalent Statement
`CALL CATS(OF X1-X4);`	`X1=TRIM(LEFT(X1))||TRIM(LEFT(X2))||TRIM(LEFT(X3))||TRIM(LEFT(X4));`
`CALL CATT(OF X1-X4);`	`X1=TRIM(X1)||TRIM(X2)||TRIM(X3)||TRIM(X4);`
`CALL CATX(SP, OF X1-X4); *`	`X1=TRIM(LEFT(X1))||SP||TRIM(LEFT(X2))||SP||TRIM(LEFT(X3))||SP||TRIM(LEFT(X4));`

* If any of the arguments are blank, the results that are produced by CALL CATX differ slightly from the results that are produced by the concatenated code. In this case, CALL CATX omits the corresponding delimiter. For example, **`CALL CATX("+",newvar,"X"," ", "Z"," ");`** produces **`X+Z`**.

Examples

The following example shows how the CALL CATX routine concatenates strings.

```
data _null_;
   length answer $ 50;
   separator='%%$%%';
   x='Athens is t  ';
   y='he Olym     ';
   z='  pic site for 2004.  ';
   call catx(separator,answer,x,y,z);
   put answer;
run;
```

The following line is written to the SAS log:

```
----+----1----+----2----+----3----+----4----+----5
Athens is t%%$%%he Olym%%$%%pic site for 2004.
```

See Also

Functions and CALL Routines:

“CALL CATS Routine” on page 441
“CALL CATT Routine” on page 443
“CAT Function” on page 543
“CATQ Function” on page 546
“CATS Function” on page 550
“CATT Function” on page 552
“CATX Function” on page 554

CALL COMPCOST Routine

Sets the costs of operations for later use by the COMPGED function

Category: Character

Restriction: Use with the COMPGED function

Interaction: When invoked by the %SYSCALL macro statement, CALL COMPCOST removes quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL COMPCOST(*operation-1*, *value-1* <,*operation-2*, *value-2* …>);

Arguments

operation
is a character constant, variable, or expression that specifies an operation that is performed by the COMPGED function.

value
is a numeric constant, variable, or expression that specifies the cost of the operation that is indicated by the preceding argument.

Restriction: Must be an integer that ranges from -32767 to 32767, or a missing value

Details

Computing the Cost of Operations Each argument that specifies an operation must have a value that is a character string. The character string corresponds to one of the terms that is used to denote an operation that the COMPGED function performs. See "Computing the Generalized Edit Distance" on page 597 to view a table of operations that the COMPGED function uses.

The character strings that specify operations can be in uppercase, lowercase, or mixed case. Blanks are ignored. Each character string must end with an equal sign (=). Valid values for operations, and the default cost of the operations are listed in the following table.

Operation	Default Cost
APPEND=	**very large**
BLANK=	**very large**
DELETE=	**100**
DOUBLE=	**very large**
FDELETE=	**equal to DELETE**
FINSERT=	**equal to INSERT**
FREPLACE=	**equal to REPLACE**

Operation	Default Cost
INSERT=	**100**
MATCH=	**0**
PUNCTUATION=	**very large**
REPLACE=	**100**
SINGLE=	**very large**
SWAP=	**very large**
TRUNCATE=	**very large**

If an operation does not appear in the call to the COMPCOST routine, or if the operation appears and is followed by a missing value, then that operation is assigned a default cost. A "very large" cost indicates a cost that is sufficiently large that the COMPGED function will not use the corresponding operation.

After your program calls the COMPCOST routine, the costs that are specified remain in effect until your program calls the COMPCOST routine again, or until the step that contains the call to COMPCOST terminates.

Abbreviating Character Strings You can abbreviate character strings. That is, you can use the first one or more letters of a specific operation rather than use the entire term. You must, however, use as many letters as necessary to uniquely identify the term. For example, you can specify the INSERT= operation as "in=", and the REPLACE= operation as "r=". To specify the DELETE= or the DOUBLE= operation, you must use the first two letters because both DELETE= and DOUBLE= begin with "d". The character string must always end with an equal sign.

Examples

The following example calls the COMPCOST routine to compute the generalized edit distance for the operations that are specified.

```
options pageno=1 nodate linesize=80 pagesize=60;

data test;
   length String $8 Operation $40;
   if _n_ = 1 then call compcost('insert=',10,'DEL=',11,'r=', 12);
   input String Operation;
   GED=compged(string, 'baboon');
   datalines;
baboon  match
xbaboon insert
babon   delete
baXoon  replace
;

proc print data=test label;
   label GED='Generalized Edit Distance';
   var String Operation GED;
run;
```

The following output shows the results.

Output 4.12 Generalized Edit Distance Based on Operation

```
                          The SAS System                                1

                                          Generalized
                                              Edit
                Obs    String     Operation    Distance

                 1     baboon      match           0
                 2     xbaboon     insert         10
                 3     babon       delete         11
                 4     baXoon      replace        12
```

See Also

Functions:

"COMPGED Function" on page 596
"COMPARE Function" on page 591
"COMPLEV Function" on page 601

CALL EXECUTE Routine

Resolves the argument, and issues the resolved value for execution at the next step boundary.

Category: Macro

Syntax

CALL EXECUTE(*argument*);

Arguments

argument
specifies a character expression or a constant that yields a macro invocation or a SAS statement. *Argument* can be:

- a character string, enclosed in quotation marks.
- the name of a DATA step character variable. Do not enclose the name of the DATA step variable in quotation marks.
- a character expression that the DATA step resolves to a macro text expression or a SAS statement.

Details

If *argument* resolves to a macro invocation, the macro executes immediately and DATA step execution pauses while the macro executes. If *argument* resolves to a SAS

statement or if execution of the macro generates SAS statements, the statement(s) execute after the end of the DATA step that contains the CALL EXECUTE routine. CALL EXECUTE is fully documented in *SAS Macro Language: Reference*.

CALL GRAYCODE Routine

Generates all subsets of *n* items in a minimal change order.

Category: Combinatorial

Syntax

CALL GRAYCODE(*k*, *numeric-variable-1*, ..., *numeric-variable-n*);

CALL GRAYCODE(*k*, *character-variable* <, *n* <, *in-out*>>);

Arguments

k
specifies a numeric variable. Initialize *k* to either of the following values before executing the CALL GRAYCODE routine:

- a negative number to cause CALL GRAYCODE to initialize the subset to be empty
- the number of items in the initial set indicated by *numeric-variable-1* through *numeric-variable-n*, or *character-variable*, which must be an integer value between 0 and N inclusive

The value of *k* is updated when CALL GRAYCODE is executed. The value that is returned is the number of items in the subset.

numeric-variable
specifies numeric variables that have values of 0 or 1 which are updated when CALL GRAYCODE is executed. A value of 1 for *numeric-variable-j* indicates that the j^{th} item is in the subset. A value of 0 for *numeric-variable-j* indicates that the j^{th} item is not in the subset.

If you assign a negative value to *k* before you execute CALL GRAYCODE, then you do not need to initialize *numeric-variable-1* through *numeric-variable-n* before executing CALL GRAYCODE unless you want to suppress the note about uninitialized variables.

If you assign a value between 0 and *n* inclusive to *k* before you execute CALL GRAYCODE, then you must initialize *numeric-variable-1* through *numeric-variable-n* to *k* values of 1 and *n-k* values of 0.

character-variable
specifies a character variable that has a length of at least *n* characters. The first *n* characters indicate which items are in the subset. By default, an "I" in the j^{th} position indicates that thej^{th} item is in the subset, and an "O" in the j^{th} position indicates that the j^{th} item is out of the subset. You can change the two characters by specifying the *in-out* argument.

If you assign a negative value to *k* before you execute CALL GRAYCODE, then you do not need to initialize *character-variable* before executing CALL GRAYCODE unless you want to suppress the note about an uninitialized variable.

If you assign a value between 0 and *n* inclusive to *k* before you execute CALL GRAYCODE, then you must initialize *character-variable* to *k* characters that indicate an item is in the subset, and *k*-*k* characters that indicate an item is out of the subset.

n
specifies a numeric constant, variable, or expression. By default, *n* is the length of *character-variable*.

in-out
specifies a character constant, variable, or expression. The default value is "IO." The first character is used to indicate that an item is in the subset. The second character is used to indicate that an item is out of the subset.

Details

Using CALL GRAYCODE in a DATA Step When you execute the CALL GRAYCODE routine with a negative value of *k*, the subset is initialized to be empty.

When you execute the CALL GRAYCODE routine with an integer value of *k* between 0 and *n* inclusive, one item is either added to the subset or removed from the subset, and the value of *k* is updated to equal the number of items in the subset.

To generate all subsets of *n* items, you can initialize *k* to a negative value and execute CALL GRAYCODE in a loop that iterates 2**n* times. If you want to start with a non-empty subset, then initialize *k* to be the number of items in the subset, initialize the other arguments to specify the desired initial subset, and execute CALL GRAYCODE in a loop that iterates 2**n*-1 times. The sequence of subsets that are generated by CALL GRAYCODE is cyclical, so you can begin with any subset you want.

Using the CALL GRAYCODE Routine with Macros You can call the GRAYCODE routine when you use the %SYSCALL macro. Differences exist when you use CALL GRAYCODE in a DATA step and when you use the routine with macros. The following list describes usage with macros:

- All arguments must be initialized to nonblank values.
- If you use the *character-variable* argument, then it must be initialized to a nonblank, nonnumeric character string that contains at least *n* characters.
- If you use the *in-out* argument, then it must be initialized to a string that contains two characters that are not blanks, digits, decimal points, or plus and minus signs.

If %SYSCALL identifies an argument as being the wrong type, or if %SYSCALL is unable to identify the type of argument, then &SYSERR and &SYSINFO are *not* set.

Otherwise, if an error occurs during the execution of the CALL GRAYCODE routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than –100.

If there are no errors, then &SYSERR is set to zero, and &SYSINFO is set to one of the following values:

- 0 if the value of *k* on input is negative
- the index of the item that was added or removed from the subset if the value of *k* on input is a valid nonnegative integer.

Examples

Example 1: Using a Character Variable and Positive Initial *k* with CALL GRAYCODE The following example uses the CALL GRAYCODE routine to generate subsets in a minimal change order.

```
data _null_;
   x='++++';
   n=length(x);
   k=countc(x, '+');
   put '    1' +3 k= +2 x=;
   nsubs=2**n;
   do i=2 to nsubs;
      call graycode(k, x, n, '+-');
      put i 5. +3 k= +2 x=;
   end;
run;
```

SAS writes the following output to the log:

```
    1   k=4   x=++++
    2   k=3   x=-+++
    3   k=2   x=-+-+
    4   k=3   x=++-+
    5   k=2   x=+--+
    6   k=1   x=---+
    7   k=0   x=----
    8   k=1   x=+---
    9   k=2   x=++--
   10   k=1   x=-+--
   11   k=2   x=-++-
   12   k=3   x=+++-
   13   k=2   x=+-+-
   14   k=1   x=--+-
   15   k=2   x=--++
   16   k=3   x=+-++
```

Example 2: Using %SYSCALL with Numeric Variables and Negative *k* The following example uses the %SYSCALL macro with numeric variables to generate subsets in a minimal change order.

```
%macro test;
   %let n=3;
   %let x1=.;
   %let x2=.;
   %let x3=.;
   %let k=-1;
   %let nsubs=%eval(2**&n + 1);
   %put nsubs=&nsubs k=&k x: &x1 &x2 &x3;
   %do j=1 %to &nsubs;
      %syscall graycode(k, x1, x2, x3);
      %put &j: k=&k x: &x1 &x2 &x3 sysinfo=&sysinfo;
   %end;
%mend;

%test;
```

SAS writes the following output to the log:

```
nsubs=9 k=-1 x: . . .
1: k=0 x: 0 0 0 sysinfo=0
2: k=1 x: 1 0 0 sysinfo=1
3: k=2 x: 1 1 0 sysinfo=2
4: k=1 x: 0 1 0 sysinfo=1
5: k=2 x: 0 1 1 sysinfo=3
6: k=3 x: 1 1 1 sysinfo=1
7: k=2 x: 1 0 1 sysinfo=2
8: k=1 x: 0 0 1 sysinfo=1
9: k=0 x: 0 0 0 sysinfo=3
```

Example 3: Using %SYSCALL with a Character Variable and Negative *k* The following example uses the %SYSCALL macro with a character variable to generate subsets in a minimal change order.

```
%macro test(n);
   %*** Initialize the character variable to a
        sufficiently long nonblank, nonnumeric value. ;
   %let x=%sysfunc(repeat(_, &n-1));
   %let k=-1;
   %let nsubs=%eval(2**&n + 1);
   %put nsubs=&nsubs k=&k x="&x";
   %do j=1 %to &nsubs;
      %syscall graycode(k, x, n);
      %put &j: k=&k x="&x" sysinfo=&sysinfo;
   %end;
%mend;

%test(3);
```

SAS writes the following output to the log:

```
nsubs=9 k=-1 x="___"
1: k=0 x="OOO" sysinfo=0
2: k=1 x="IOO" sysinfo=1
3: k=2 x="IIO" sysinfo=2
4: k=1 x="OIO" sysinfo=1
5: k=2 x="OII" sysinfo=3
6: k=3 x="III" sysinfo=1
7: k=2 x="IOI" sysinfo=2
8: k=1 x="OOI" sysinfo=1
9: k=0 x="OOO" sysinfo=3
```

See Also

Functions:

"GRAYCODE Function" on page 797

CALL IS8601_CONVERT Routine

Converts an ISO 8601 interval to datetime and duration values, or converts datetime and duration values to an ISO 8601 interval.

Category: Date and Time

Syntax

CALL IS8601_CONVERT(*convert-from*, *convert-to*, <*from-variables*>, <*to-variables*>, <*date_time_replacements*>

Arguments

convert-from
specifies a keyword in single quotation marks that indicates whether the source for the conversion is an interval, a datetime and duration value, or a duration value. *convert-from* can have one of the following values:

'intvl'	specifies that the source value for the conversion is an interval value.
'dt/du'	specifies that the source value for the conversion is a datetime/duration value.
'du/dt'	specifies that the source value for the conversion is a duration/datetime value.
'dt/dt'	specifies that the source value for the conversion is a datetime/datetime value.
'du'	specifies that the source value for the conversion is a duration value.

convert-to
specifies a keyword in single quotation marks that indicates the results of the conversion. *convert-to* can have one of the following values:

'intvl'	specifies to create an interval value.
'dt'/du'	specifies to create a datetime/duration interval.
'du/dt'	specifies to create a duration/datetime interval.
'dt/dt'	specifies to create a datetime/datetime interval.
'du'	specifies to create a duration.
'start'	specifies to create a value that is the beginning datetime or duration of an interval value.
'end'	specifies to create a value that is the ending datetime or duration of an interval value.

from-variable
specifies one or two variables that contain the source value. Specify one variable for an interval value and two variables, one each, for datetime and duration values. The datetime and duration values are interval components where the first value is the

beginning value of the interval and the second value is the ending value of the interval.

Requirement: An integer variable must be at least a 16-byte character variable whose value is determined by reading the value using either the $N8601B informat or the $N8601E informat, or the integer variable is an integer value returned from invoking the CALL ISO8601_CONVERT routine.

Requirement: A datetime value must be either a SAS datetime value or an 8-byte character value that is read by the $N8601B informat or the $N8601E informat, or by invoking the CALL ISO8601_CONVERT routine.

Requirement: A duration value must be a numeric value that represents the number of seconds in the duration or an 8–byte character value whose value is determined by reading the value using either the $N8601B informat or the $N8601E informat, or by invoking the CALL ISO8601_CONVERT routine.

to-variable

specifies one or two variables that contain converted values. Specify one variable for in interval value and two variables, one each, for datetime and duration values.

Requirement: The interval variable must be at least a 16-byte character variable.

Tip: The datetime and duration variables can be either numeric or character. To avoid losing precision of a numeric value, the length of a numeric variable needs to be at least eight characters. Datetime and duration character variables must be at least 16 bytes; they are padded with blank characters for values that are less than the length of the variable.

date_time_replacements

specifies date or time component values to use when a month, day, or time component is omitted from an interval, datetime, or duration value. *date_time_replacements* is specified as a series of numbers separated by a comma to represent, in this order, the year, month, day, hour, minute, or second. Components of *date_time_replacements* can be omitted only in the reverse order, seconds, minutes, hours, day, and month. If no substitute values are specified, the conversion is done using default values.

Defaults: The following are default values for omitted date and time components:

month	1
day	1
hour	0
minute	0
second	0

Requirements: A year component must be part of the datetime or duration value, and therefore is not valid in *date_time_replacements*. A comma is required as a placeholder for the year in *date_time_replacements*. For example, in the replacement value string, **`,9,4,,2,'`**, the first comma is a placeholder for a year value.

Examples

This DATA step uses the ISO8601_CONVERT function to do the following:

- create an interval by using datetime and duration values
- create datetime and duration values from an interval that was created using the CALL IS8601_CONVERT routine
- create an interval from datetime and duration values, using replacement values for omitted date and time components in the datetime value

For easier reading, numeric variables end with an N and character variables end with a C.

```
data _null_;

   /**  declare variable length and type                          **/
   /**  Character datetime and duration values must be at least  **/
   /**  16 characters. In order not to lose precision, the      **/
   /**  numeric datetime value has a length of 8.                 **/

   length dtN duN 8 dtC duC $16 intervalC $32;

   /** assign a numeric datetime value and a      **/
   /** character duration value.                  **/

   dtN='15Sep2008:09:00:00'dt;
   duC=input('P2y3m4dT5h6m7s', $n8601b.);
   put dtN=;
   put duC=;

   /** Create an interval from a datetime and duration value   **/
   /** and format it using the ISO 8601 extended notation for  **/
   /** character values.                                       **/

   call is8601_convert('dt/du', 'intvl', dtN, duC, intervalC);

   put '**  Character interval created from datetime and duration values **/';
   put intervalC $n8601e.;
   put ' ';

   /** Create numeric datetime and duration values from an interval    **/
   /** and format it using the ISO 8601 extended notation for          **/
   /** numeric values.                                                 **/

   call is8601_convert('intvl', 'dt/du', intervalC, dtN, duN);

   put '**  Character datetime and duration created from an interval  **/';
   put dtN=;
   put duN=;
   put ' ';

   /** assign a new datetime value with omitted components              **/

   dtC=input('2009---15T10:-:-', $n8601b.);

   put '** This datetime is a character value. **';
   put dtC $n8601h.;
   put ' ';

   /** Create an interval by reading in a datetime value       **/
   /** with omitted date and time components.  Use replacement **/
   /** values for the month, minutes, and seconds.             **/


   call is8601_convert('du/dt', 'intvl', duC, dtC, intervalC,,7,,,35,45);
```

```
   put '**  Interval created using a datetime with omitted values,      **';
   put '**  inserting replacement values for month (7), minute (35)     **';
   put '**  seconds (45).                                               **';
   put intervalC $n8601e.;
   put ' ';

run;
```

The following output appears in the SAS log:

```
dtN=1537088400
duC=0002304050607FFC
**  Character interval created from datetime and duration values **/
2008-09-15T09:00:00.000/P2Y3M4DT5H6M7S

**  Character datetime and duration created from an interval  **/
dtN=1537088400
duN=71211967

** This datetime is a character value. **
2009---15T10:-:-

**  Interval created using a datetime with omitted values,      **
**  inserting replacement values for month (7), minute (35)     **
**  seconds (45).                                               **

P2Y3M4DT5H6M7S/2009-07-15T10:35:45
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.03 seconds
```

CALL LABEL Routine

Assigns a variable label to a specified character variable.

Category: Variable Control

Syntax

CALL LABEL(*variable-1*,*variable-2*);

Arguments

variable-1
specifies any SAS variable. If *variable-1* does not have a label, the variable name is assigned as the value of *variable-2*.

variable-2
specifies any SAS character variable. Variable labels can be up to 256 characters long. Therefore, the length of *variable-2* should be at least 256 characters to avoid truncating variable labels.

Note: To conserve space, you should set the length of *variable-2* to the length of the label for *variable-1*, if it is known. △

Details

The CALL LABEL routine assigns the label of the *variable-1* variable to the character variable *variable-2*.

Examples

This example uses the CALL LABEL routine with array references to assign the labels of all variables in the data set OLD as values of the variable LAB in data set NEW:

```
data new;
   set old;
      /* lab is not in either array */
   length lab $256;
      /* all character variables in old */
   array abc{*} _character_;
      /* all numeric variables in old */
   array def{*} _numeric_;
   do i=1 to dim(abc);
         /* get label of character variable */
      call label(abc{i},lab);
         /* write label to an observation */
      output;
   end;
   do j=1 to dim(def);
         /* get label of numeric variable */
      call label(def{j},lab);
         /* write label to an observation */
      output;
   end;
   stop;
   keep lab;
run;
```

See Also

Function:

"VLABEL Function" on page 1214

CALL LEXCOMB Routine

Generates all distinct combinations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.

Category: Combinatorial

Interaction: When invoked by the %SYSCALL macro statement, CALL LEXCOMB removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL LEXCOMB(*count*, *k*, *variable-1*, …, *variable-n*);

Arguments

count
specifies an integer value that is assigned values from 1 to the number of combinations in a loop.

k
specifies an integer constant, variable, or expression between 1 and *n*, inclusive, that specifies the number of items in each combination.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before you call the LEXCOMB routine.

Tip: After calling LEXCOMB, the first *k* variables contain the values in one combination.

Details

The Basics Use the CALL LEXCOMB routine in a loop where the first argument to CALL LEXCOMB takes each integral value from 1 to the number of distinct combinations of the non-missing values of the variables. In each call to LEXCOMB within this loop, *k* should have the same value.

Number of Combinations When all of the variables have non-missing, unequal values, then the number of combinations is COMB(*n*,*k*). If the number of variables that have missing values is *m*, and all the non-missing values are unequal, then LEXCOMB produces COMB(*n-m*,*k*) combinations because the missing values are omitted from the combinations.

When some of the variables have equal values, the exact number of combinations is difficult to compute. If you cannot compute the exact number of combinations, use the LEXCOMB function instead of the CALL LEXCOMB routine.

CALL LEXCOMB Processing On the first call to the LEXCOMB routine, the following actions occur:

- The argument types and lengths are checked for consistency.
- The *m* missing values are assigned to the last *m* arguments.
- The *n-m* non-missing values are assigned in ascending order to the first *n-m* arguments following *count*.

On subsequent calls, up to and including the last combination, the next distinct combination of the non-missing values is generated in lexicographic order.

If you call the LEXCOMB routine with the first argument out of sequence, then the results are not useful. In particular, if you initialize the variables and then immediately call the LEXCOMB routine with a first argument of *j*, you will not get the j^{th} combination (except when *j* is 1). To get the j^{th} combination, you must call the LEXCOMB routine *j* times, with the first argument taking values from 1 through *j* in that exact order.

Using the CALL LEXCOMB Routine with Macros You can call the LEXCOMB routine when you use the %SYSCALL macro. In this case, the *variable* arguments are not

required to be the same length, but they are required to be the same type. If %SYSCALL identifies an argument as numeric, then %SYSCALL reformats the returned value.

If an error occurs during the execution of the CALL LEXCOMB routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than –100.

If there are no errors, then &SYSERR is set to zero, and &SYSINFO is set to one of the following values:

- 1 if *count*=1 and at lease one variable has a non-missing value
- 1 if the value of *variable-1* changed
- *j* if *variable-1* through *variable-i* did not change, but *variable-j* did change, where *j*=*i*+1
- –1 if all distinct combinations have already been generated

Comparisons

The CALL LEXCOMB routine generates all distinct combinations of the non-missing values of *n* variables taken *k* at a time in lexicographic order. The CALL ALLCOMB routine generates all combinations of the values of *n* variables taken *k* at a time in a minimal change order.

Examples

Example 1: Using CALL LEXCOMB in a DATA Step The following example calls the LEXCOMB routine to generate distinct combinations in lexicographic order.

```
data _null_;
   array x[5] $3 ('ant' 'bee' 'cat' 'dog' 'ewe');
   n=dim(x);
   k=3;
   ncomb=comb(n,k);
   do j=1 to ncomb;
      call lexcomb(j, k, of x[*]);
      put j 5. +3 x1-x3;
   end;
run;
```

SAS writes the following output to the log:

```
    1   ant bee cat
    2   ant bee dog
    3   ant bee ewe
    4   ant cat dog
    5   ant cat ewe
    6   ant dog ewe
    7   bee cat dog
    8   bee cat ewe
    9   bee dog ewe
   10   cat dog ewe
```

Example 2: Using CALL LEXCOMB with Macros The following is an example of the CALL LEXCOMB routine that is used with macros. The output includes values for the %SYSINFO macro.

```
%macro test;
   %let x1=ant;
   %let x2=baboon;
   %let x3=baboon;
   %let x4=hippopotamus;
   %let x5=zebra;
   %let k=2;
   %let ncomb=%sysfunc(comb(5,&k));
      %do j=1 %to &ncomb;
         %syscall lexcomb(j, k, x1, x2, x3, x4, x5);
         %let jfmt=%qsysfunc(putn(&j, 5. ));
         %let pad=%qsysfunc(repeat(%str( ), 20-%length(&x1 &x2)));
         %put &jfmt: &x1 &x2 &pad sysinfo=&sysinfo;
         %if &sysinfo < 0 %then %let j=%eval(&ncomb+1);
      %end;
%mend;

%test
```

SAS writes the following output to the log:

```
1: ant baboon             sysinfo=1
2: ant hippopotamus       sysinfo=2
3: ant zebra              sysinfo=2
4: baboon baboon          sysinfo=1
5: baboon hippopotamus    sysinfo=2
6: baboon zebra           sysinfo=2
7: hippopotamus zebra     sysinfo=1
8: hippopotamus zebra     sysinfo=-1
```

See Also

Functions and CALL Routines:

"LEXCOMB Function" on page 889

"CALL ALLCOMB Routine" on page 432

CALL LEXCOMBI Routine

Generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order.

Category: Combinatorial

Syntax

CALL LEXCOMBI(*n*, *k*, *index-1*, ..., *index-k*);

Arguments

n
is a numeric constant, variable, or expression that specifies the total number of objects.

k
is a numeric constant, variable, or expression that specifies the number of objects in each combination.

index
is a numeric variable that contains indices of the objects in the combination that is returned. Indices are integers between 1 and *n*, inclusive.

Tip: If *index-1* is missing or zero, then the CALL LEXCOMBI routine initializes the indices to **`index-1=1`** through **`index-k=k`**. Otherwise, CALL LEXCOMBI creates a new combination by removing one index from the combination and adding another index.

Details

CALL LEXCOMBI Processing Before the first call to the LEXCOMBI routine, complete one of the following tasks:

- Set *index-1* equal to zero or to a missing value.
- Initialize *index-1* through *index-k* to distinct integers between 1 and *n* inclusive.

The number of combinations of *n* objects taken *k* at a time can be computed as COMB(*n*,*k*). To generate all combinations of *n* objects taken *k* at a time, call LEXCOMBI in a loop that executes COMB(*n*,*k*) times.

Using the CALL LEXCOMBI Routine with Macros If you call the LEXCOMBI routine from the macro processor with %SYSCALL, then you must initialize all arguments to numeric values. %SYSCALL reformats the values that are returned.

If an error occurs during the execution of the CALL LEXCOMBI routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than –100.

If there are no errors, then &SYSERR is set to zero, and &SYSINFO is set to one of the following values:

- 1 if the value of *variable-1* changed
- *j* if *variable-1* through *variable-i* did not change, but *variable-j* did change, where *j*=*i*+1
- –1 if all distinct combinations have already been generated

Comparisons

The CALL LEXCOMBI routine generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order. The CALL ALLCOMBI routine generates all combinations of the indices of *n* objects taken *k* at a time in a minimum change order.

Examples

Example 1: Using the CALL LEXCOMBI Routine with the DATA Step The following example uses the CALL LEXCOMBI routine to generate combinations of indices in lexicographic order.

```
data _null_;
   array x[5] $3 ('ant' 'bee' 'cat' 'dog' 'ewe');
   array c[3] $3;
   array i[3];
   n=dim(x);
   k=dim(i);
   i[1]=0;
   ncomb=comb(n,k);
   do j=1 to ncomb;
      call lexcombi(n, k, of i[*]);
      do h=1 to k;
         c[h]=x[i[h]];
      end;
      put @4 j= @10 'i= ' i[*] +3 'c= ' c[*];
   end;
run;
```

SAS writes the following output to the log:

```
j=1   i= 1 2 3    c= ant bee cat
j=2   i= 1 2 4    c= ant bee dog
j=3   i= 1 2 5    c= ant bee ewe
j=4   i= 1 3 4    c= ant cat dog
j=5   i= 1 3 5    c= ant cat ewe
j=6   i= 1 4 5    c= ant dog ewe
j=7   i= 2 3 4    c= bee cat dog
j=8   i= 2 3 5    c= bee cat ewe
j=9   i= 2 4 5    c= bee dog ewe
j=10  i= 3 4 5    c= cat dog ewe
```

Example 2: Using the CALL LEXCOMBI Routine with Macros and Displaying the Return Code The following example uses the CALL LEXCOMBI routine with macros. The output includes values for the %SYSINFO macro.

```
%macro test;
   %let x1=0;
   %let x2=0;
   %let x3=0;
   %let n=5;
   %let k=3;
   %let ncomb=%sysfunc(comb(&n,&k));
   %do j=1 %to &ncomb+1;
      %syscall lexcombi(n,k,x1,x2,x3);
      %let jfmt=%qsysfunc(putn(&j,5.));
      %let pad=%qsysfunc(repeat(%str(),6-%length(&x1 &x2 &x3)));
      %put &jfmt: &x1 &x2 &x3 &pad sysinfo=&sysinfo;
   %end;
%mend;

%test
```

SAS writes the following output to the log:

```
    1: 1 2 3  sysinfo=1
    2: 1 2 4  sysinfo=3
    3: 1 2 5  sysinfo=3
    4: 1 3 4  sysinfo=2
    5: 1 3 5  sysinfo=3
    6: 1 4 5  sysinfo=2
    7: 2 3 4  sysinfo=1
    8: 2 3 5  sysinfo=3
    9: 2 4 5  sysinfo=2
   10: 3 4 5  sysinfo=1
   11: 3 4 5  sysinfo=-1
```

See Also

Functions and CALL Routines:

“CALL LEXCOMB Routine” on page 458

“CALL ALLCOMBI Routine” on page 434

CALL LEXPERK Routine

Generates all distinct permutations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.

Category: Combinatorial

Interaction: When invoked by THE %SYSCALL macro statement, CALL LEXPERK removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL LEXPERK(*count*, *k*, *variable-1*, ..., *variable-n*);

Arguments

count
specifies an integer variable that is assigned a value from 1 to the number of permutations in a loop.

k
specifies an integer constant, variable, or expression between 1 and *n*, inclusive, that specifies the number of items in each permutation.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before you call the LEXPERK routine.

Tip: After calling LEXPERK, the first *k* variables contain the values in one permutation.

Details

The Basics Use the CALL LEXPERK routine in a loop where the first argument to CALL LEXPERK accepts each integral value from 1 to the number of distinct permutations of *k* non-missing values of the variables. In each call to LEXPERK within this loop, *k* should have the same value.

Number of Permutations When all of the variables have non-missing, unequal values, the number of permutations is PERM(,*k*). If the number of variables that have missing values is *m*, and all the non-missing values are unequal, CALL LEXPERK produces PERM(*n-m*,*k*) permutations because the missing values are omitted from the permutations. When some of the variables have equal values, the exact number of permutations is difficult to compute. If you cannot compute the exact number of permutations, use the LEXPERK function instead of the CALL LEXPERK routine.

CALL LEXPERK Processing On the first call to the LEXPERK routine, the following actions occur:

- The argument types and lengths are checked for consistency.
- The *m* missing values are assigned to the last *m* arguments.
- The *n-m* non-missing values are assigned in ascending order to the first *n-m* arguments following *count*.

On subsequent calls, up to and including the last permutation, the next distinct permutation of *k* non-missing values is generated in lexicographic order.

If you call the LEXPERK routine with the first argument out of sequence, then the results are not useful. In particular, if you initialize the variables and then immediately call the LEXPERK routine with a first argument of *j*, you will not get the j^{th} permutation (except when *j* is 1). To get the j^{th} permutation, you must call LEXPERK *j* times, with the first argument taking values from 1 through *j* in that exact order.

Using the CALL LEXPERK Routine with Macros You can call the LEXPERK routine when you use the %SYSCALL macro. In this case, the *variable* arguments are not required to be the same length, but they are required to be the same type. If %SYSCALL identifies an argument as numeric, then %SYSCALL reformats the returned value.

If an error occurs during the execution of the CALL LEXPERK routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than –100.

If there are no errors, then &SYSERR is set to zero, and &SYSINFO is set to one of the following values:

- 1 if *count*=1 and at least one variable has a non-missing value
- 1 if *count*>1 and the value of *variable-1* changed
- *j* if *count*>1 and *variable-1* through *variable-i* did not change, but *variable-j* did change, where *j*=*i*+1
- –1 if all distinct permutations were already generated

Comparisons

The CALL LEXPERK routine generates all distinct permutations of the non-missing values of *n* variables taken *k* at a time in lexicographic order. The CALL ALLPERM routine generates all permutations of the values of several variables in a minimal change order.

Examples

Example 1: Using CALL LEXPERK in a DATA Step The following is an example of the CALL LEXPERK routine.

```
data _null_;
   array x[5] $3 ('V' 'W' 'X' 'Y' 'Z');
```

```
   n=dim(x);
   k=3;
   nperm=perm(n,k);
   do j=1 to nperm;
      call lexperk(j, k, of x[*]);
      put j 5. +3 x1-x3;
   end;
run;
```

SAS writes the following output to the log:

```
    1   V W X
    2   V W Y
    3   V W Z
    4   V X W
    5   V X Y
    6   V X Z
    7   V Y W
    8   V Y X
    9   V Y Z
   10   V Z W
   11   V Z X
   12   V Z Y
   13   W V X
   14   W V Y
   15   W V Z
   16   W X V
   17   W X Y
   18   W X Z
   19   W Y V
   20   W Y X
   21   W Y Z
   22   W Z V
   23   W Z X
   24   W Z Y
   25   X V W
   26   X V Y
   27   X V Z
   28   X W V
   29   X W Y
   30   X W Z
   31   X Y V
   32   X Y W
   33   X Y Z
   34   X Z V
   35   X Z W
   36   X Z Y
   37   Y V W
   38   Y V X
   39   Y V Z
   40   Y W V
   41   Y W X
   42   Y W Z
   43   Y X V
   44   Y X W
```

```
45   Y X Z
46   Y Z V
47   Y Z W
48   Y Z X
49   Z V W
50   Z V X
51   Z V Y
52   Z W V
53   Z W X
54   Z W Y
55   Z X V
56   Z X W
57   Z X Y
58   Z Y V
59   Z Y W
60   Z Y X
```

Example 2: Using CALL LEXPERK with Macros The following is an example of the CALL LEXPERK routine that is used with macros. The output includes values for the %SYSINFO macro.

```
%macro test;
   %let x1=ant;
   %let x2=baboon;
   %let x3=baboon;
   %let x4=hippopotamus;
   %let x5=zebra;
   %let k=2;
   %let nperk=%sysfunc(perm(5,&k));
   %do j=1 %to &nperk;
      %syscall lexperk(j, k, x1, x2, x3, x4, x5);
      %let jfmt=%qsysfunc(putn(&j,5.));
      %let pad=%qsysfunc(repeat(%str(),20-%length(&x1 &x2)));
      %put &jfmt: &x1 &x2 &pad sysinfo=&sysinfo;
      %if &sysinfo<0 %then %let j=%eval(&nperk+1);
   %end;
%mend;

%test
```

SAS writes the following output to the log:

```
 1: ant baboon  sysinfo=1
 2: ant hippopotamus  sysinfo=2
 3: ant zebra  sysinfo=2
 4: baboon ant  sysinfo=1
 5: baboon baboon  sysinfo=2
 6: baboon hippopotamus  sysinfo=2
 7: baboon zebra  sysinfo=2
 8: hippopotamus ant  sysinfo=1
 9: hippopotamus baboon  sysinfo=2
10: hippopotamus zebra  sysinfo=2
11: zebra ant  sysinfo=1
12: zebra baboon  sysinfo=2
13: zebra hippopotamus  sysinfo=2
14: zebra hippopotamus  sysinfo=-1
```

See Also

Functions and CALL Routines:

"CALL ALLPERM Routine" on page 437
"LEXPERM Function" on page 896
"CALL RANPERK Routine" on page 503
"CALL RANPERM Routine" on page 505

CALL LEXPERM Routine

Generates all distinct permutations of the non-missing values of several variables in lexicographic order.

Category: Combinatorial

Interaction: When invoked by the %SYSCALL macro statement, CALL LEXPERM removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL LEXPERM(*count*, *variable-1* <, …, *variable-N*>);

Arguments

count
specifies a numeric variable that has an integer value that ranges from 1 to the number of permutations.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted by LEXPERM.

Requirement: Initialize these variables before you call the LEXPERM routine.

Details

Determine the Number of Distinct Permutations These variables are defined for use in the equation that follows:

N	specifies the number of variables that are being permuted—that is, the number of arguments minus one.
M	specifies the number of missing values among the variables that are being permuted.
d	specifies the number of distinct non-missing values among the arguments.
N_i	for i=1 through i=d, N_i specifies the number of instances of the *ith* distinct value.

The number of distinct permutations of non-missing values of the arguments is expressed as follows:

$$P = \frac{(N_1 + N_2 + ... + N_d)!}{N_1!N_2!...N_d!} <= N!$$

CALL LEXPERM Processing Use the CALL LEXPERM routine in a loop where the argument *count* accepts each integral value from 1 to P. You do not need to compute P provided you exit the loop when CALL LEXPERM returns a value that is less than zero.

For 1=*count*<P, the following actions occur:

- The argument types and lengths are checked for consistency.
- The M missing values are assigned to the last M arguments.
- The N-M non-missing values are assigned in ascending order to the first N-M arguments following *count*.
- CALL LEXPERM returns 1.

For 1<*count*<=P, the following actions occur:

- The next distinct permutation of the non-missing values is generated in lexicographic order.
- If *variable-1* through *variable-I* did not change, but *variable-J* did change, where J=I+1, then CALL LEXPERM returns J.

For *count*>P, CALL LEXPERM returns –1.

If the CALL LEXPERM routine is executed with the first argument out of sequence, the results might not be useful. In particular, if you initialize the variables and then immediately execute CALL LEXPERM with a first argument of K, you will not get the K*th* permutation (except when K is 1). To get the K*th* permutation, you must execute CALL LEXPERM K times, with the first argument accepting values from 1 through K in that exact order.

Using the CALL LEXPERM Routine with Macros You can call the LEXPERM routine when you use the %SYSCALL macro. In this case, the *variable* arguments are not required to be the same length, but they must be the same type. If %SYSCALL identifies an argument as numeric, then %SYSCALL reformats the returned value.

If an error occurs during the execution of the CALL LEXPERM routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than –100.

If there are no errors, then &SYSERR is set to zero, and &SYSINFO is set to one of the following values:

- 1 if 1=*count*<P
- 1 if 1<*count*<=P and the value of *variable-1* changed
- J if 1<*count*<=P and *variable-1* through *variable-I* did not change, but *variable-J* did change, where J=I+1
- –1 if *count*>P

Comparisons

SAS provides three functions or CALL routines for generating all permutations:

- ALLPERM generates all *possible* permutations of the values, *missing or non-missing*, of several variables. Each permutation is formed from the previous permutation by interchanging two consecutive values.

- LEXPERM generates all *distinct* permutations of the *non-missing* values of several variables. The permutations are generated in lexicographic order.
- LEXPERK generates all *distinct* permutations of K of the *non-missing* values of N variables. The permutations are generated in lexicographic order.

ALLPERM is the fastest of these functions and CALL routines. LEXPERK is the slowest.

Examples

Example 1: Using CALL LEXPERM in a DATA Step The following example uses the DATA step to generate all distinct permutations of the non-missing values of several variables in lexicographic order.

```
data _null_;
   array x[4] $3 ('ant' 'bee' 'cat' 'dog');
   n=dim(x);
   nfact=fact(n);
   do i=1 to nfact;
      call lexperm(i, of x[*]);
      put i 5. +2 x[*];
   end;
run;
```

SAS writes the following output to the log:

```
    1  ant bee cat dog
    2  ant bee dog cat
    3  ant cat bee dog
    4  ant cat dog bee
    5  ant dog bee cat
    6  ant dog cat bee
    7  bee ant cat dog
    8  bee ant dog cat
    9  bee cat ant dog
   10  bee cat dog ant
   11  bee dog ant cat
   12  bee dog cat ant
   13  cat ant bee dog
   14  cat ant dog bee
   15  cat bee ant dog
   16  cat bee dog ant
   17  cat dog ant bee
   18  cat dog bee ant
   19  dog ant bee cat
   20  dog ant cat bee
   21  dog bee ant cat
   22  dog bee cat ant
   23  dog cat ant bee
   24  dog cat bee ant
```

Example 2: Using CALL LEXPERM with Macros The following is an example of the CALL LEXPERM routine that is used with macros. The output includes values for the %SYSINFO macro.

```
%macro test;
   %let x1=ant;
   %let x2=baboon;
   %let x3=baboon;
   %let x4=hippopotamus;
   %let n=4;
   %let nperm=%sysfunc(perm(4));
   %do j=1 %to &nperm;
      %syscall lexperm(j,x1,x2,x3,x4);
      %let jfmt=%qsysfunc(putn(&j,5.));
      %put &jfmt: &x1 &x2 &x3 &x4 sysinfo=&sysinfo;
      %if &sysinfo<0 %then %let j=%eval(&nperm+1);
   %end;
 %mend;

 %test;
```

SAS writes the following output to the log:

```
    1: ant baboon baboon hippopotamus sysinfo=1
    2: ant baboon hippopotamus baboon sysinfo=3
    3: ant hippopotamus baboon baboon sysinfo=2
    4: baboon ant baboon hippopotamus sysinfo=1
    5: baboon ant hippopotamus baboon sysinfo=3
    6: baboon baboon ant hippopotamus sysinfo=2
    7: baboon baboon hippopotamus ant sysinfo=3
    8: baboon hippopotamus ant baboon sysinfo=2
    9: baboon hippopotamus baboon ant sysinfo=3
   10: hippopotamus ant baboon baboon sysinfo=1
   11: hippopotamus baboon ant baboon sysinfo=2
   12: hippopotamus baboon baboon ant sysinfo=3
   13: hippopotamus baboon baboon ant sysinfo=-1
```

See Also

Functions and CALL Routines:

“LEXPERM Function” on page 896
“CALL ALLPERM Routine” on page 437
“LEXPERK Function” on page 894
“CALL RANPERK Routine” on page 503
“CALL RANPERM Routine” on page 505

CALL LOGISTIC Routine

Applies the logistic function to each argument.

Category: Mathematical

Syntax

CALL LOGISTIC(*argument*<, *argument, ...*>)

Arguments

argument
is a numeric variable.

Restriction The CALL LOGISTIC routine only accepts variables as valid arguments. Do not use a constant or a SAS expression because the CALL routine is unable to update these arguments.

Details

The CALL LOGISTIC routine replaces each argument by the logistic value of that argument. For example x_j is replaced by

$$\frac{e^{x_j}}{1+e^{x_j}}$$

If any argument contains a missing value, then CALL LOGISTIC returns missing values for all the arguments.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`x=0.5;`
`y=-0.5;`
`call logistic(x,y);`
`put x= y=;`
	`x=0.6224593312 y=0.3775406688`

CALL MISSING Routine

Assigns missing values to the specified character or numeric variables.

Category: Character

Syntax

CALL MISSING(*varname1*<, *varname2*, *...*>);

Arguments

varname
specifies the name of SAS character or numeric variables.

Details

The CALL MISSING routine assigns an ordinary numeric missing value (.) to each numeric variable in the argument list.

The CALL MISSING routine assigns a character missing value (a blank) to each character variable in the argument list. If the current length of the character variable equals the maximum length, the current length is not changed. Otherwise, the current length is set to 1.

You can mix character and numeric variables in the argument list.

Comparison

The MISSING function checks whether the argument has a missing value but does not change the value of the argument.

Examples

SAS Statements	Results

`prod='shoes';`
`invty=7498;`
`sales=23759;`
`call missing(sales);`
`put prod= invty= sales=;`
	`prod=shoes invty=7498 sales=.`

`prod='shoes';`
`invty=7498;`
`sales=23759;`
`call missing(prod,invty);`
`put prod= invty= sales=;`
	`prod= invty=. sales=23759`

`prod='shoes';`
`invty=7498;`
`sales=23759;`
`call missing(of _all_);`
`put prod= invty= sales=;`
	`prod= invty=. sales=.`

See Also

Function:

"MISSING Function" on page 929

"How to Set Variable Values to Missing in a Data Step" in *SAS Language Reference: Concepts*

CALL MODULE Routine

Calls an external routine without any return code.

Category: External Routines

Syntax

CALL MODULE(<*cntl-string*,>*module-name*<,*argument-1, ..., argument-n*>);

Arguments

cntl-string
is an optional control string whose first character must be an asterisk (*), followed by any combination of the following characters:

I	prints the hexadecimal representations of all arguments to the CALL MODULE routine. You can use this option to help diagnose problems caused by incorrect arguments or attribute tables. If you specify the I option, the E option is implied.
E	prints detailed error messages. Without the E option (or the I option, which supersedes it), the only error message that the CALL MODULE routine generates is "Invalid argument to function," which is usually not enough information to determine the cause of the error. The E option is useful for a production environment, while the I option is preferable for a development or debugging environment.
H	provides brief help information about the syntax of the CALL MODULE routine, the attribute file format, and suggested SAS formats and informats.

module-name
is the name of the external module to use.

argument
is one or more arguments to pass to the requested routine.

CAUTION:
Be sure to use the correct arguments and attributes. If you use incorrect arguments or attributes, you can cause the SAS System, and possibly your operating system, to fail. △

Details

The CALL MODULE routine executes a routine *module-name* that resides in an external library with the specified arguments.

CALL MODULE builds a parameter list using the information in the arguments and a routine description and argument attribute table that you define in a separate file. The attribute table is a sequential text file that contains descriptions of the routines that you can invoke with the CALL MODULE routine. The purpose of the table is to define how CALL MODULE should interpret its supplied arguments when it builds a parameter list to pass to the external routine. The attribute table should contain a

description for each external routine that you intend to call, and descriptions of each argument associated with that routine.

Before you invoke CALL MODULE, you must define the fileref of SASCBTBL to point to the external file that contains the attribute table. You can name the file whatever you want when you create it. This way, you can use SAS variables and formats as arguments to CALL MODULE and ensure that these arguments are properly converted before being passed to the external routine. If you do not define this fileref, CALL MODULE calls the requested routine without altering the arguments.

CAUTION:
Using the CALL MODULE routine without a defined attribute table can cause the SAS System to fail or force you to reset your computer. You need to use an attribute table for all external functions that you want to invoke. △

Comparisons

The two CALL routines and four functions share identical syntax:

- The MODULEN and MODULEC functions return a number and a character, respectively, while the routine CALL MODULE does not return a value.
- The CALL MODULEI routine and the functions MODULEIC and MODULEIN permit vector and matrix arguments. Their return values are scalar. You can invoke CALL MODULEI, MODULEIC, and MODULEIN only from the IML procedure.

Examples

Example 1: Using the CALL MODULE Routine This example calls the **xyz** routine. Use the following attribute table:

```
routine xyz minarg=2 maxarg=2;
arg 1 input num byvalue format=ib4.;
arg 2 output char format=$char10.;
```

The following is the sample SAS code that calls the **xyz** function:

```
data _null_;
   call module('xyz',1,x);
run;
```

Example 2: Using the MODULEIN Function in the IML Procedure This example invokes the **changi** routine from the TRYMOD.DLL module on a Windows platform. Use the following attribute table:

```
routine changi module=trymod returns=long;
arg 1 input num format=ib4. byvalue;
arg 2 update num format=ib4.;
```

The following PROC IML code calls the **changi** function:

```
proc iml;
   x1=J(4,5,0);
   do i=1 to 4;
      do j=1 to 5;
         x1[i,j]=i*10+j+3;
      end;
   end;
```

```
   y1=x1;
   x2=x1;
   y2=y1;
   rc=modulein('*i','changi',6,x2);
```

Example 3: Using the MODULEN Function This example calls the **Beep** routine, which is part of the Win32 API in the KERNEL32 Dynamic Link Library on a Windows platform. Use the following attribute table:

```
routine Beep
   minarg=2
   maxarg=2
   stackpop=called
   callseq=byvalue
   module=kernel32;
arg 1 num format=pib4.;
arg 2 num format=pib4.;
```

Assume that you name the attribute table file 'myatttbl.dat'. The following is the sample SAS code that calls the **Beep** function:

```
filename sascbtbl 'myatttbl.dat';
data _null_;
   rc=modulen("*e","Beep",1380,1000);
run;
```

The previous code causes the computer speaker to beep.

See Also

Functions and CALL Routines:

"MODULEC Function" on page 933
"MODULEN Function" on page 933

CALL POKE Routine

Writes a value directly into memory on a 32-bit platform.

Category: Special
Restriction: Use on 32-bit platforms only.

Syntax

CALL POKE(*source*,*pointer*,<*length*>,<*floating-point*>);

Arguments

source
specifies a constant, variable, or expression that contains a value to write into memory.

pointer
specifies a numeric expression that contains the virtual address of the data that the CALL POKE routine alters.

length
specifies a numeric constant, variable, or expression that contains the number of bytes to write from the *source* to the address that is indicated by *pointer*. If you omit *length*, the action that the CALL POKE routine takes depends on whether *source* is a character value or a numeric value:

- If *source* is a character value, the CALL POKE routine copies the entire value of *source* to the specified memory location.
- If *source* is a numeric value, the CALL POKE routine converts *source* into a long integer and writes into memory the number of bytes that constitute a pointer.

  Operating Environment Information: Under z/OS, pointers are 3 or 4 bytes long, depending on the situation. △

floating-point
specifies that the value of *source* is stored as a floating-point number. The value of *floating-point* can be any number.

Tip: If you do not use the *floating-point* argument, then *source* is stored as an integer value.

Details

CAUTION:
The CALL POKE routine is intended only for experienced programmers in specific cases. If you plan to use this routine, use extreme care both in your programming and in your typing. Writing directly into memory can cause devastating problems. This routine bypasses the normal safeguards that prevent you from destroying a vital element in your SAS session or in another piece of software that is active at the time. △

If you do not have access to the memory location that you specify, the CALL POKE routine returns an "Invalid argument" error.

You cannot use the CALL POKE routine on 64-bit platforms. If you attempt to use it, SAS writes a message to the log stating that this restriction applies. If you have legacy applications that use CALL POKE, change the applications and use CALL POKELONG instead. You can use CALL POKELONG on both 32–bit and 64–bit platforms.

If you use the fourth argument, then a floating-point number is assumed to be the value that is stored. If you do not use the fourth argument, then an integer value is assumed to be stored.

See Also

Functions and CALL Routines:

"ADDR Function" on page 371
"CALL POKELONG Routine" on page 478
"PEEK Function" on page 1001
"PEEKC Function" on page 1002

CALL POKELONG Routine

Writes a value directly into memory on 32-bit and 64-bit platforms.

Category: Special

Syntax

CALL POKELONG(*source*,*pointer*<,*length*>,<*floating-point*>)

Arguments

source
specifies a character constant, variable, or expression that contains a value to write into memory.

pointer
specifies a character string that contains the virtual address of the data that the CALL POKELONG routine alters.

length
specifies a numeric SAS expression that contains the number of bytes to write from the *source* to the address that is indicated by the *pointer*. If you omit *length*, the CALL POKELONG routine copies the entire value of *source* to the specified memory location.

floating-point
specifies that the value of *source* is stored as a floating-point number. The value of *floating-point* can be any number.

Tip: If you do not use the *floating-point* argument, then *source* is stored as an integer value.

Details

CAUTION:
The CALL POKELONG routine is intended only for experienced programmers in specific cases. If you plan to use this routine, use extreme care both in your programming and in your typing. *Writing directly into memory can cause devastating problems.* It bypasses the normal safeguards that prevent you from destroying a vital element in your SAS session or in another piece of software that is active at the time. △

If you do not have access to the memory location that you specify, the CALL POKELONG routine returns an "Invalid argument" error.

If you use the fourth argument, then a floating-point number is assumed to be the value that is stored. If you do not use the fourth argument, then an integer value is assumed to be stored.

See Also

Functions and CALL Routines:

"CALL POKE Routine" on page 477

CALL PRXCHANGE Routine

Performs a pattern-matching replacement.

Category: Character String Matching

Restriction: Use with the PRXPARSE function.

Interaction: When invoked by the %SYSCALL macro statement, CALL PRXCHANGE removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL PRXCHANGE (*regular-expression-id*, *times*, *old-string* <, *new-string* <, *result-length* <, *truncation-value* <, *number-of-changes*>>>>);

Arguments

regular-expression-id
specifies a numeric variable with a value that is a pattern identifier that is returned from the PRXPARSE function.

times
is a numeric constant, variable, or expression that specifies the number of times to search for a match and replace a matching pattern.

Tip: If the value of *times* is -1, then all matching patterns are replaced.

old-string
specifies the character expression on which to perform a search and replace.

Tip: All changes are made to *old-string* if you do not use the *new-string* argument.

new-string
specifies a character variable in which to place the results of the change to *old-string*.

Tip: If you use the *new-string* argument in the call to the PRXCHANGE routine, then *old-string* is not modified.

result-length
is a numeric variable with a return value that is the number of characters that are copied into the result.

Tip: Trailing blanks in the value of *old-string* are not copied to *new-string*, and are therefore not included as part of the length in *result-length*.

truncation-value
is a numeric variable with a returned value that is either 0 or 1, depending on the result of the change operation:

0	if the entire replacement result is not longer than the length of *new-string*.
1	if the entire replacement result is longer than the length of *new-string*.

number-of-changes
is a numeric variable with a returned value that is the total number of replacements that were made. If the result is truncated when it is placed into *new-string*, the value of *number-of-changes* is not changed.

Details

The CALL PRXCHANGE routine matches and replaces a pattern. If the value of *times* is -1, the replacement is performed as many times as possible.

For more information about pattern matching, see “Pattern Matching Using Perl Regular Expressions (PRX)” on page 333.

Comparisons

The CALL PRXCHANGE routine is similar to the PRXCHANGE function except that the CALL routine returns the value of the pattern matching replacement as one of its parameters instead of as a return argument.

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in “Functions and CALL Routines by Category” on page 345.

Examples

The following example replaces all occurrences of cat, rat, or bat with the value TREE.

```
data _null_;
      /* Use a pattern to replace all occurrences of cat,      */
      /* rat, or bat with the value TREE.                      */
   length text $ 46;
   RegularExpressionId = prxparse('s/[crb]at/tree/');
   text = 'The woods have a bat, cat, bat, and a rat!';
      /* Use CALL PRXCHANGE to perform the search and replace. */
      /* Because the argument times has a value of -1, the     */
      /* replacement is performed as many times as possible.   */
   call prxchange(RegularExpressionId, -1, text);
   put text;
run;
```

SAS writes the following line to the log:

```
The woods have a tree, tree, tree, and a tree!
```

See Also

Functions and CALL routines:

“CALL PRXDEBUG Routine” on page 482

“CALL PRXFREE Routine” on page 484

“CALL PRXNEXT Routine” on page 485

“CALL PRXPOSN Routine” on page 487

“CALL PRXSUBSTR Routine” on page 490

“PRXCHANGE Function” on page 1039

“PRXPAREN Function” on page 1049

“PRXMATCH Function” on page 1045

“PRXPARSE Function” on page 1051

“PRXPOSN Function” on page 1053

CALL PRXDEBUG Routine

Enables Perl regular expressions in a DATA step to send debugging output to the SAS log.

Category: Character String Matching

Restriction: Use with the CALL PRXCHANGE, CALL PRXFREE, CALL PRXNEXT, CALL PRXPOSN, CALL PRXSUBSTR, PRXPARSE, PRXPAREN, and PRXMATCH functions and CALL routines. The PRXPARSE function is not DBCS compatible.

Syntax

CALL PRXDEBUG (*on-off*);

Arguments

on-off
specifies a numeric constant, variable, or expression. If the value of *on-off* is positive and non-zero, then debugging is turned on. If the value of *on-off* is zero, then debugging is turned off.

Details

The CALL PRXDEBUG routine provides information about how a Perl regular expression is compiled, and about which steps are taken when a pattern is matched to a character value.

You can turn debugging on and off multiple times in your program if you want to see debugging output for particular Perl regular expression function calls.

For more information about pattern matching, see “Pattern Matching Using Perl Regular Expressions (PRX)” on page 333.

Comparisons

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in “Functions and CALL Routines by Category” on page 345.

Examples

The following example produces debugging output.

```
data _null_;

      /* Turn the debugging option on. */
   call prxdebug(1);
   putlog 'PRXPARSE: ';
   re = prxparse('/[bc]d(ef*g)+h[ij]k$/');
   putlog 'PRXMATCH: ';
   pos = prxmatch(re, 'abcdefg_gh_');
```

```
      /* Turn the debugging option off. */
   call prxdebug(0);
run;
```

The following lines are written to the SAS log.

Output 4.13 SAS Log Results from CALL PRXDEBUG

```
PRXPARSE:
Compiling REx '[bc]d(ef*g)+h[ij]k$' ❶
size 41 first at 1 ❷
rarest char g at 0 ❺
rarest char d at 0
   1: ANYOF[bc](10) ❸
  10: EXACT <d>(12)
  12: CURLYX[0] {1,32767}(26)
  14:   OPEN1(16)
  16:     EXACT <e>(18)
  18:     STAR(21)
  19:       EXACT <f>(0)
  21:     EXACT <g>(23)
  23:   CLOSE1(25)
  25:   WHILEM[1/1](0)
  26: NOTHING(27)
  27: EXACT <h>(29)
  29: ANYOF[ij](38)
  38: EXACT <k>(40)
  40: EOL(41)
  41: END(0)
anchored 'de' at 1 floating 'gh' at 3..2147483647 (checking floating)  ❹
stclass 'ANYOF[bc]' minlen 7 ❻                                        ❹

PRXMATCH:
Guessing start of match, REx '[bc]d(ef*g)+h[ij]k$' against 'abcdefg_gh_'...
Did not find floating substr 'gh'...
Match rejected by optimizer
```

The following items correspond to the lines that are numbered in the SAS log that is shown above.

❶ This line shows the precompiled form of the Perl regular expression.

❷ Size specifies a value in arbitrary units of the compiled form of the Perl regular expression. 41 is the label ID of the first node that performs a match.

❸ This line begins a list of program nodes in compiled form for regular expressions.

❹ These two lines provide optimizer information. In the example above, the optimizer found that the match should contain the substring **de** at offset 1, and the substring **gh** at an offset between 3 and infinity. To rule out a pattern match quickly, Perl checks substring **gh** before it checks substring **de**.

The optimizer might use the information that the match begins at the *first* ID (❷), with a character class (❺), and cannot be shorter than seven characters (❻).

See Also

Functions and CALL routines:

"CALL PRXCHANGE Routine" on page 479

"CALL PRXFREE Routine" on page 484

"CALL PRXNEXT Routine" on page 485

"CALL PRXPOSN Routine" on page 487
"CALL PRXSUBSTR Routine" on page 490
"CALL PRXCHANGE Routine" on page 479
"PRXCHANGE Function" on page 1039
"PRXPAREN Function" on page 1049
"PRXMATCH Function" on page 1045
"PRXPARSE Function" on page 1051
"PRXPOSN Function" on page 1053

CALL PRXFREE Routine

Frees memory that was allocated for a Perl regular expression.

Category: Character String Matching
Restriction: Use with the PRXPARSE function.

Syntax

CALL PRXFREE (*regular-expression-id*);

Arguments

regular-expression-id
specifies a numeric variable with a value that is the identification number that is returned by the PRXPARSE function. *regular-expression-id* is set to missing if the call to the PRXFREE routine occurs without error.

Details

The CALL PRXFREE routine frees unneeded resources that were allocated for a Perl regular expression.

For more information about pattern matching, see "Pattern Matching Using Perl Regular Expressions (PRX)" on page 333.

Comparisons

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in "Functions and CALL Routines by Category" on page 345.

See Also

Functions and CALL routines:

"CALL PRXCHANGE Routine" on page 479

"CALL PRXDEBUG Routine" on page 482
"CALL PRXNEXT Routine" on page 485
"CALL PRXPOSN Routine" on page 487
"CALL PRXSUBSTR Routine" on page 490
"CALL PRXCHANGE Routine" on page 479
"PRXCHANGE Function" on page 1039
"PRXPAREN Function" on page 1049
"PRXPAREN Function" on page 1049
"PRXPARSE Function" on page 1051
"PRXPOSN Function" on page 1053

CALL PRXNEXT Routine

Returns the position and length of a substring that matches a pattern, and iterates over multiple matches within one string.

Category: Character String Matching

Restriction: Use with the PRXPARSE function.

Interaction: When invoked by the %SYSCALL macro statement, CALL PRXNEXT removes the quotation marks from arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL PRXNEXT (*regular-expression-id*, *start*, *stop*, *source*, *position*, *length*);

Arguments

regular-expression-id
specifies a numeric variable with a value that is the identification number that is returned by the PRXPARSE function.

start
is a numeric variable that specifies the position at which to start the pattern matching in *source*. If the match is successful, CALL PRXNEXT returns a value of *position* + MAX(1, *length*). If the match is not successful, the value of *start* is not changed.

stop
is a numeric constant, variable, or expression that specifies the last character to use in *source*. If *stop* is -1, then the last character is the last non-blank character in *source*.

source
specifies a character constant, variable, or expression that you want to search.

position
is a numeric variable with a returned value that is the position in *source* at which the pattern begins. If no match is found, CALL PRXNEXT returns zero.

length
is a numeric variable with a returned value that is the length of the string that is matched by the pattern. If no match is found, CALL PRXNEXT returns zero.

Details

The CALL PRXNEXT routine searches the variable *source* with a pattern. It returns the position and length of a pattern match that is located between the *start* and the *stop* positions in *source*. Because the value of the *start* parameter is updated to be the position of the next character that follows a match, CALL PRXNEXT enables you to search a string for a pattern multiple times in succession.

For more information about pattern matching, see "Pattern Matching Using Perl Regular Expressions (PRX)" on page 333.

Comparisons

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in "Functions and CALL Routines by Category" on page 345.

Examples

The following example finds all instances of cat, rat, or bat in a text string.

```
data _null_;
   ExpressionID = prxparse('/[crb]at/');
   text = 'The woods have a bat, cat, and a rat!';
   start = 1;
   stop = length(text);

      /* Use PRXNEXT to find the first instance of the pattern, */
      /* then use DO WHILE to find all further instances.       */
      /* PRXNEXT changes the start parameter so that searching  */
      /* begins again after the last match.                     */
   call prxnext(ExpressionID, start, stop, text, position, length);
      do while (position > 0);
         found = substr(text, position, length);
         put found= position= length=;
         call prxnext(ExpressionID, start, stop, text, position, length);
      end;
run;
```

The following lines are written to the SAS log:

```
found=bat position=18 length=3
found=cat position=23 length=3
found=rat position=34 length=3
```

See Also

Functions and CALL routines:

"CALL PRXCHANGE Routine" on page 479
"CALL PRXDEBUG Routine" on page 482

"CALL PRXFREE Routine" on page 484
"CALL PRXPOSN Routine" on page 487
"CALL PRXSUBSTR Routine" on page 490
"CALL PRXCHANGE Routine" on page 479
"PRXCHANGE Function" on page 1039
"PRXPAREN Function" on page 1049
"PRXMATCH Function" on page 1045
"PRXPARSE Function" on page 1051
"PRXPOSN Function" on page 1053

CALL PRXPOSN Routine

Returns the start position and length for a capture buffer.

Category: Character String Matching
Restriction: Use with the PRXPARSE function.

Syntax

CALL PRXPOSN (*regular-expression-id*, *capture-buffer*, *start* <, *length*>);

Arguments

regular-expression-id
specifies a numeric variable with a value that is a pattern identifier that is returned by the PRXPARSE function.

capture-buffer
is a numeric constant, variable, or expression with a value that identifies the capture buffer from which to retrieve the start position and length:

- If the value of *capture-buffer* is zero, CALL PRXPOSN returns the start position and length of the entire match.
- If the value of *capture-buffer* is between 1 and the number of open parentheses, CALL PRXPOSN returns the start position and length for that capture buffer.
- If the value of *capture-buffer* is greater than the number of open parentheses, CALL PRXPOSN returns missing values for the start position and length.

start
is a numeric variable with a returned value that is the position at which the capture buffer is found:

- If the value of *capture-buffer* is not found, CALL PRXPOSN returns a zero value for the start position.
- If the value of *capture-buffer* is greater than the number of open parentheses in the pattern, CALL PRXPOSN returns a missing value for the start position.

length
is a numeric variable with a returned value that is the pattern length of the previous pattern match:

- If the pattern match is not found, CALL PRXPOSN returns a zero value for the length.
- If the value of *capture-buffer* is greater than the number of open parentheses in the pattern, CALL PRXPOSN returns a missing value for length.

Details

The CALL PRXPOSN routine uses the results of PRXMATCH, PRXSUBSTR, PRXCHANGE, or PRXNEXT to return a capture buffer. A match must be found by one of these functions for the CALL PRXPOSN routine to return meaningful information.

A capture buffer is part of a match, enclosed in parentheses, that is specified in a regular expression. CALL PRXPOSN does not return the text for the capture buffer directly. It requires a call to the SUBSTR function to return the text.

For more information about pattern matching, see "Pattern Matching Using Perl Regular Expressions (PRX)" on page 333.

Comparisons

The CALL PRXPOSN routine is similar to the PRXPOSN function, except that CALL PRXPOSN returns the position and length of the capture buffer rather than the capture buffer itself.

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in "Functions and CALL Routines by Category" on page 345.

Examples

Example 1: Finding Submatches within a Match The following example searches a regular expression and calls the PRXPOSN routine to find the position and length of three submatches.

```
data _null_;
   patternID = prxparse('/(\d\d):(\d\d)(am|pm)/');
   text = 'The time is 09:56am.';

   if prxmatch(patternID, text) then do;
      call prxposn(patternID, 1, position, length);
      hour = substr(text, position, length);
      call prxposn(patternID, 2, position, length);
      minute = substr(text, position, length);
      call prxposn(patternID, 3, position, length);
      ampm = substr(text, position, length);

      put hour= minute= ampm=;
      put text=;
   end;
run;
```

SAS writes the following lines to the log:

```
hour=09 minute=56 ampm=am
text=The time is 09:56am.
```

Example 2: Parsing Time Data The following example parses time data and writes the results to the SAS log.

```
data _null_;
   if _N_ = 1 then
   do;
      retain patternID;
      pattern = "/(\d+):(\d\d)(?:\.(\d+))?/";
      patternID = prxparse(pattern);
   end;

   array match[3] $ 8;
   input minsec $80.;
   position = prxmatch(patternID, minsec);
   if position ^= 0 then
   do;
      do i = 1 to prxparen(patternID);
         call prxposn(patternID, i, start, length);
         if start ^= 0 then
            match[i] = substr(minsec, start, length);
      end;
      put match[1] "minutes, " match[2] "seconds" @;
      if ^missing(match[3]) then
         put ", " match[3] "milliseconds";
   end;
   datalines;
14:56.456
45:32
;
```

SAS writes the following lines to the log:

```
14 minutes, 56 seconds, 456 milliseconds
45 minutes, 32 seconds
```

See Also

Functions and CALL routines:

- “CALL PRXCHANGE Routine” on page 479
- “CALL PRXDEBUG Routine” on page 482
- “CALL PRXFREE Routine” on page 484
- “CALL PRXNEXT Routine” on page 485
- “CALL PRXSUBSTR Routine” on page 490
- “CALL PRXCHANGE Routine” on page 479
- “PRXCHANGE Function” on page 1039
- “PRXPAREN Function” on page 1049
- “PRXMATCH Function” on page 1045
- “PRXPARSE Function” on page 1051
- “PRXPOSN Function” on page 1053

CALL PRXSUBSTR Routine

Returns the position and length of a substring that matches a pattern.

Category: Character String Matching

Restriction: Use with the PRXPARSE function.

Interaction: When invoked by the %SYSCALL macro statement, CALL PRXSUBSTR removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL PRXSUBSTR (*regular-expression-id*, *source*, *position* <, *length*>);

Arguments

regular-expression-id
specifies a numeric variable with a value that is an identification number that is returned by the PRXPARSE function.

source
specifies a character constant, variable, or expression that you want to search.

position
is a numeric variable with a returned value that is the position in *source* where the pattern begins. If no match is found, CALL PRXSUBSTR returns zero.

length
is a numeric variable with a returned value that is the length of the substring that is matched by the pattern. If no match is found, CALL PRXSUBSTR returns zero.

Details

The CALL PRXSUBSTR routine searches the variable *source* with the pattern from PRXPARSE, returns the position of the start of the string, and if specified, returns the length of the string that is matched. By default, when a pattern matches more than one character that begins at a specific position, CALL PRXSUBSTR selects the longest match.

For more information about pattern matching, see "Pattern Matching Using Perl Regular Expressions (PRX)" on page 333.

Comparisons

CALL PRXSUBSTR performs the same matching as PRXMATCH, but CALL PRXSUBSTR additionally enables you to use the *length* argument to receive more information about the match.

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in "Functions and CALL Routines by Category" on page 345.

Examples

Example 1: Finding the Position and Length of a Substring The following example searches a string for a substring, and returns its position and length in the string.

```
data _null_;
      /* Use PRXPARSE to compile the Perl regular expression. */
   patternID = prxparse('/world/');
      /* Use PRXSUBSTR to find the position and length of the string. */
   call prxsubstr(patternID, 'Hello world!', position, length);
   put position= length=;
run;
```

The following line is written to the SAS log:

```
position=7 length=5
```

Example 2: Finding a Match in a Substring The following example searches for addresses that contain avenue, drive, or road, and extracts the text that was found.

```
data _null_;
   if _N_ = 1 then
   do;
      retain ExpressionID;

         /* The i option specifies a case insensitive search. */
      pattern = "/ave|avenue|dr|drive|rd|road/i";
      ExpressionID = prxparse(pattern);
   end;

   input street $80.;
   call prxsubstr(ExpressionID, street, position, length);
   if position ^= 0 then
   do;
      match = substr(street, position, length);
      put match:$QUOTE. "found in " street:$QUOTE.;
   end;
   datalines;
153 First Street
6789 64th Ave
4 Moritz Road
7493 Wilkes Place
;

run;
```

The following lines are written to the SAS log:

```
"Ave" found in "6789 64th Ave"
"Road" found in "4 Moritz Road"
```

See Also

Functions and CALL routines:

"CALL PRXCHANGE Routine" on page 479

"CALL PRXDEBUG Routine" on page 482
"CALL PRXFREE Routine" on page 484
"CALL PRXNEXT Routine" on page 485
"CALL PRXPOSN Routine" on page 487
"PRXCHANGE Function" on page 1039
"PRXPAREN Function" on page 1049
"PRXMATCH Function" on page 1045
"PRXPARSE Function" on page 1051
"PRXPOSN Function" on page 1053

CALL RANBIN Routine

Returns a random variate from a binomial distribution.

Category: Random Number

Syntax

CALL RANBIN(*seed,n,p,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANBIN is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

n
is an integer number of independent Bernoulli trials.

Range: $n > 0$

p
is a numeric probability of success parameter.

Range: $0<p<1$

x
is a numeric SAS variable. A new value for the random variate *x* is returned each time CALL RANBIN is executed.

Details

The CALL RANBIN routine updates *seed* and returns a variate *x* that is generated from a binomial distribution with mean *np* and variance $np(1–p)$. If $n \leq 50$, $np \leq 5$, or $n(1–p) \leq 5$, SAS uses an inverse transform method applied to a RANUNI uniform variate. If $n>50$, $np>5$, and $n(1–p)>5$, SAS uses the normal approximation to the

binomial distribution. In that case, the Box-Muller transformation of RANUNI uniform variates is used.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANBIN routine gives greater control of the seed and random number streams than does the RANBIN function.

Examples

The following example uses the CALL RANBIN routine:

```
options pageno=1 nodate ls=80 ps=64;

data u1 (keep = x);
   seed = 104;
   do i = 1 to 5;
      call ranbin(seed, 2000, 0.2 ,x);
      output;
   end;
   call symputx('seed', seed);
run;

data u2 (keep = x);
   seed = &seed;
   do i = 1 to 5;
      call ranbin(seed, 2000, 0.2 ,x);
      output;
   end;
run;

data all;
   set u1 u2;
   z = ranbin(104, 2000, 0.2);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.14 Output from the CALL RANBIN Routine

```
                              The SAS System                                1

                               Separate    Single
                          Obs   Streams    Stream

                            1      423       423
                            2      418       418
                            3      403       403
                            4      394       394
                            5      429       429
                            6      369       369
                            7      413       413
                            8      417       417
                            9      400       400
                           10      383       383
```

See Also

Functions:

"RAND Function" on page 1069
"RANBIN Function" on page 1067

CALL RANCAU Routine

Returns a random variate from a Cauchy distribution.

Category: Random Number

Syntax

CALL RANCAU(*seed,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANCAU is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

x
is a numeric SAS variable. A new value for the random variate *x* is returned each time CALL RANCAU is executed.

Details

The CALL RANCAU routine updates *seed* and returns a variate *x* that is generated from a Cauchy distribution that has a location parameter of 0 and scale parameter of 1.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

An acceptance-rejection procedure applied to RANUNI uniform variates is used. If u and v are independent uniform (–1/2, 1/2) variables and $u^2+v^2 \leq 1/4$, then u/v is a Cauchy variate.

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANCAU routine gives greater control of the seed and random number streams than does the RANCAU function.

Examples

```
options nodate pageno=1 linesize=80 pagesize=60;

data case;
   retain Seed_1 Seed_2 Seed_3 45;
   do i=1 to 10;
      call rancau(Seed_1,X1);
      call rancau(Seed_2,X2);
      X3=rancau(Seed_3);
      if i=5 then
          do;
             Seed_2=18;
             Seed_3=18;
          end;
      output;
   end;
run;

proc print;
   id i;
   var Seed_1-Seed_3 X1-X3;
run;
```

This example uses the CALL RANCAU routine:

```
options pageno=1 ls=80 ps=64 nodate;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call rancau(seed, X);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
```

```
   do i = 1 to 5;
      call rancau(seed, X);
      output;
   end;
run;

data all;
   set u1 u2;
   z = rancau(104);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.15 Output from the CALL RANCAU Routine

```
                              The SAS System                                1

                              Separate      Single
                          Obs  Streams      Stream

                            1   -0.6780    -0.6780
                            2    0.1712     0.1712
                            3    1.1372     1.1372
                            4    0.1478     0.1478
                            5   16.6536    16.6536
                            6    0.0747     0.0747
                            7   -0.5872    -0.5872
                            8    1.4713     1.4713
                            9    0.1792     0.1792
                           10   -0.0473    -0.0473
```

See Also

Functions:

"RAND Function" on page 1069
"RANCAU Function" on page 1068

CALL RANEXP Routine

Returns a random variate from an exponential distribution.

Category: Random Number

Syntax

CALL RANEXP(*seed,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANEXP is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

x
is a numeric variable. A new value for the random variate *x* is returned each time CALL RANEXP is executed.

Details

The CALL RANEXP routine updates *seed* and returns a variate *x* that is generated from an exponential distribution that has a parameter of 1.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

The CALL RANEXP routine uses an inverse transform method applied to a RANUNI uniform variate.

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANEXP routine gives greater control of the seed and random number streams than does the RANEXP function.

Examples

This example uses the CALL RANEXP routine:

```
options pageno=1 ls=80 ps=64 nodate;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call ranexp(seed, x);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
      call ranexp(seed, x);
      output;
   end;
run;

data all;
   set u1 u2;
   z = ranexp(104);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.16 Output from the CALL RANEXP Routine

```
                              The SAS System                              1

                          Separate     Single
                    Obs    Streams     Stream

                     1     1.44347    1.44347
                     2     0.11740    0.11740
                     3     0.54175    0.54175
                     4     0.02280    0.02280
                     5     0.16645    0.16645
                     6     0.21711    0.21711
                     7     0.75538    0.75538
                     8     1.21760    1.21760
                     9     1.72273    1.72273
                    10     0.08021    0.08021
```

See Also

Functions:

"RAND Function" on page 1069
"RANEXP Function" on page 1082

CALL RANGAM Routine

Returns a random variate from a gamma distribution.

Category: Random Number

Syntax

CALL RANGAM(*seed,a,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANGAM is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

a
is a numeric shape parameter.

Range: $a > 0$

x
is a numeric variable. A new value for the random variate *x* is returned each time CALL RANGAM is executed.

Details

The CALL RANGAM routine updates *seed* and returns a variate *x* that is generated from a gamma distribution with parameter *a*.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

For $a>1$, an acceptance-rejection method by Cheng is used (Cheng, 1977; see in "References" on page 1255). For $a\leq 1$, an acceptance-rejection method by Fishman is used (Fishman, 1978; see in "References" on page 1255).

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANGAM routine gives greater control of the seed and random number streams than does the RANGAM function.

Examples

This example uses the CALL RANGAM routine:

```
options nodate pageno=1 linesize=80 pagesize=60;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call rangam(seed, x);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
      call rangam(seed, x);
      output;
   end;
run;

data all;
   set u1 u2;
   z = rangam(104);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.17 Output from the CALL RANGAM Routine

The SAS System 1

Obs	Separate Streams	Single Stream
1	1.44347	1.44347
2	0.11740	0.11740
3	0.54175	0.54175
4	0.02280	0.02280
5	0.16645	0.16645
6	0.21711	0.21711
7	0.75538	0.75538
8	1.21760	1.21760
9	1.72273	1.72273
10	0.08021	0.08021

Output 4.18 The RANGAM Example

The SAS System 1

i	Seed_1	Seed_2	Seed_3	X1	X2	X3
1	1404437564	1404437564	45	1.30569	1.30569	1.30569
2	1326029789	1326029789	45	1.87514	1.87514	1.87514
3	1988843719	1988843719	45	1.71597	1.71597	1.71597
4	50049159	50049159	45	1.59304	1.59304	1.59304
5	802575599	18	18	0.43342	0.43342	0.43342
6	100573943	991271755	18	1.11812	1.32646	1.11812
7	1986749826	1437043694	18	0.68415	0.88806	0.68415
8	52428589	959908645	18	1.62296	2.46091	1.62296
9	1216356463	1225034217	18	2.26455	4.06596	2.26455
10	805366679	425626811	18	2.16723	6.94703	2.16723

Changing Seed_2 for the CALL RANGAM statement, when I=5, forces the stream of the variates for X2 to deviate from the stream of the variates for X1. Changing Seed_3 on the RANGAM function, however, has no effect.

See Also

Functions:

"RAND Function" on page 1069

"RANGAM Function" on page 1083

CALL RANNOR Routine

Returns a random variate from a normal distribution.

Category: Random Number

Syntax

CALL RANNOR(*seed,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANNOR is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

x
is a numeric variable. A new value for the random variate *x* is returned each time CALL RANNOR is executed.

Details

The CALL RANNOR routine updates *seed* and returns a variate *x* that is generated from a normal distribution, with mean 0 and variance 1.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

The CALL RANNOR routine uses the Box-Muller transformation of RANUNI uniform variates.

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANNOR routine gives greater control of the seed and random number streams than does the RANNOR function.

Examples

This example uses the CALL RANNOR routine:

```
options pageno=1 ls=80 ps=64 nodate;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call rannor(seed, X);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
```

```
      call rannor(seed, X);
      output;
   end;
run;

data all;
   set u1 u2;
   z = rannor(104);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.19 Output from the CALL RANNOR Routine

```
                              The SAS System                                1

                              Separate     Single
                        Obs    Streams     Stream

                          1     1.30390     1.30390
                          2     1.03049     1.03049
                          3     0.19491     0.19491
                          4    -0.34987    -0.34987
                          5     1.64273     1.64273
                          6    -1.75842    -1.75842
                          7     0.75080     0.75080
                          8     0.94375     0.94375
                          9     0.02436     0.02436
                         10    -0.97256    -0.97256
```

See Also

Functions:

"RAND Function" on page 1069

"RANNOR Function" on page 1086

CALL RANPERK Routine

Randomly permutes the values of the arguments, and returns a permutation of *k* out of *n* values.

Category: Combinatorial

Syntax

CALL RANPERK(*seed*, *k*, *variable-1*<, *variable–2, ...*>);

Arguments

seed
is a numeric variable that contains the random number seed. For more information about seeds, see "Seed Values" on page 314.

k
is the number of values that you want to have in the random permutation.

variable
specifies all numeric variables, or all character variables that have the same length. *K* values of these variables are randomly permuted.

Details

Using CALL RANPERK with Macros You can call the RANPERK routine when you use the %SYSCALL macro. In this case, the *variable* arguments are not required to be the same type or length. If %SYSCALL identifies an argument as numeric, then %SYSCALL reformats the returned value.

If an error occurs during the execution of the CALL RANPERK routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than -100.

If there are no errors, then &SYSERR and &SYSINFO are set to zero.

Examples

Example 1: Using CALL RANPERK in a DATA Step The following example shows how to generate random permutations of given values by using the CALL RANPERK routine.

```
data _null_;
   array x x1-x5 (1 2 3 4 5);
   seed = 1234567890123;
   do n=1 to 10;
      call ranperk(seed, 3, of x1-x5);
      put seed= @20 ' x= ' x1-x3;
   end;
run;
```

Output 4.20 Log Output from Using the CALL RANPERK Routine in a DATA Step

```
seed=1332351321     x= 5 4 2
seed=829042065      x= 4 1 3
seed=767738639      x= 5 1 2
seed=1280236105     x= 3 2 5
seed=670350431      x= 4 3 5
seed=1956939964     x= 3 1 2
seed=353939815      x= 4 2 1
seed=1996660805     x= 3 4 5
seed=1835940555     x= 5 1 4
seed=910897519      x= 5 1 2
```

Example 2: Using CALL RANPERK with a Macro The following is an example of the CALL RANPERK routine that is used with macros.

```
%macro test;
   %let x1=ant;
   %let x2=-.1234;
   %let x3=1e10;
   %let x4=hippopotamus;
   %let x5=zebra;
   %let k=3;
   %let seed = 12345;
   %do j=1 %to 10;
      %syscall ranperk(seed, k, x1, x2, x3, x4, x5);
      %put j=&j   &x1 &x2 &x3;
   %end;
   %mend;

%test;
```

Output 4.21 Output from Using the CALL RANPERK Routine with a Macro

```
j=1   -0.1234 hippopotamus zebra
j=2   hippopotamus -0.1234 10000000000
j=3   hippopotamus ant zebra
j=4   -0.1234 zebra ant
j=5   -0.1234 ant hippopotamus
j=6   10000000000 hippopotamus ant
j=7   10000000000 hippopotamus ant
j=8   ant 10000000000 -0.1234
j=9   zebra -0.1234 10000000000
j=10   zebra hippopotamus 10000000000
```

See Also

Functions and CALL Routines:

"RAND Function" on page 1069
"CALL ALLPERM Routine" on page 437
"CALL RANPERM Routine" on page 505

CALL RANPERM Routine

Randomly permutes the values of the arguments.

Category: Combinatorial

Syntax

CALL RANPERM(*seed*, *variable-1*<, *variable–2, ...*>);

Arguments

seed
is a numeric variable that contains the random number seed. For more information about seeds, see "Seed Values" on page 314.

variable
specifies all numeric variables or all character variables that have the same length. The values of these variables are randomly permuted.

Details

Using CALL RANPERM with Macros You can call the RANPERM routine when you use the %SYSCALL macro. In this case, the *variable* arguments are not required to be the same type or length. If %SYSCALL identifies an argument as numeric, then %SYSCALL reformats the returned value.

If an error occurs during the execution of the CALL RANPERM routine, then both of the following values are set:

- &SYSERR is assigned a value that is greater than 4.
- &SYSINFO is assigned a value that is less than -100.

If there are no errors, then &SYSERR and &SYSINFO are set to zero.

Examples

Example 1: Using CALL RANPERM in a DATA Step The following example generates random permutations of given values by using the CALL RANPERM routine.

```
data _null_;
      array x x1-x4 (1 2 3 4);
      seed = 1234567890123;
      do n=1 to 10;
         call ranperm(seed, of x1-x4);
         put seed= @20 ' x= ' x1-x4;
      end;
run;
```

Output 4.22 Output from Using the CALL RANPERM Routine in a DATA Step

```
seed=1332351321     x= 1 3 2 4
seed=829042065      x= 3 4 2 1
seed=767738639      x= 4 2 3 1
seed=1280236105     x= 1 2 4 3
seed=670350431      x= 2 1 4 3
seed=1956939964     x= 2 4 3 1
seed=353939815      x= 4 1 2 3
seed=1996660805     x= 4 3 1 2
seed=1835940555     x= 4 3 2 1
seed=910897519      x= 3 2 1 4
```

Example 2: Using CALL RANPERM with a Macro The following is an example of the CALL RANPERM routine that is used with the %SYSCALL macro.

```
%macro test;
   %let x1=ant;
   %let x2=-.1234;
   %let x3=1e10;
   %let x4=hippopotamus;
   %let x5=zebra;
   %let seed = 12345;
   %do j=1 %to 10;
      %syscall ranperm(seed, x1, x2, x3, x4, x5);
      %put j=&j   &x1 &x2 &x3;
   %end;
   %mend;

%test;
```

Output 4.23 Output from Using the CALL RANPERM Routine with a Macro

```
j=1   zebra ant hippopotamus
j=2   10000000000 ant -0.1234
j=3   -0.1234 10000000000 ant
j=4   hippopotamus ant zebra
j=5   -0.1234 zebra 10000000000
j=6   -0.1234 hippopotamus ant
j=7   zebra ant -0.1234
j=8   -0.1234 hippopotamus ant
j=9   ant -0.1234 hippopotamus
j=10   -0.1234 zebra 10000000000
```

See Also

Functions and CALL Routines:

"RAND Function" on page 1069
"CALL ALLPERM Routine" on page 437
"CALL RANPERK Routine" on page 503

CALL RANPOI Routine

Returns a random variate from a Poisson distribution.

Category: Random Number

Syntax

CALL RANPOI(*seed,m,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANPOI is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

m
is a numeric mean parameter.

Range: $m \geq 0$

x
is a numeric variable. A new value for the random variate *x* is returned each time CALL RANPOI is executed.

Details

The CALL RANPOI routine updates *seed* and returns a variate *x* that is generated from a Poisson distribution, with mean *m*.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

For $m < 85$, an inverse transform method applied to a RANUNI uniform variate is used (Fishman, 1976; see in "References" on page 1255). For $m \geq 85$, the normal approximation of a Poisson random variable is used. To expedite execution, internal variables are calculated only on initial calls (that is, with each new *m*).

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANPOI routine gives greater control of the seed and random number streams than does the RANPOI function.

Examples

This example uses the CALL RANPOI routine:

```
options pageno=1 ls=80 ps=64 nodate;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call ranpoi(seed, 2000, x);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
      call ranpoi(seed, 2000, x);
      output;
   end;
run;

data all;
   set u1 u2;
   z = ranpoi(104, 2000);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.24 Output from the CALL RANPOI Routine

```
                              The SAS System                                1

                               Separate    Single
                          Obs   Streams    Stream

                            1      2058      2058
                            2      2046      2046
                            3      2009      2009
                            4      1984      1984
                            5      2073      2073
                            6      1921      1921
                            7      2034      2034
                            8      2042      2042
                            9      2001      2001
                           10      1957      1957
```

See Also

Functions:

"RAND Function" on page 1069
"RANPOI Function" on page 1087

CALL RANTBL Routine

Returns a random variate from a tabled probability distribution.

Category: Random Number

Syntax

CALL RANTBL(*seed*,p_1,...p_i,...p_n,*x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANTBL is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

***p**ᵢ*
is a numeric SAS value.

Range: $0 \leq p_i \leq 1$ for $0 < i \leq n$

x
is a numeric SAS variable. A new value for the random variate *x* is returned each time CALL RANTBL is executed.

Details

The CALL RANTBL routine updates *seed* and returns a variate *x* generated from the probability mass function defined by p_1 through p_n.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

An inverse transform method applied to a RANUNI uniform variate is used. The CALL RANTBL routine returns these data:

$$\begin{array}{ll} 1 & \text{with probability } p_1 \\ 2 & \text{with probability } p_2 \\ \vdots & \\ n & \text{with probability } p_n \\ n+1 & \text{with probability } 1 - \sum_{i=1}^{n} p_i \quad \text{if } \sum_{i=1}^{n} p_i \leq 1 \end{array}$$

If, for some index $j<n$,

$$\sum_{i=1}^{j} p_i \geq 1$$

RANTBL returns only the indices 1 through *j*, with the probability of occurrence of the index *j* equal to

$$1 - \sum_{i=1}^{j-1} p_i$$

For a discussion and example of an effective use of the random number CALL routines, see “Starting, Stopping, and Restarting a Stream” on page 326.

Comparisons

The CALL RANTBL routine gives greater control of the seed and random number streams than does the RANTBL function.

Examples

This example uses the CALL RANTBL routine:

```
options pageno=1 ls=80 ps=64 nodate;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call rantbl(seed, .02, x);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
      call rantbl(seed, .02, x);
      output;
   end;
run;

data all;
   set u1 u2;
   z = rantbl(104, .02);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.25 Output from the CALL RANTBL Routine

```
                              The SAS System                              1

                                Separate    Single
                          Obs    Streams    Stream

                            1        2          2
                            2        2          2
                            3        2          2
                            4        2          2
                            5        2          2
                            6        2          2
                            7        2          2
                            8        2          2
                            9        2          2
                           10        2          2
```

See Also

Functions:

"RAND Function" on page 1069
"RANTBL Function" on page 1088

CALL RANTRI Routine

Returns a random variate from a triangular distribution.

Category: Random Number

Syntax

CALL RANTRI(*seed,h,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANTRI is executed.

Range: $seed < 2^{31} - 1$

Note: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

h
is a numeric SAS value.

Range: $0<h<1$

x
is a numeric SAS variable. A new value for the random variate *x* is returned each time CALL RANTRI is executed.

Details

The CALL RANTRI routine updates *seed* and returns a variate *x* generated from a triangular distribution on the interval (0,1) with parameter *h*, which is the modal value of the distribution.

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

The CALL RANTRI routine uses an inverse transform method applied to a RANUNI uniform variate.

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANTRI routine gives greater control of the seed and random number streams than does the RANTRI function.

Examples

This example uses the CALL RANTRI routine:

```
options pageno=1 ls=80 ps=64 nodate;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call rantri(seed, .5, x);
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
      call rantri(seed, .5, x);
      output;
   end;
run;

data all;
   set u1 u2;
   z = rantri(104, .5);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

Output 4.26 Output from the CALL RANTRI Routine

The SAS System 1

Obs	Separate Streams	Single Stream
1	0.34359	0.34359
2	0.76466	0.76466
3	0.54269	0.54269
4	0.89384	0.89384
5	0.72311	0.72311
6	0.68763	0.68763
7	0.48468	0.48468
8	0.38467	0.38467
9	0.29881	0.29881
10	0.80369	0.80369

See Also

Functions:

"RAND Function" on page 1069
"RANTRI Function" on page 1089

CALL RANUNI Routine

Returns a random variate from a uniform distribution.

Category: Random Number

Syntax

CALL RANUNI(*seed,x*);

Arguments

seed
is the seed value. A new value for *seed* is returned each time CALL RANUNI is executed.

Range: $seed < 2^{31} - 1$

Tip: If $seed \leq 0$, the time of day is used to initialize the seed stream.

See: "Seed Values" on page 314 and "Comparison of Seed Values in Random-Number Functions and CALL Routines" on page 319 for more information about seed values

x
is a numeric variable. A new value for the random variate *x* is returned each time CALL RANUNI is executed.

Details

The CALL RANUNI routine updates *seed* and returns a variate *x* that is generated from the uniform distribution on the interval (0,1), using a prime modulus multiplicative generator with modulus 2^{31}–1 and multiplier 397204094 (Fishman and Moore 1982) (See "References" on page 1255).

By adjusting the seeds, you can force streams of variates to agree or disagree for some or all of the observations in the same, or in subsequent, DATA steps.

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANUNI routine gives greater control of the seed and random number streams than does the RANUNI function.

Examples

This example uses the CALL RANUNI routine:

```
options pageno=1 nodate ls=80 ps=64;

data u1(keep=x);
   seed = 104;
   do i = 1 to 5;
      call ranuni(seed, x);
```

```
      output;
   end;
   call symputx('seed', seed);
run;

data u2(keep=x);
   seed = &seed;
   do i = 1 to 5;
      call ranuni(seed, x);
      output;
   end;
run;

data all;
   set u1 u2;
   z = ranuni(104);
run;

proc print label;
   label x = 'Separate Streams' z = 'Single Stream';
run;
```

The following output shows the results:

Output 4.27 Output from the CALL RANUNI Routine

```
                              The SAS System                                1

                       Separate     Single
                  Obs   Streams     Stream

                    1    0.23611    0.23611
                    2    0.88923    0.88923
                    3    0.58173    0.58173
                    4    0.97746    0.97746
                    5    0.84667    0.84667
                    6    0.80484    0.80484
                    7    0.46983    0.46983
                    8    0.29594    0.29594
                    9    0.17858    0.17858
                   10    0.92292    0.92292
```

See Also

Functions:

"RAND Function" on page 1069
"RANUNI Function" on page 1090

CALL SCAN Routine

Returns the position and length of the *n*th word from a character string.

Category: Character

Interaction: When invoked by the %SYSCALL macro statement, CALL SCAN removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL SCAN(<*string*>, *count*, *position*, *length* <, <*charlist*> <, <*modifier(s)*>>>);

Arguments

string
specifies a character constant, variable, or expression.

count
is a non-zero numeric constant, variable, or expression that has an integer value that specifies the number of the word in the character string that you want the CALL SCAN routine to select. For example, a value of 1 indicates the first word, a value of 2 indicates the second word, and so on. The following rules apply:

- If *count* is positive, then CALL SCAN counts words from left to right in the character string.
- If *count* is negative, then CALL SCAN counts words from right to left in the character string.

position
specifies a numeric variable in which the position of the word is returned. If *count* exceeds the number of words in the string, then the value that is returned in *position* is zero. If *count* is zero or missing, then the value that is returned in *position* is missing.

length
specifies a numeric variable in which the length of the word is returned. If *count* exceeds the number of words in the string, then the value that is returned in *length* is zero. If *count* is zero or missing, then the value that is returned in *length* is missing.

charlist
specifies an optional character constant, variable, or expression that initializes a list of characters. This list determines which characters are used as the delimiters that separate words. The following rules apply:

- By default, all characters in *charlist* are used as delimiters.
- If you specify the K modifier in the *modifier* argument, then all characters that are not in *charlist* are used as delimiters.

Tip: You can add more characters to *charlist* by using other modifiers.

modifier
specifies a character constant, variable, or expression in which each non-blank character modifies the action of the CALL SCAN routine. Blanks are ignored. You can use the following characters as modifiers:

a or A	adds alphabetic characters to the list of characters.
b or B	scans backwards, from right to left instead of from left to right, regardless of the sign of the *count* argument.

c or C	adds control characters to the list of characters.
d or D	adds digits to the list of characters.
f or F	adds an underscore and English letters (that is, valid first characters in a SAS variable name using VALIDVARNAME=V7) to the list of characters.
g or G	adds graphic characters to the list of characters. Graphic characters are those that, when printed, produce an image on paper.
h or H	adds a horizontal tab to the list of characters.
i or I	ignores the case of the characters.
k or K	causes all characters that are not in the list of characters to be treated as delimiters. That is, if K is specified, then characters that are in the list of characters are kept in the returned value rather than being omitted because they are delimiters. If K is not specified, then all characters that are in the list of characters are treated as delimiters.
l or L	adds lower case letters to the list of characters.
m or M	specifies that multiple consecutive delimiters, and delimiters at the beginning or end of the *string* argument, refer to words that have a length of zero. If the M modifier is not specified, then multiple consecutive delimiters are treated as one delimiter, and delimiters at the beginning or end of the *string* argument are ignored.
n or N	adds digits, an underscore, and English letters (that is, the characters that can appear in a SAS variable name using VALIDVARNAME=V7) to the list of characters.
o or O	processes the *charlist* and *modifier* arguments only once, rather than every time the CALL SCAN routine is called. **Tip:** Using the O modifier in the DATA step can make CALL SCAN run faster when you call it in a loop where the *charlist* and *modifier* arguments do not change. The O modifier applies separately to each instance of the CALL SCAN routine in your SAS code, and does not cause all instances of the CALL SCAN routine to use the same delimiters and modifiers.
p or P	adds punctuation marks to the list of characters.
q or Q	ignores delimiters that are inside of substrings that are enclosed in quotation marks. If the value of the *string* argument contains unmatched quotation marks, then scanning from left to right will produce different words than scanning from right to left.
s or S	adds space characters to the list of characters (blank, horizontal tab, vertical tab, carriage return, line feed, and form feed).
t or T	trims trailing blanks from the *string* and *charlist* arguments. **Tip:** If you want to remove trailing blanks from just one character argument instead of both character arguments, then use the TRIM function instead of the CALL SCAN routine with the T modifier.
u or U	adds upper case letters to the list of characters.

w or W	adds printable (writable) characters to the list of characters.
x or X	adds hexadecimal characters to the list of characters.

Tip: If the *modifier* argument is a character constant, then enclose it in quotation marks. Specify multiple modifiers in a single set of quotation marks. A *modifier* argument can also be expressed as a character variable or expression.

Details

Definition of "Delimiter" and "Word" A delimiter is any of several characters that are used to separate words. You can specify the delimiters in the *charlist* and *modifier* arguments.

If you specify the Q modifier, then delimiters inside of substrings that are enclosed in quotation marks are ignored.

In the CALL SCAN routine, "word" refers to a substring that has all of the following characteristics:

- is bounded on the left by a delimiter or the beginning of the string
- is bounded on the right by a delimiter or the end of the string
- contains no delimiters

A word can have a length of zero if there are delimiters at the beginning or end of the string, or if the string contains two or more consecutive delimiters. However, the CALL SCAN routine ignores words that have a length of zero unless you specify the M modifier.

Using Default Delimiters in ASCII and EBCDIC Environments If you use the CALL SCAN routine with only four arguments, then the default delimiters depend on whether your computer uses ASCII or EBCDIC characters.

- If your computer uses ASCII characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ^ ¦

  In ASCII environments that do not contain the ^ character, the CALL SCAN routine uses the ~ character instead.
- If your computer uses EBCDIC characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ¬ | ¢ ¦

If you use the *modifier* argument without specifying any characters as delimiters, then the only delimiters that will be used are those that are defined by the *modifier* argument. In this case, the lists of default delimiters for ASCII and EBCDIC environments are not used. In other words, modifiers add to the list of delimiters that are explicitly specified by the *charlist* argument. Modifiers do not add to the list of default modifiers.

Using the CALL SCAN Routine with the M Modifier If you specify the M modifier, then the number of words in a string is defined as one plus the number of delimiters in the string. However, if you specify the Q modifier, delimiters that are inside quotation marks are ignored.

If you specify the M modifier, the CALL SCAN routine returns a positive position and a length of zero if one of the following conditions is true:

- The string begins with a delimiter and you request the first word.
- The string ends with a delimiter and you request the last word.
- The string contains two consecutive delimiters and you request the word that is between the two delimiters.

In you specify a count that is greater in absolute value than the number of words in the string, then the CALL SCAN routine returns a position and length of zero.

Using the CALL SCAN Routine without the M Modifier If you do not specify the M modifier, then the number of words in a string is defined as the number of maximal substrings of consecutive non-delimiters. However, if you specify the Q modifier, delimiters that are inside quotation marks are ignored.

If you do not specify the M modifier, then the CALL SCAN routine does the following:

- ignores delimiters at the beginning or end of the string
- treats two or more consecutive delimiters as if they were a single delimiter

If the string contains no characters other than delimiters, or if you specify a count that is greater in absolute value than the number of words in the string, then the CALL SCAN routine returns a position and length of zero.

Finding the Word as a Character String To find the designated word as a character string after calling the CALL SCAN routine, use the SUBSTRN function with the *string*, *position*, and *length* arguments:

```
substrn(string, position, length);
```

Because CALL SCAN can return a length of zero, using the SUBSTR function can cause an error.

Using Null Arguments The CALL SCAN routine allows character arguments to be null. Null arguments are treated as character strings with a length of zero. Numeric arguments cannot be null.

Examples

Example 1: Scanning for a Word in a String The following example shows how you can use the CALL SCAN routine to find the position and length of a word in a string.

```
options pageno=1 nodate ls=80 ps=64;

data artists;
   input string $60.;
   drop string;
   do i=1 to 99;
      call scan(string, i, position, length);
      if not position then leave;
      Name=substrn(string, position, length);
      output;
   end;
   datalines;
Picasso Toulouse-Lautrec Turner "Van Gogh" Velazquez
;

proc print data=artists;
run;
```

Output 4.28 SAS Output: Scanning for a Word in a String

```
                              The SAS System                                1

                Obs    i    position    length    Name

                 1     1       1          7      Picasso
                 2     2       9          8      Toulouse
                 3     3      18          7      Lautrec
                 4     4      26          6      Turner
                 5     5      33          4      "Van
                 6     6      38          5      Gogh"
                 7     7      44          9      Velazquez
```

Example 2: Finding the First and Last Words in a String The following example scans a string for the first and last words. Note the following:

- A negative count instructs the CALL SCAN routine to scan from right to left.
- Leading and trailing delimiters are ignored because the M modifier is not used.
- In the last observation, all characters in the string are delimiters, so no words are found.

```
options pageno=1 nodate ls=80 ps=64;

data firstlast;
   input String $60.;
   call scan(string, 1, First_Pos, First_Length);
   First_Word = substrn(string, First_Pos, First_Length);
   call scan(string, --1, Last_Pos, Last_Length);
   Last_Word = substrn(string, Last_Pos, Last_Length);
   datalines4;
Jack and Jill
& Bob & Carol & Ted & Alice &
Leonardo
! $ % & ( ) * + , - . / ;
;;;;

proc print data=firstlast;
   var First: Last:;
run;
```

Output 4.29 Results of Finding the First and Last Words in a String

```
                              The SAS System                                1

          First_    First_    First_                    Last_    Last_
    Obs     Pos     Length    Word       Last_Pos     Length   Word

     1        1        4      Jack          10          4      Jill
     2        3        3      Bob           23          5      Alice
     3        1        8      Leonardo       1          8      Leonardo
     4        0        0                     0          0
```

Example 3: Finding All Words in a String without Using the M Modifier The following example scans a string from left to right until no more words are found. Because the M modifier is not used, the CALL SCAN routine does not return any words that have a length of zero. Because blanks are included among the default delimiters, the CALL SCAN routine returns a position or length of zero only when the count exceeds the number of words in the string. The loop can be stopped when the returned position is less than or equal to zero. It is safer to use an inequality comparison to end the loop, rather than to use a strict equality comparison with zero, in case an error causes the position to be missing. (In SAS, a missing value is considered to have a lesser value than any nonmissing value.)

```
options pageno=1 nodate ls=80 ps=64;

data all;
   length word $20;
   drop string;
   string = ' The quick brown fox jumps over the lazy dog.   ';
   do until(position <= 0);
      count+1;
      call scan(string, count, position, length);
      word = substrn(string, position, length);
      output;
   end;
run;

proc print data=all noobs;
   var count position length word;
run;
```

Output 4.30 Results of Finding All Words in a String without Using the M Modifier

```
                          The SAS System                                1

    count    position    length    word

       1          2          3      The
       2          6          5      quick
       3         12          5      brown
       4         18          3      fox
       5         22          5      jumps
       6         28          4      over
       7         33          3      the
       8         37          4      lazy
       9         42          3      dog
      10          0          0
```

Example 4: Finding All Words in a String by Using the M and O Modifiers The following example shows the results of using the M modifier with a comma as a delimiter. With the M modifier, leading, trailing, and multiple consecutive delimiters cause the CALL SCAN routine to return words that have a length of zero.

The O modifier is used for efficiency because the delimiters and modifiers are the same in every call to the CALL SCAN routine.

```
options pageno=1 nodate ls=80 ps=64;

data comma;
```

```
   length word $30;
   string = ',leading,  trailing,and multiple,,delimiters,,';
   do until(position <= 0);
      count + 1;
      call scan(string, count, position, length, ',', 'mo');
      word = substrn(string, position, length);
      output;
   end;
run;

proc print data=comma noobs;
   var count position length word;
run;
```

Output 4.31 Results of Finding All Words in a String by Using the M and O Modifiers

The SAS System 1

count	position	length	word
1	1	0	
2	2	7	leading
3	10	10	trailing
4	21	12	and multiple
5	34	0	
6	35	10	delimiters
7	46	0	
8	47	0	
9	0	0	

Example 5: Using Comma-Separated Values, Substrings in Quotation Marks, and the O Modifier The following example uses the CALL SCAN routine with the O modifier and a comma as a delimiter.

The O modifier is used for efficiency because in each call of the CALL SCAN routine, the delimiters and modifiers do not change.

```
options pageno=1 nodate ls=80 ps=64;

data test;
   length word word_r $30;
   string = 'He said, "She said, ""No!""", not "Yes!"';
   do until(position <= 0);
      count + 1;
      call scan(string, count, position, length, ',', 'oq');
      word = substrn(string, position, length);
      output;
   end;
run;

proc print data=test noobs;
   var count position length word;
run;
```

Output 4.32 Results of Comma-Separated Values and Substrings in Quotation Marks

```
                              The SAS System                                1

              count    position    length    word

                1          1          7      He said
                2          9         20       "She said, ""No!"""
                3         30         11       not "Yes!"
                4          0          0
```

Example 6: Finding Substrings of Digits by Using the D and K Modifiers The following example finds substrings of digits. The *charlist* argument is null, and consequently the list of characters is initially empty. The D modifier adds digits to the list of characters. The K modifier treats all characters that are not in the list as delimiters. Therefore, all characters except digits are delimiters.

```
options pageno=1 nodate ls=80 ps=64;

data digits;
   length digits $20;
   string = 'Call (800) 555--1234 now!';
   do until(position <= 0);
      count+1;
      call scan(string, count, position, length, , 'dko');
      digits = substrn(string, position, length);
      output;
   end;
run;

proc print data=digits noobs;
   var count position length digits;
run;
```

Output 4.33 Results of Finding Substrings of Digits by Using the D and K Modifiers

```
                              The SAS System                                1

                    count    position    length    digits

                      1          7          3       800
                      2         12          3       555
                      3         16          4       1234
                      4          0          0
```

See Also

Function:

"SCAN Function" on page 1111
"FINDW Function" on page 743
"COUNTW Function" on page 621

CALL SET Routine

Links SAS data set variables to DATA step or macro variables that have the same name and data type.

Category: Variable Control

Syntax

CALL SET(*data-set-id*);

Arguments

data-set-id
is the identifier that is assigned by the OPEN function when the data set is opened.

Details

Using SET can significantly reduce the coding that is required for accessing variable values for modification or verification when you use functions to read or to manipulate a SAS file. After a CALL SET, whenever a read is performed from the SAS data set, the values of the corresponding macro or DATA step variables are set to the values of the matching SAS data set variables. If the variable lengths do not match, the values are truncated or padded according to need. If you do not use SET, then you must use the GETVARC and GETVARN functions to move values explicitly between data set variables and macro or DATA step variables.

As a general rule, use CALL SET immediately following OPEN if you want to link the data set and the macro and DATA step variables.

Examples

This example uses the CALL SET routine:

- The following statements automatically set the values of the macro variables PRICE and STYLE when an observation is fetched:

```
%macro setvar;
   %let dsid=%sysfunc(open(sasuser.houses,i));
      /* No leading ampersand with %SYSCALL */
   %syscall set(dsid);
   %let rc=%sysfunc(fetchobs(&dsid,10));
   %let rc=%sysfunc(close(&dsid));
%mend setvar;

%global price style;
%setvar
%put _global_;
```

- The %PUT statement writes these lines to the SAS log:

```
GLOBAL PRICE 127150
GLOBAL STYLE CONDO
```

□ The following statements obtain the values for the first 10 observations in SASUSER.HOUSES and store them in MYDATA:

```
data mydata;
      /* create variables for assignment */
      /*by CALL SET */
   length style $8 sqfeet bedrooms baths 8
      street $16 price 8;
   drop rc dsid;
   dsid=open("sasuser.houses","i");
   call set (dsid);
   do i=1 to 10;
      rc=fetchobs(dsid,i);
      output;
   end;
run;
```

See Also

Functions:

"FETCH Function" on page 684
"FETCHOBS Function" on page 685
"GETVARC Function" on page 794
"GETVARN Function" on page 795

CALL SLEEP Routine

For a specified period of time, suspends the execution of a program that invokes this CALL routine.

Category: Special

See: CALL SLEEP Routine in the documentation for your operating environment.

Syntax

CALL SLEEP(*n*<, *unit*>)

Arguments

n
is a numeric constant that specifies the number of units of time for which you want to suspend execution of a program.

Range: $n \geq 0$

unit
specifies the unit of time, as a power of 10, which is applied to *n*. For example, 1 corresponds to a second, and .001 corresponds to a millisecond.

Default: .001

Details

The CALL SLEEP routine suspends the execution of a program that invokes this call routine for a period of time that you specify. The program can be a DATA step, macro, IML, SCL, or anything that can invoke a call routine. The maximum sleep period for the CALL SLEEP routine is 46 days.

Examples

Example 1: Suspending Execution for a Specified Period of Time The following example tells SAS to suspend the execution of the DATA step PAYROLL for 1 minute and 10 seconds:

```
data payroll;
   call sleep(7000,.01);
   ...more SAS statements...
run;
```

Example 2: Suspending Execution Based on a Calculation of Sleep Time The following example tells SAS to suspend the execution of the DATA step BUDGET until March 1, 2004, at 3:00 AM. SAS calculates the length of the suspension based on the target date and the date and time that the DATA step begins to execute.

```
data budget;
   sleeptime='01mar2004:03:00'dt-datetime();
   call sleep(sleeptime,1);
   ...more SAS statements...;
run;
```

See Also

Functions:

"SLEEP Function" on page 1127

CALL SOFTMAX Routine

Returns the softmax value.

Category: Mathematical

Syntax

CALL SOFTMAX(*argument*<,*argument*,...>);

Arguments

argument
is numeric.

Restriction: The CALL SOFTMAX routine only accepts variables as valid arguments. Do not use a constant or a SAS expression because the CALL routine is unable to update these arguments.

Details

The CALL SOFTMAX routine replaces each argument with the softmax value of that argument. For example x_j is replaced by

$$\frac{e^{x_j}}{\sum_{i=1}^{i=n} e^{x_i}}$$

If any argument contains a missing value, then CALL SOFTMAX returns missing values for all the arguments. Upon a successful return, the sum of all the values is equal to 1.

Examples

The following SAS statements produce these results:

SAS Statements	Results

`x=0.5;`
`y=-0.5;`
`z=1;`
`call softmax(x,y,z);`
`put x= y= z=;`
	`x=0.3314989604 y=0.1219516523 z=0.5465493873`

CALL SORTC Routine

Sorts the values of character arguments.

Category: Sort

Interaction: When invoked by the %SYSCALL macro statement, CALL SORTC removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL SORTC(*variable-1*<, ..., *variable-n*>)

Arguments

variable
specifies a character variable.

Details

The values of *variable* are sorted in ascending order by the CALL SORTC routine.

Comparisons

The CALL SORTC routine is used with character variables, and the CALL SORTN routine is used with numeric variables.

Examples

The following example sorts the character variables in the array in ascending order.

```
data _null_;
   array x(8) $10
      ('tweedledum' 'tweedledee' 'baboon' 'baby'
       'humpty' 'dumpty' 'banana' 'babylon');
   call sortc(of x(*));
   put +3 x(*);
run;
```

SAS writes the following output to the log:

```
   baboon baby babylon banana dumpty humpty tweedledee tweedledum
```

See Also

Functions and CALL routines:
"CALL SORTN Routine" on page 530

CALL SORTN Routine

Sorts the values of numeric arguments.

Category: Sort

Syntax

CALL SORTN(*variable-1*<, ..., *variable-n*>)

Arguments

variable
specifies a numeric variable.

Details

The values of *variable* are sorted in ascending order by the CALL SORTN routine.

Comparisons

The CALL SORTN routine is used with numeric variables, and the CALL SORTC routine is used with character variables.

Examples

The following example sorts the numeric variables in the array in ascending order.

```
data _null_;
   array x(10) (0, ., .a, 1e-12, -1e-8, .z, -37, 123456789, 1e20, 42);
   call sortn(of x(*));
   put +3 x(*);
run;
```

SAS writes the following output to the log:

```
   . A Z -37 -1E-8 0 1E-12 42 123456789 1E20
```

See Also

Functions and CALL routines:

"CALL SORTC Routine" on page 528

CALL STDIZE Routine

Standardizes the values of one or more variables.

Category: Mathematical

Interaction: When invoked by the %SYSCALL macro statement, CALL STDIZE removes the quotation marks from its arguments. For more information, see "Using CALL Routines and the %SYSCALL Macro Statement" on page 312.

Syntax

CALL STDIZE(<*option-1, option-2, …*,>*variable-1*<,*variable-2, …*>);

Arguments

option
specifies a character expression whose values can be uppercase, lowercase, or mixed case letters. Leading and trailing blanks are ignored.

Restriction: Use a separate argument for each option because you cannot specify more than one option in a single argument.

Tip: Character expressions can end with an equal sign that is followed by another argument that is a numeric constant, variable, or expression.

See Also: PROC STDIZE in *SAS/STAT User's Guide* for information about formulas and other details. The options that are used in CALL STDIZE are the same as those used in PROC STDIZE.

option includes the following three categories:

standardization-options
specify how to compute the location and scale measures that are used to standardize the variables. The following standardization options are available:

ABW=
must be followed by an argument that is a numeric expression specifying the tuning constant.

AGK=
must be followed by an argument that is a numeric expression that specifies the proportion of pairs to be included in the estimation of the within-cluster variances.

AHUBER=
must be followed by an argument that is a numeric expression specifying the tuning constant.

AWAVE=
must be followed by an argument that is a numeric expression specifying the tuning constant.

EUCLEN
specifies the Euclidean length.

IQR
specifies the interquartile range.

L=
must be followed by an argument that is a numeric expression with a value greater than or equal to 1 specifying the power to which differences are to be raised in computing an L(p) or Minkowski metric.

MAD
specifies the median absolute deviation from the median.

MAXABS
specifies the maximum absolute values.

MEAN
specifies the arithmetic mean (average).

MEDIAN
specifies the middle number in a set of data that is ordered according to rank.

MIDRANGE
specifies the midpoint of the range.

RANGE
specifies a range of values.

SPACING=
must be followed by an argument that is a numeric expression that specifies the proportion of data to be contained in the spacing.

STD
specifies the standard deviation.

SUM
specifies the result that you obtain when you add numbers.

USTD
specifies the standard deviation about the origin, based on the uncorrected sum of squares.

VARDEF-options
specify the divisor to be used in the calculation of variances. VARDEF options can have the following values:

DF
specifies degrees of freedom.

N
specifies the number of observations.

Default: DF

miscellaneous-options
Miscellaneous options can have the following values:

ADD=
is followed by a numeric argument that specifies a number to add to each value after standardizing and multiplying by the value from the MULT= option. The default value is 0.

FUZZ=
is followed by a numeric argument that specifies the relative fuzz factor.

MISSING=
is followed by a numeric argument that specifies a value to be assigned to variables that have a missing value.

MULT=
is followed by a numeric argument that specifies a number by which to multiply each value after standardizing. The default value is 1.

NORM
normalizes the scale estimator to be consistent for the standard deviation of a normal distribution. This option affects only the methods AGK=, IQR, MAD, and SPACING=.

PSTAT
writes the values of the location and scale measures in the log.

REPLACE
replaces missing values with the value 0 in the standardized data (this value corresponds to the location measure before standardizing). To replace missing values by other values, see the MISSING= option.

SNORM
normalizes the scale estimator to have an expectation of approximately 1 for a standard normal distribution.

Tip: This option affects only the SPACING= method.

variable
is numeric. These values will be standardized according to the method that you use.

Details

The CALL STDIZE routine transforms one or more arguments that are numeric variables by subtracting a location measure and dividing by a scale measure. You can use a variety of location and scale measures. The default location option is MEAN, and the default scale option is STD.

In addition, you can multiply each standardized value by a constant and you can add a constant. The final output value would be

$$result = add + mult * \left(\frac{(original - location)}{scale} \right)$$

where

result	specifies the final value that is returned for each variable.
add	specifies the constant to add (ADD= option).
mult	specifies the constant to multiply by (MULT= option).
original	specifies the original input value.
location	specifies the location measure.
scale	specifies the scale measure.

You can replace missing values by any constant. If you do not specify the MISSING= or the REPLACE option, variables that have missing values are not altered. The initial estimation method for the ABW=, AHUBER=, and AWAVE= methods is MAD. Percentiles are computed using definition 5. For more information about percentile calculations, see "SAS Elementary Statistics Procedures" in *Base SAS Procedures Guide*.

Comparisons

The CALL STDIZE routine is similar to the STDIZE procedure in the SAS/STAT product. However, the CALL STDIZE routine is primarily useful for standardizing the rows of a SAS data set, whereas the STDIZE procedure can standardize only the columns of a SAS data set. For more information, see PROC STDIZE in *SAS/STAT User's Guide*.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`retain x 1 y 2 z 3;`
`call stdize(x,y,z);`
`put x= y= z=;`
	`x=-1 y=0 z=1`

`retain w 10 x 11 y 12 z 13;`
`call stdize('iqr',w,x,y,z);`
`put w= x= y= z=;`
	`w=-0.75 x=-0.25 y=0.25 z=0.75`

`retain w . x 1 y 2 z 3;`
`call stdize('range',w,x,y,z);`
`put w= x= y= z=;`
	`w=. x=0 y=0.5 z=1`

`retain w . x 1 y 2 z 3;`
`call stdize('mult=',10,'missing=',`
`-1,'range',w,x,y,z);`
`put w= x= y= z=;`
	`w=-1 x=0 y=5 z=10`

CALL STREAMINIT Routine

Specifies a seed value to use for subsequent random number generation by the RAND function.

Category: Random Number

Syntax

CALL STREAMINIT(*seed*);

Arguments

seed
is an integer seed value.

Range: $seed < 2^{31} - 1$

Tip: If you specify a nonpositive seed, then CALL STREAMINIT is ignored. Any subsequent random number generation seeds itself from the system clock.

Details

If you want to create reproducible streams of random numbers, then specify CALL STREAMINIT before any calls to the RAND random number function. If you call the RAND function before you specify a seed with the CALL STREAMINIT routine (or if you specify a nonpositive seed value in the CALL STREAMINIT routine), then the RAND function uses a call to the system clock to seed itself. For more information about seed values see "Seed Values" on page 314.

Examples

Example 1: Creating a Reproducible Stream of Random Numbers The following example shows how to specify a seed value with CALL STREAMINIT to create a reproducible stream of random numbers with the RAND function.

```
options nodate ps=60 ls=80 pageno=1;

data random;
   call streaminit(123);
   do i=1 to 10;
      x1=rand('cauchy');
      output;
   end;

proc print data=random;
   id i;
run;
```

Output 4.34 Number String Seeded with CALL STREAMINIT

```
                                The SAS System                                1

 i       x1

 1    -0.17593
 2     3.76106
 3     1.23427
 4     0.49095
 5    -0.05094
 6     0.72496
 7    -0.51646
 8     7.61304
 9     0.89784
10     1.69348
```

See Also

Functions:

"RAND Function" on page 1069

CALL SYMPUT Routine

Assigns DATA step information to a macro variable.

Category: Macro

Syntax

CALL SYMPUT(*argument-1*,*argument-2*);

Arguments

argument-1
specifies a character expression that identifies the macro variable that is assigned a value. If the macro variable does not exist, the routine creates it.

argument-2
specifies a character constant, variable, or expression that contains the value that is assigned.

Details

The CALL SYMPUT routine either creates a macro variable whose value is information from the DATA step or assigns a DATA step value to an existing macro variable. CALL SYMPUT is fully documented in "SYMPUT Routine" in *SAS Macro Language: Reference*.

See Also

Function:

"SYMGET Function" on page 1151

CALL SYMPUTX Routine

Assigns a value to a macro variable, and removes both leading and trailing blanks.

Category: Macro

Syntax

CALL SYMPUTX(*macro-variable*, *value* <,*symbol-table*>);

Arguments

macro-variable

can be one of the following:

- a character string that is a SAS name, enclosed in quotation marks.
- the name of a character variable whose values are SAS names.
- a character expression that produces a macro variable name. This form is useful for creating a series of macro variables.

a character constant, variable, or expression. Leading and trailing blanks are removed from the value of *name*, and the result is then used as the name of the macro variable.

value

specifies a character or numeric constant, variable, or expression. If *value* is numeric, SAS converts the value to a character string using the BEST. format and does not issue a note to the SAS log. Leading and trailing blanks are removed, and the resulting character string is assigned to the macro variable.

symbol-table

specifies a character constant, variable, or expression. The value of *symbol-table* is not case sensitive. The first non-blank character in *symbol-table* specifies the symbol table in which to store the macro variable. The following values are valid as the first non-blank character in *symbol-table*:

G	specifies that the macro variable is stored in the global symbol table, even if the local symbol table exists.
L	specifies that the macro variable is stored in the most local symbol table that exists, which will be the global symbol table, if used outside a macro.
F	specifies that if the macro variable exists in any symbol table, CALL SYMPUTX uses the version in the most local symbol table in which it exists. If the macro variable does not exist, CALL SYMPUTX stores the variable in the most local symbol table.

Note: If you omit *symbol-table*, or if *symbol-table* is blank, CALL SYMPUTX stores the macro variable in the same symbol table as does the CALL SYMPUT routine. △

Comparisons

CALL SYMPUTX is similar to CALL SYMPUT except that

- CALL SYMPUTX does not write a note to the SAS log when the second argument is numeric. CALL SYMPUT, however, writes a note to the log stating that numeric values were converted to character values.
- CALL SYMPUTX uses a field width of up to 32 characters when it converts a numeric second argument to a character value. CALL SYMPUT uses a field width of up to 12 characters.
- CALL SYMPUTX left-justifies both arguments and trims trailing blanks. CALL SYMPUT does not left-justify the arguments, and trims trailing blanks from the first argument only. Leading blanks in the value of *name* cause an error.
- CALL SYMPUTX enables you to specify the symbol table in which to store the macro variable, whereas CALL SYMPUT does not.

Examples

The following example shows the results of using CALL SYMPUTX.

```
data _null_;
   call symputx('  items  ', '   leading and trailing blanks removed   ',
                'lplace');
   call symputx('  x   ', 123.456);
run;

%put items=!&items!;
%put x=!&x!;
```

The following lines are written to the SAS log:

```
----+----1----+----2----+----3----+----4----+----5
items=!leading and trailing blanks removed!
x=!123.456!
```

See Also

Functions and CALL Routines:

SYMGET Function“SYMGET Function” on page 1151

CALL SYMPUT“CALL SYMPUT Routine” on page 536

CALL SYSTEM Routine

Submits an operating environment command for execution.

Category: Special

Interaction: When invoked by the %SYSCALL macro statement, CALL SYSTEM removes quotation marks from its arguments. For more information, see “Using CALL Routines and the %SYSCALL Macro Statement” on page 312.

See: CALL SYSTEM Routine in the documentation for your operating environment.

Syntax

CALL SYSTEM(*command*);

Arguments

command
specifies any of the following: a system command that is enclosed in quotation marks (character string), an expression whose value is a system command, or the name of a character variable whose value is a system command that is executed.

Operating Environment Information: See the SAS documentation for your operating environment for information about what you can specify. △

Restriction: The length of the command cannot be greater than 1024 characters, including trailing blanks.

Comparisons

The behavior of the CALL SYSTEM routine is similar to that of the X command, the X statement, and the SYSTEM function. It is useful in certain situations because it can be conditionally executed, it accepts an expression as an argument, and it is executed at run time.

See Also

Function:

"SYSTEM Function" on page 1159

CALL TANH Routine

Returns the hyperbolic tangent.

Category: Mathematical

Syntax

CALL TANH(*argument*<, *argument,...*>);

Arguments

argument
is numeric.

Restriction: The CALL TANH routine only accepts variables as valid arguments. Do not use a constant or a SAS expression, because the CALL routine is unable to update these arguments.

Details

The subroutine TANH replaces each argument by the tanh of that argument. For example x_j is replaced by

$$tanh\,(x_j) = \frac{e^{x_j} - e^{-x_j}}{e^{x_j} + e^{-x_j}}$$

If any argument contains a missing value, then CALL TANH returns missing values for all the arguments.

Examples

The following SAS statements produce these results:

SAS Statements	Results

`x=0.5;`
`y=-0.5;`
`call tanh(x,y);`
`put x= y=;`
	`x=0.4621171573 y=-0.462117157`

See Also

Function:

"TANH Function" on page 1161

CALL VNAME Routine

Assigns a variable name as the value of a specified variable.

Category: Variable Control

Syntax

CALL VNAME(*variable-1*,*variable-2*);

Arguments

variable-1
specifies any SAS variable.

variable-2
specifies any SAS character variable. Because SAS variable names can contain up to 32 characters, the length of *variable-2* should be at least 32.

Details

The CALL VNAME routine assigns the name of the *variable-1* variable as the value of the *variable-2* variable.

Examples

This example uses the CALL VNAME routine with array references to return the names of all variables in the data set OLD:

```
data new(keep=name);
   set old;
      /* all character variables in old */
   array abc{*} _character_;
      /* all numeric variables in old */
   array def{*} _numeric_;
      /* name is not in either array */
   length name $32;
   do i=1 to dim(abc);
         /* get name of character variable */
      call vname(abc{i},name);
         /* write name to an observation */
      output;
   end;
   do j=1 to dim(def);
         /* get name of numeric variable */
      call vname(def{j},name);
         /* write name to an observation */
      output;
   end;
   stop;
run;
```

See Also

Functions:

"VNAME Function" on page 1219
"VNAMEX Function" on page 1220

CALL VNEXT Routine

Returns the name, type, and length of a variable that is used in a DATA step.

Category: Variable Information

Syntax

CALL VNEXT(*varname* <,*vartype* <, *varlength*>>);

Arguments

varname
is a character variable that is updated by the CALL VNEXT routine. The following rules apply:

- If the input value of *varname* is blank, the value that is returned in *varname* is the name of the first variable in the DATA step's list of variables.
- If the CALL VNEXT routine is executed for the first time in the DATA step, the value that is returned in *varname* is the name of the first variable in the DATA step's list of variables.

If neither of the above conditions exists, the input value of *varname* is ignored. Each time the CALL VNEXT routine is executed, the value that is returned in *varname* is the name of the next variable in the list.

After the names of all the variables in the DATA step are returned, the value that is returned in *varname* is blank.

vartype
is a character variable whose input value is ignored. The value that is returned is "N" or "C." The following rules apply:

- If the value that is returned in *varname* is the name of a numeric variable, the value that is returned in *vartype* is "N."
- If the value that is returned in *varname* is the name of a character variable, the value that is returned in *vartype* is "C."
- If the value that is returned in *varname* is blank, the value that is returned in *vartype* is also blank.

varlength
is a numeric variable. The input value of *varlength* is ignored.

The value that is returned is the length of the variable whose name is returned in *varname*. If the value that is returned in *varname* is blank, the value that is returned in *varlength* is zero.

Details

The variable names that are returned by the CALL VNEXT routine include automatic variables such as _N_ and _ERROR_. If the DATA step contains a BY statement, the variable names that are returned by CALL VNEXT include the FIRST.*variable* and LAST.*variable* names. CALL VNEXT also returns the names of the variables that are used as arguments to CALL VNEXT.

Note: The order in which variable names are returned by CALL VNEXT can vary in different releases of SAS and in different operating environments. △

Examples

The following example shows the results from using the CALL VNEXT routine.

```
data test;
   x=1;
   y='abc';
   z=.;
   length z 5;
run;

data attributes;
   set test;
   by x;
   input a b $ c;
   length name $32 type $3;
   name=' ';
   length=666;
   do i=1 to 99 until(name=' ');
      call vnext(name,type,length);
      put i= name @40 type= length=;
   end;
   this_is_a_long_variable_name=0;
   datalines;
1 q 3
;
```

Output 4.35 Partial SAS Log Output for the CALL VNEXT Routine

```
i=1 x                                    type=N length=8
i=2 y                                    type=C length=3
i=3 z                                    type=N length=5
i=4 FIRST.x                              type=N length=8
i=5 LAST.x                               type=N length=8
i=6 a                                    type=N length=8
i=7 b                                    type=C length=8
i=8 c                                    type=N length=8
i=9 name                                 type=C length=32
i=10 type                                type=C length=3
i=11 length                              type=N length=8
i=12 i                                   type=N length=8
i=13 this_is_a_long_variable_name        type=N length=8
i=14 _ERROR_                             type=N length=8
i=15 _N_                                 type=N length=8
i=16                                     type=  length=0
```

CAT Function

Does not remove leading or trailing blanks, and returns a concatenated character string.

Category: Character

Restriction: "I18N Level 2" on page 314

Tip: DBCS equivalent function is KSTRCAT in *SAS National Language Support (NLS): Reference Guide*.

Syntax

CAT(*item-1* <, …, *item-n*>)

Arguments

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string by using the BEST*w*. format. In this case, leading blanks are removed and SAS does not write a note to the log.

Details

Length of Returned Variable

In a DATA step, if the CAT function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes. If the concatenation operator (||) returns a value to a variable that has not previously been assigned a length, then that variable is given a length that is the sum of the lengths of the values which are being concatenated.

Length of Returned Variable: Special Cases The CAT function returns a value to a variable, or returns a value in a temporary buffer. The value that is returned from the CAT function has the following length:

- up to 200 characters in WHERE clauses and in PROC SQL
- up to 32767 characters in the DATA step except in WHERE clauses
- up to 65534 characters when CAT is called from the macro processor

If CAT returns a value in a temporary buffer, the length of the buffer depends on the calling environment, and the value in the buffer can be truncated after CAT finishes processing. In this case, SAS does not write a message about the truncation to the log.

If the length of the variable or the buffer is not large enough to contain the result of the concatenation, SAS does the following:

- changes the result to a blank value in the DATA step, and in PROC SQL
- writes a warning message to the log stating that the result was either truncated or set to a blank value, depending on the calling environment
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation
- sets _ERROR_ to 1 in the DATA step

The CAT function removes leading and trailing blanks from numeric arguments after it formats the numeric value with the BEST*w*. format.

Comparisons

The results of the CAT, CATS, CATT, and CATX functions are *usually* equivalent to results that are produced by certain combinations of the concatenation operator (||) and the TRIM and LEFT functions. However, the default length for the CAT, CATS,

CATT, and CATX functions is different from the length that is obtained when you use the concatenation operator. For more information, see “Length of Returned Variable” on page 544.

Using the CAT, CATS, CATT, and CATX functions is faster than using TRIM and LEFT, and you can use them with the OF syntax for variable lists in calling environments that support variable lists.

The following table shows equivalents of the CAT, CATS, CATT, and CATX functions. The variables X1 through X4 specify character variables, and SP specifies a delimiter, such as a blank or comma.

Function	Equivalent Code
CAT(OF X1-X4)	**X1||X2||X3||X4**
CATS(OF X1-X4)	**TRIM(LEFT(X1))||TRIM(LEFT(X2))||TRIM(LEFT(X3))|| TRIM(LEFT(X4))**
CATT(OF X1-X4)	**TRIM(X1)||TRIM(X2)||TRIM(X3)||TRIM(X4)**
CATX(SP, OF X1-X4)	**TRIM(LEFT(X1))||SP||TRIM(LEFT(X2))||SP|| TRIM(LEFT(X3))||SP||TRIM(LEFT(X4))**

Examples

The following example shows how the CAT function concatenates strings.

```
data _null_;
   x='  The 2002 Olym';
   y='pic Arts Festi';
   z='  val included works by D  ';
   a='ale Chihuly.';
   result=cat(x,y,z,a);
   put result $char.;
run;
```

SAS writes the following line to the log:

```
   ----+----1----+----2----+----3----+----4----+----5----+----6----+----7
     The 2002 Olympic Arts Festi  val included works by D  ale Chihuly.
```

See Also

Functions and CALL Routines:

“CALL CATS Routine” on page 441
“CALL CATT Routine” on page 443
“CALL CATX Routine” on page 445
“CATQ Function” on page 546
“CATS Function” on page 550
“CATT Function” on page 552
“CATX Function” on page 554

CATQ Function

Concatenates character or numeric values by using a delimiter to separate items and by adding quotation marks to strings that contain the delimiter.

Category: Character

Syntax

CATQ(*modifiers*<, *delimiter*>, *item-1* <, ..., *item-n*>)

Arguments

modifier
specifies a character constant, variable, or expression in which each non-blank character modifies the action of the CATQ function. Blanks are ignored. You can use the following characters as modifiers:

1 or ’	uses single quotation marks when CATQ adds quotation marks to a string.
2 or ”	uses double quotation marks when CATQ adds quotation marks to a string.
a or A	adds quotation marks to all of the item arguments.
b or B	adds quotation marks to item arguments that have leading or trailing blanks that are not removed by the S or T modifiers.
c or C	uses a comma as a delimiter.
d or D	indicates that you have specified the delimiter argument.
h or H	uses a horizontal tab as the delimiter.
m or M	inserts a delimiter for every item argument after the first. If you do not use the M modifier, then CATQ does not insert delimiters for item arguments that have a length of zero after processing that is specified by other modifiers. The M modifier can cause delimiters to appear at the beginning or end of the result and can cause multiple consecutive delimiters to appear in the result.
n or N	converts item arguments to name literals when the value does not conform to the usual syntactic conventions for a SAS name. A name literal is a string in quotation marks that is followed by the letter “n” without any intervening blanks. To use name literals in SAS statements, you must specify the SAS option, VALIDVARNAME=ANY.
q or Q	adds quotation marks to item arguments that already contain quotation marks.
s or S	strips leading and trailing blanks from subsequently processed arguments: □ To strip leading and trailing blanks from the delimiter argument, specify the S modifier *before* the D modifier.

- To strip leading and trailing blanks from the item arguments but *not* from the delimiter argument, specify the S modifier *after* the D modifier.

t or T trims trailing blanks from subsequently processed arguments:

- To trim trailing blanks from the delimiter argument, specify the T modifier *before* the D modifier.
- To trim trailing blanks from the item arguments but *not* from the delimiter argument, specify the T modifier *after* the D modifier.

x or X converts item arguments to hexadecimal literals when the value contains nonprintable characters.

Tip: If *modifier* is a constant, enclose it in quotation marks. You can also express *modifier* as a variable name or an expression.

Tip: The A, B, N, Q, S, T, and X modifiers operate internally to the CATQ function. If an item argument is a variable, then the value of that variable is not changed by CATQ unless the result is assigned to that variable.

delimiter
specifies a character constant, variable, or expression that is used as a delimiter between concatenated strings. If you specify this argument, then you must also specify the D modifier.

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string by using the BEST*w*. format. In this case, leading blanks are removed and SAS does not write a note to the log.

Details

Length of Returned Variable The CATQ function returns a value to a variable or if CATQ is called inside an expression, CATQ returns a value to a temporary buffer. The value that is returned has the following length:

- up to 200 characters in WHERE clauses and in PROC SQL
- up to 32767 characters in the DATA step except in WHERE clauses
- up to 65534 characters when CATQ is called from the macro processor

If the length of the variable or the buffer is not large enough to contain the result of the concatenation, then SAS does the following steps:

- changes the result to a blank value in the DATA step and in PROC SQL
- writes a warning message to the log stating that the result was either truncated or set to a blank value, depending on the calling environment
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation
- sets _ERROR_ to 1 in the DATA step

If CATQ returns a value in a temporary buffer, then the length of the buffer depends on the calling environment, and the value in the buffer can be truncated after CATQ finishes processing. In this case, SAS does not write a message about the truncation to the log.

The Basics If you do not use the C, D, or H modifiers, then CATQ uses a blank as a delimiter.

If you specify neither a quotation mark in *modifier* nor the 1 or 2 modifiers, then CATQ decides independently for each item argument which type of quotation mark to use, if quotation marks are required. The following rules apply:

- CATQ uses single quotation marks for strings that contain an ampersand (&) or percent (%) sign, or that contain more double quotation marks than single quotation marks.
- CATQ uses double quotation marks for all other strings.

The CATQ function initializes the result to a length of zero and then performs the following actions for each item argument:

1. If *item* is not a character string, then CATQ converts *item* to a character string by using the BEST*w*. format and removes leading blanks.
2. If you used the S modifier, then CATQ removes leading blanks from the string.
3. If you used the S or T modifiers, then CATQ removes trailing blanks from the string.
4. CATQ determines whether to add quotation marks based on the following conditions:
   - If you use the X modifier and the string contains control characters, then the string is converted to a hexadecimal literal.
   - If you use the N modifier, then the string is converted to a name literal if either of the following conditions is true:
     - The first character in the string is not an underscore or an English letter.
     - The string contains any character that is not a digit, underscore, or English letter.
   - If you did not use the X or the N modifiers, then CATQ adds quotation marks to the string if any of the following conditions is true:
     - You used the A modifier.
     - You used the B modifier and the string contains leading or trailing blanks that were not removed by the S or T modifiers.
     - You used the Q modifier and the string contains quotation marks.
     - The string contains a substring that equals the delimiter with leading and trailing blanks omitted.
5. For the second and subsequent item arguments, CATQ appends the delimiter to the result if either of the following conditions is true:
   - You used the M modifier.
   - The string has a length greater than zero after it has been processed by the preceding steps.
6. CATQ appends the string to the result.

Comparisons

The CATX function is similar to the CATQ function except that CATX does not add quotation marks.

Examples

The following example shows how the CATQ function concatenates strings.

```
options ls=110;

data _null_;
   result1=CATQ(' ',
                'noblanks',
                'one blank',
                12345,
                '  lots of blanks    ');
   result2=CATQ('CS',
                'Period (.)                ',
                'Ampersand (&)             ',
                'Comma (,)                 ',
                'Double quotation marks (") ',
                '   Leading Blanks');
   result3=CATQ('BCQT',
                'Period (.)                ',
                'Ampersand (&)             ',
                'Comma (,)                 ',
                'Double quotation marks (") ',
                '   Leading Blanks');
   result4=CATQ('ADT',
                '#=#',
                'Period (.)                ',
                'Ampersand (&)             ',
                'Comma (,)                 ',
                'Double quotation marks (") ',
                '   Leading Blanks');
   result5=CATQ('N',
                'ABC_123   ',
                '123    ',
                'ABC 123');
   put (result1-result5) (=/);
run;
```

SAS writes the following output to the log.

```
result1=noblanks "one blank" 12345 "  lots of blanks    "
result2=Period (.),Ampersand (&),"Comma (,)",Double quotation marks ("),Leading Blanks
result3=Period (.),Ampersand (&),"Comma (,)",'Double quotation marks (")',"   Leading Blanks"
result4="Period (.)"#=#'Ampersand (&)'#=#"Comma (,)"#=#'Double quotation marks (")'#=#"   Leading Blanks"
result5=ABC_123 "123"n "ABC 123"n
```

See Also

Functions and CALL Routines:

"CALL CATS Routine" on page 441
"CALL CATT Routine" on page 443
"CALL CATX Routine" on page 445
"CAT Function" on page 543
"CATS Function" on page 550

"CATT Function" on page 552
"CATX Function" on page 554

CATS Function

Removes leading and trailing blanks, and returns a concatenated character string.

Category: Character
Restriction: "I18N Level 0" on page 313

Syntax

CATS(*item-1* <, …, *item-n*>)

Arguments

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string by using the BEST*w*. format. In this case, SAS does not write a note to the log.

Details

Length of Returned Variable

In a DATA step, if the CATS function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes. If the concatenation operator (||) returns a value to a variable that has not previously been assigned a length, then that variable is given a length that is the sum of the lengths of the values which are being concatenated.

Length of Returned Variable: Special Cases The CATS function returns a value to a variable, or returns a value in a temporary buffer. The value that is returned from the CATS function has the following length:

- up to 200 characters in WHERE clauses and in PROC SQL
- up to 32767 characters in the DATA step except in WHERE clauses
- up to 65534 characters when CATS is called from the macro processor

If CATS returns a value in a temporary buffer, the length of the buffer depends on the calling environment, and the value in the buffer can be truncated after CATS finishes processing. In this case, SAS does not write a message about the truncation to the log.

If the length of the variable or the buffer is not large enough to contain the result of the concatenation, SAS does the following:

- changes the result to a blank value in the DATA step, and in PROC SQL
- writes a warning message to the log stating that the result was either truncated or set to a blank value, depending on the calling environment
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation

□ sets _ERROR_ to 1 in the DATA step

The CATS function removes leading and trailing blanks from numeric arguments after it formats the numeric value with the BEST*w*. format.

Comparisons

The results of the CAT, CATS, CATT, and CATX functions are *usually* equivalent to results that are produced by certain combinations of the concatenation operator (||) and the TRIM and LEFT functions. However, the default length for the CAT, CATS, CATT, and CATX functions is different from the length that is obtained when you use the concatenation operator. For more information, see "Length of Returned Variable" on page 550.

Using the CAT, CATS, CATT, and CATX functions is faster than using TRIM and LEFT, and you can use them with the OF syntax for variable lists in calling environments that support variable lists.

The following table shows equivalents of the CAT, CATS, CATT, and CATX functions. The variables X1 through X4 specify character variables, and SP specifies a delimiter, such as a blank or comma.

Function	Equivalent Code
CAT(OF X1-X4)	**X1||X2||X3||X4**
CATS(OF X1-X4)	**TRIM(LEFT(X1))||TRIM(LEFT(X2))||TRIM(LEFT(X3))|| TRIM(LEFT(X4))**
CATT(OF X1-X4)	**TRIM(X1)||TRIM(X2)||TRIM(X3)||TRIM(X4)**
CATX(SP, OF X1-X4)	**TRIM(LEFT(X1))||SP||TRIM(LEFT(X2))||SP|| TRIM(LEFT(X3))||SP||TRIM(LEFT(X4))**

Examples

The following example shows how the CATS function concatenates strings.

```
data _null_;
   x='  The   Olym';
   y='pic Arts Festi';
   z='  val includes works by D  ';
   a='ale Chihuly.';
   result=cats(x,y,z,a);
   put result $char.;
run;
```

The following line is written to the SAS log:

```
   ----+----1----+----2----+----3----+----4----+----5----+----6
   The   Olympic Arts Festival includes works by Dale Chihuly.
```

See Also

Functions and CALL Routines:

"CALL CATS Routine" on page 441
"CALL CATT Routine" on page 443
"CALL CATX Routine" on page 445
"CAT Function" on page 543
"CATQ Function" on page 546
"CATT Function" on page 552
"CATX Function" on page 554

CATT Function

Removes trailing blanks, and returns a concatenated character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

CATT(*item-1* <, … *item-n*>)

Arguments

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string by using the BEST*w*. format. In this case, leading blanks are removed and SAS does not write a note to the log.

Details

Length of Returned Variable

In a DATA step, if the CATT function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes. If the concatenation operator (||) returns a value to a variable that has not previously been assigned a length, then that variable is given a length that is the sum of the lengths of the values which are being concatenated.

Length of Returned Variable: Special Cases The CATT function returns a value to a variable, or returns a value in a temporary buffer. The value that is returned from the CATT function has the following length:

- up to 200 characters in WHERE clauses and in PROC SQL
- up to 32767 characters in the DATA step except in WHERE clauses
- up to 65534 characters when CATT is called from the macro processor

If CATT returns a value in a temporary buffer, the length of the buffer depends on the calling environment, and the value in the buffer can be truncated after CATT finishes processing. In this case, SAS does not write a message about the truncation to the log.

If the length of the variable or the buffer is not large enough to contain the result of the concatenation, SAS does the following:

- changes the result to a blank value in the DATA step, and in PROC SQL
- writes a warning message to the log stating that the result was either truncated or set to a blank value, depending on the calling environment
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation
- sets _ERROR_ to 1 in the DATA step

The CATT function removes leading and trailing blanks from numeric arguments after it formats the numeric value with the BEST*w*. format.

Comparisons

The results of the CAT, CATS, CATT, and CATX functions are *usually* equivalent to results that are produced by certain combinations of the concatenation operator (||) and the TRIM and LEFT functions. However, the default length for the CAT, CATS, CATT, and CATX functions is different from the length that is obtained when you use the concatenation operator. For more information, see "Length of Returned Variable" on page 552.

Using the CAT, CATS, CATT, and CATX functions is faster than using TRIM and LEFT, and you can use them with the OF syntax for variable lists in calling environments that support variable lists.

The following table shows equivalents of the CAT, CATS, CATT, and CATX functions. The variables X1 through X4 specify character variables, and SP specifies a delimiter, such as a blank or comma.

Function	Equivalent Code
`CAT(OF X1-X4)`	`X1||X2||X3||X4`
`CATS(OF X1-X4)`	`TRIM(LEFT(X1))||TRIM(LEFT(X2))||TRIM(LEFT(X3))||TRIM(LEFT(X4))`
`CATT(OF X1-X4)`	`TRIM(X1)||TRIM(X2)||TRIM(X3)||TRIM(X4)`
`CATX(SP, OF X1-X4)`	`TRIM(LEFT(X1))||SP||TRIM(LEFT(X2))||SP||TRIM(LEFT(X3))||SP||TRIM(LEFT(X4))`

Examples

The following example shows how the CATT function concatenates strings.

```
data _null_;
   x='  The   Olym';
   y='pic Arts Festi';
   z='  val includes works by D  ';
   a='ale Chihuly.';
   result=catt(x,y,z,a);
   put result $char.;
run;
```

The following line is written to the SAS log:

```
   ----+----1----+----2----+----3----+----4----+----5----+----6----+----7
     The   Olympic Arts Festi  val includes works by Dale Chihuly.
```

See Also

Functions and CALL Routines:

"CALL CATS Routine" on page 441
"CALL CATT Routine" on page 443
"CALL CATX Routine" on page 445
"CAT Function" on page 543
"CATQ Function" on page 546
"CATS Function" on page 550
"CATX Function" on page 554

CATX Function

Removes leading and trailing blanks, inserts delimiters, and returns a concatenated character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

CATX(*delimiter*, *item-1* <, … *item-n*>)

Arguments

delimiter
specifies a character string that is used as a delimiter between concatenated items.

item
specifies a constant, variable, or expression, either character or numeric. If *item* is numeric, then its value is converted to a character string by using the BEST*w*. format. In this case, SAS does not write a note to the log. For more information, see "The Basics" on page 555.

Details

The Basics

The CATX function first copies *item-1* to the result, omitting leading and trailing blanks. Then for each subsequent argument *item-i*, i=2, …, n, if *item-i* contains at least one non-blank character, then CATX appends *delimiter* and *item-i* to the result, omitting leading and trailing blanks from *item-i*. CATX does not insert the delimiter at the beginning or end of the result. Blank items do not produce delimiters at the beginning or end of the result, nor do blank items produce multiple consecutive delimiters.

Length of Returned Variable

In a DATA step, if the CATX function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes. If the concatenation operator (||) returns a value to a variable that has not previously been assigned a length, then that variable is given a length that is the sum of the lengths of the values which are being concatenated.

Length of Returned Variable: Special Cases The CATX function returns a value to a variable, or returns a value in a temporary buffer. The value that is returned from the CATX function has the following length:

- up to 200 characters in WHERE clauses and in PROC SQL
- up to 32767 characters in the DATA step except in WHERE clauses
- up to 65534 characters when CATX is called from the macro processor

If CATX returns a value in a temporary buffer, the length of the buffer depends on the calling environment, and the value in the buffer can be truncated after CATX finishes processing. In this case, SAS does not write a message about the truncation to the log.

If the length of the variable or the buffer is not large enough to contain the result of the concatenation, SAS does the following:

- changes the result to a blank value in the DATA step, and in PROC SQL
- writes a warning message to the log stating that the result was either truncated or set to a blank value, depending on the calling environment
- writes a note to the log that shows the location of the function call and lists the argument that caused the truncation
- sets _ERROR_ to 1 in the DATA step

Comparisons

The results of the CAT, CATS, CATT, and CATX functions are *usually* equivalent to results that are produced by certain combinations of the concatenation operator (||) and the TRIM and LEFT functions. However, the default length for the CAT, CATS, CATT, and CATX functions is different from the length that is obtained when you use the concatenation operator. For more information, see "Length of Returned Variable" on page 555.

Using the CAT, CATS, CATT, and CATX functions is faster than using TRIM and LEFT, and you can use them with the OF syntax for variable lists in calling environments that support variable lists.

Note: In the case of variables that have missing values, the concatenation produces different results. See Example 2 on page 557. △

The following table shows equivalents of the CAT, CATS, CATT, and CATX functions. The variables X1 through X4 specify character variables, and SP specifies a delimiter, such as a blank or comma.

Function	Equivalent Code
`CAT(OF X1-X4)`	`X1||X2||X3||X4`
`CATS(OF X1-X4)`	`TRIM(LEFT(X1))||TRIM(LEFT(X2))||TRIM(LEFT(X3))|| TRIM(LEFT(X4))`
`CATT(OF X1-X4)`	`TRIM(X1)||TRIM(X2)||TRIM(X3)||TRIM(X4)`
`CATX(SP, OF X1-X4)`	`TRIM(LEFT(X1))||SP||TRIM(LEFT(X2))||SP|| TRIM(LEFT(X3))||SP||TRIM(LEFT(X4))`

Examples

Example 1: Concatenating Strings That Have No Missing Values

The following example shows how the CATX function concatenates strings the have no missing values.

```
data _null_;
   separator='%%$%%';
   x='The Olympic  ';
   y='   Arts Festival ';
   z='   includes works by ';
   a='Dale Chihuly.';
   result=catx(separator,x,y,z,a);
   put result $char.;
run;
```

The following line is written to the SAS log:

```
   ----+----1----+----2----+----3----+----4----+----5----+----6----+----7
   The Olympic%%$%%Arts Festival%%$%%includes works by%%$%%Dale Chihuly.
```

Example 2: Concatenating Strings That Have Missing Values

The following example shows how the CATX function concatenates strings that contain missing values.

```
options nodate nostimer ls=78 ps=60;

data one;
   length x1--x4 $1;
   input x1--x4;
   datalines;
A B C D
E . F G
H . . J
;
run;

data two;
   set one;
   SP='^';
   test1=catx(sp, of x1--x4);
   test2=trim(left(x1)) || sp || trim(left(x2)) || sp || trim(left(x3)) || sp ||
      trim(left(x4));
run;

proc print data=two;
run;
```

SAS creates the following output:

Output 4.36 Using CATX with Missing Values

```
                          The SAS System                                1

          Obs    x1    x2    x3    x4    SP    test1             test2

           1        A     B     C     D     ^     A^B^C^D    A^B^C^D
           2        E               F     G     ^     E^F^G          E^ ^F^G
           3        H                     J     ^     H^J                 H^ ^ ^J
```

See Also

Functions and CALL Routines:

"CALL CATS Routine" on page 441
"CALL CATT Routine" on page 443
"CALL CATX Routine" on page 445
"CAT Function" on page 543
"CATQ Function" on page 546
"CATS Function" on page 550
"CATT Function" on page 552

CDF Function

Returns a value from a cumulative probability distribution.

Category: Probability

Syntax

CDF (*distribution,quantile<,parm-1, … ,parm-k>*)

Arguments

distribution
is a character constant, variable, or expression that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	**BERNOULLI**
Beta	**BETA**
Binomial	**BINOMIAL**
Cauchy	**CAUCHY**
Chi-Square	**CHISQUARE**
Exponential	**EXPONENTIAL**
F	**F**
Gamma	**GAMMA**
Geometric	**GEOMETRIC**
Hypergeometric	**HYPERGEOMETRIC**
Laplace	**LAPLACE**
Logistic	**LOGISTIC**
Lognormal	**LOGNORMAL**
Negative binomial	**NEGBINOMIAL**
Normal	**NORMAL|GAUSS**
Normal mixture	**NORMALMIX**
Pareto	**PARETO**
Poisson	**POISSON**
T	**T**
Uniform	**UNIFORM**
Wald (inverse Gaussian)	**WALD|IGAUSS**
Weibull	**WEIBULL**

Note: Except for T, F, and NORMALMIX, you can minimally identify any distribution by its first four characters. △

quantile
is a numeric constant, variable, or expression that specifies the value of the random variable.

parm-1, ... ,parm-k
are optional constants, variables, or expressions that specify *shape*, *location*, or *scale* parameters appropriate for the specific distribution.

See: "Details" on page 559 for complete information about these parameters

Details

The CDF function computes the left cumulative distribution function from various continuous and discrete probability distributions.

Note: The QUANTILE function returns the quantile from a distribution that you specify. The QUANTILE function is the inverse of the CDF function. For more information, see "QUANTILE Function" on page 1064 . △

Bernoulli Distribution

CDF('BERNOULLI',*x*,*p*)

where

x
is a numeric random variable.

p
is a numeric probability of success.

Range: $0 \le p \le 1$

The CDF function for the Bernoulli distribution returns the probability that an observation from a Bernoulli distribution, with probability of success equal to p, is less than or equal to x. The equation follows:

$$CDF\left('BERN', x, p\right) = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

Note: There are no *location* or *scale* parameters for this distribution. △

Beta Distribution

CDF('BETA',*x*,*a*,*b*<,*l*,*r*>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.
Range: $a > 0$

b
is a numeric shape parameter.
Range: $b > 0$

l
is the numeric left location parameter.
Default: 0

r
is the right location parameter.
Default: 1
Range: $r > l$

The CDF function for the beta distribution returns the probability that an observation from a beta distribution, with shape parameters a and b, is less than or equal to v. The following equation describes the CDF function of the beta distribution:

$$CDF\left('BETA', x, a, b, l, r\right) = \begin{cases} 0 & x \leq l \\ \frac{1}{\beta(a,b)} \int_{l}^{x} \frac{(v-l)^{a-1}(r-v)^{b-1}}{(r-l)^{a+b-1}} dv & l < x \leq r \\ 1 & x > r \end{cases}$$

where

$$\beta\left(a, b\right) = \frac{\Gamma\left(a\right)\Gamma\left(b\right)}{\Gamma\left(a+b\right)}$$

and

$$\Gamma\left(a\right) = \int_{0}^{\infty} x^{a-1} e^{-x} dx$$

Binomial Distribution

CDF('BINOMIAL',*m*,*p*,*n*)

where

m
is an integer random variable that counts the number of successes.

Range: *m* = 0, 1, ...

p
is a numeric probability of success.

Range: $0 \leq p \leq 1$

n
is an integer parameter that counts the number of independent Bernoulli trials.

Range: *n* = 0, 1, ...

The CDF function for the binomial distribution returns the probability that an observation from a binomial distribution, with parameters *p* and *n*, is less than or equal to *m*. The equation follows:

$$CDF\left('BINOM', m, p, n\right) = \begin{cases} 0 & m < 0 \\ \sum_{j=0}^{m} \binom{N}{j} p^{j} \left(1-p\right)^{n-j} & 0 \leq m \leq n \\ 1 & m > n \end{cases}$$

Note: There are no *location* or *scale* parameters for the binomial distribution. △

Cauchy Distribution

CDF('CAUCHY',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter.

Default: 0

λ
is a numeric scale parameter.

Default: 1

Range: $\lambda > 0$

The CDF function for the Cauchy distribution returns the probability that an observation from a Cauchy distribution, with the location parameter θ and the scale parameter λ, is less than or equal to *x*. The equation follows:

$$CDF\left('CAUCHY', x, \theta, \lambda\right) = \frac{1}{2} + \frac{1}{\pi} tan^{-1}\left(\frac{x-\theta}{\lambda}\right)$$

Chi-Square Distribution

CDF('CHISQUARE',*x*,*df* <,*nc*>)

where

x
is a numeric random variable.

df
is a numeric degrees of freedom parameter.

Range: $df > 0$

nc
is an optional numeric non-centrality parameter.

Range: $nc \geq 0$

The CDF function for the chi-square distribution returns the probability that an observation from a chi-square distribution, with *df* degrees of freedom and non-centrality parameter *nc*, is less than or equal to *x*. This function accepts non-integer degrees of freedom. If *nc* is omitted or equal to zero, the value returned is from the central chi-square distribution. In the following equation, let ν = *df* and let λ = *nc*. The following equation describes the CDF function of the chi-square distribution:

$$CDF\left('CHISQ', x, \nu, \lambda\right) = \begin{cases} 0 & x < 0 \\ \sum_{j=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^j}{j!} P_c\left(x, \nu + 2j\right) & x \geq 0 \end{cases}$$

where $P_c(.,.)$ denotes the probability from the central chi-square distribution:

$$P_c\left(x, a\right) = P_g\left(\frac{x}{2}, \frac{a}{2}\right)$$

and where $P_g(y,b)$ is the probability from the gamma distribution given by

$$P_g\left(y, b\right) = \frac{1}{\Gamma\left(b\right)} \int_0^y e^{-v} v^{b-1} dv$$

Exponential Distribution

CDF('EXPONENTIAL',*x* <,λ>)

where

x
is a numeric random variable.

λ
is a scale parameter.

Default: 1

Range: $\lambda > 0$

The CDF function for the exponential distribution returns the probability that an observation from an exponential distribution, with the scale parameter λ, is less than or equal to x. The equation follows:

$$CDF\left('EXPO', x, \lambda\right) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\frac{x}{\lambda}} & x \geq 0 \end{cases}$$

F Distribution

CDF('F',*x*,*ndf*,*ddf* <,*nc*>)

where

x
is a numeric random variable.

ndf
is a numeric numerator degrees of freedom parameter.

Range: *ndf* > 0

ddf
is a numeric denominator degrees of freedom parameter.

Range: *ddf* > 0

nc
is a numeric non-centrality parameter.

Range: *nc* ≥ 0

The CDF function for the *F* distribution returns the probability that an observation from an *F* distribution, with *ndf* numerator degrees of freedom, *ddf* denominator degrees of freedom, and non-centrality parameter *nc*, is less than or equal to *x*. This function accepts non-integer degrees of freedom for *ndf* and *ddf*. If *nc* is omitted or equal to zero, the value returned is from a central *F* distribution. In the following equation, let ν_1 = *ndf*, let ν_2 = *ddf*, and let λ = *nc*. The following equation describes the CDF function of the *F* distribution:

$$CDF\left('F', x, v_1, v_2, \lambda\right) = \begin{cases} 0 & x < 0 \\ \sum_{j=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^j}{j!} P_F\left(x, v_1 + 2j, v_2\right) & x \geq 0 \end{cases}$$

where $P_f(f,u_1,u_2)$ is the probability from the central *F* distribution with

$$P_F\left(x, u_1, u_2\right) = P_B\left(\frac{u_1 x}{u_1 x + u_2}, \frac{u_1}{2}, \frac{u_2}{2}\right)$$

and $P_B(x,a,b)$ is the probability from the standard beta distribution.

Note: There are no *location* or *scale* parameters for the *F* distribution. △

Gamma Distribution

CDF('GAMMA',*x*,*a*<,λ>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.
Range: $a > 0$

λ
is a numeric scale parameter.
Default: 1
Range: $\lambda > 0$

The CDF function for the gamma distribution returns the probability that an observation from a gamma distribution, with shape parameter a and scale parameter λ, is less than or equal to x. The equation follows:

$$CDF\left('GAMMA', x, a, \lambda\right) = \begin{cases} 0 & x < 0 \\ \frac{1}{\lambda^a \Gamma(a)} \int_0^x v^{a-1} e^{-\frac{v}{\lambda}} dv & x \geq 0 \end{cases}$$

Geometric Distribution

CDF('GEOMETRIC',*m*,*p*)

where

m
is a numeric random variable that denotes the number of failures.
Range: $m = 0, 1, \ldots$

p
is a numeric probability of success.
Range: $0 \leq p \leq 1$

The CDF function for the geometric distribution returns the probability that an observation from a geometric distribution, with parameter p, is less than or equal to m. The equation follows:

$$CDF\left('GEOM', m, p\right) = \begin{cases} 0 & m < 0 \\ 1 - (1-p)^{\{m+1\}} & m \geq 0 \end{cases}$$

Note: There are no *location* or *scale* parameters for this distribution. △

Hypergeometric Distribution

CDF('HYPER',*x,N,R,n*<,*o*>)

where

x
is an integer random variable.

N
is an integer population size parameter.
Range: $N = 1, 2, \ldots$

R
is an integer number of items in the category of interest.
Range: $R = 0, 1, \ldots, N$

n
is an integer sample size parameter.
Range: $n = 1, 2, \ldots, N$

o
is an optional numeric odds ratio parameter.
Range: $o > 0$

The CDF function for the hypergeometric distribution returns the probability that an observation from an extended hypergeometric distribution, with population size *N*, number of items *R*, sample size *n*, and odds ratio *o*, is less than or equal to *x*. If *o* is omitted or equal to 1, the value returned is from the usual hypergeometric distribution. The equation follows:

$$CDF\left('HYPER', x, N, R, n, o\right) =$$
$$\begin{cases} 0 & x < max\left(0, R+n-N\right) \\ \dfrac{\sum_{i=0}^{x}\binom{R}{i}\binom{N-R}{n-i}o^{i}}{\sum_{j=max(0,R+n-N)}^{min(R,n)}\binom{R}{j}\binom{N-R}{n-j}o^{j}} & max\left(0, R+n-N\right) \le x \le min\left(R, n\right) \\ 1 & x > min\left(R, n\right) \end{cases}$$

Laplace Distribution

CDF('LAPLACE',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter.

Default: 0

λ
is a numeric scale parameter.

Default: 1

Range: $\lambda > 0$

The CDF function for the Laplace distribution returns the probability that an observation from the Laplace distribution, with the location parameter θ and the scale parameter λ, is less than or equal to x. The equation follows:

$$CDF\left('LAPLACE', x, \theta, \lambda\right) = \begin{cases} \frac{1}{2}e^{\frac{(x-\theta)}{\lambda}} & x < \theta \\ 1 - \frac{1}{2}e^{\left(-\frac{(x-\theta)}{\lambda}\right)} & x \geq \theta \end{cases}$$

Logistic Distribution

CDF('LOGISTIC',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter

Default: 0

λ
is a numeric scale parameter.

Default: 1

Range: $\lambda > 0$

The CDF function for the logistic distribution returns the probability that an observation from a logistic distribution, with a location parameter θ and a scale parameter λ, is less than or equal to x. The equation follows:

$$CDF\left('LOGISTIC', x, \theta, \lambda\right) = \frac{1}{1 + e^{-\frac{x-\theta}{\lambda}}}$$

Lognormal Distribution

CDF('LOGNORMAL',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
specifies a numeric log scale parameter. (exp(θ) is a scale parameter.)
Default: 0

λ
specifies a numeric shape parameter.
Default: 1
Range: $\lambda > 0$

The CDF function for the lognormal distribution returns the probability that an observation from a lognormal distribution, with the log scale parameter θ and the shape parameter λ, is less than or equal to x. The equation follows:

$$CDF\left('LOGN', x, \theta, \lambda\right) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{\lambda\sqrt{2\pi}} \int_{-\infty}^{log(x)} exp\left(-\frac{(v-\theta)^2}{2\lambda^2}\right) dv & x > 0 \end{cases}$$

Negative Binomial Distribution

CDF('NEGBINOMIAL',*m,p,n*)

where

m
is a positive integer random variable that counts the number of failures.
Range: $m = 0, 1, \ldots$

p
is a numeric probability of success.
Range: $0 \leq p \leq 1$

n
is a numeric value that counts the number of successes.
Range: $n > 0$

The CDF function for the negative binomial distribution returns the probability that an observation from a negative binomial distribution, with probability of success p and number of successes n, is less than or equal to m. The equation follows:

$$CDF\left('NEGB', m, p, n\right) = \begin{cases} 0 & m < 0 \\ p^n \sum_{j=0}^{m} \binom{n+j-1}{n-1} (1-p)^j & m \geq 0 \end{cases}$$

Note: There are no *location* or *scale* parameters for the negative binomial distribution. △

Normal Distribution

CDF('NORMAL',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter.

Default: 0

λ
is a numeric scale parameter.

Default: 1

Range: $\lambda > 0$

The CDF function for the normal distribution returns the probability that an observation from the normal distribution, with the location parameter θ and the scale parameter λ, is less than or equal to x. The equation follows:

$$CDF\left('NORMAL', x, \theta, \lambda\right) = \frac{1}{\lambda\sqrt{2\pi}} \int_{-\infty}^{x} exp\left(-\frac{(v-\theta)^2}{2\lambda^2}\right) dv$$

Normal Mixture Distribution

CDF('NORMALMIX',*x,n,p,m,s*)

where

x
is a numeric random variable.

n
is the integer number of mixtures.

Range: *n* = 1, 2, ...

p
is the n proportions, $p_1, p_2, \ldots, p_n$, where $\sum_{i=1}^{i=n} p_i = 1$.

Range: *p* = 0, 1, ...

m
is the *n* means $m_1, m_2, \ldots, m_n$.

s
is the *n* standard deviations $s_1, s_2, \ldots, s_n$.

Range: *s* > 0

The CDF function for the normal mixture distribution returns the probability that an observation from a mixture of normal distribution is less than or equal to *x*. The equation follows:

$$CDF\left('NORMALMIX', x, n, p, m, s\right) = \sum_{i=1}^{i=n} p_i \, CDF\left('NORMAL', x, m_i, s_i\right)$$

Note: There are no *location* or *scale* parameters for the normal mixture distribution. △

Pareto Distribution

CDF('PARETO',*x,a*<,*k*>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.

Range: *a* > 0

k
is a numeric scale parameter.

Default: 1

Range: *k* > 0

The CDF function for the Pareto distribution returns the probability that an observation from a Pareto distribution, with the shape parameter a and the scale parameter k, is less than or equal to x. The equation follows:

$$CDF\left('PARETO', x, a, k\right) = \begin{cases} 0 & x < k \\ 1 - \left(\frac{k}{x}\right)^a & x \geq k \end{cases}$$

Poisson Distribution

CDF('POISSON',*n,m*)

where

n
is an integer random variable.
Range: n = 0, 1, ...

m
is a numeric mean parameter.
Range: $m > 0$

The CDF function for the Poisson distribution returns the probability that an observation from a Poisson distribution, with mean m, is less than or equal to n. The equation follows:

$$CDF\left('POISSON', n, m\right) = \begin{cases} 0 & n < 0 \\ \sum_{i=0}^{n} exp\left(-m\right) \frac{m^i}{i!} & n \geq 0 \end{cases}$$

Note: There are no *location* or *scale* parameters for the Poisson distribution. △

T Distribution

CDF('T',*t,df*<,*nc*>)

where

t
is a numeric random variable.

df
is a numeric degrees of freedom parameter.
Range: $df > 0$

nc
is an optional numeric non-centrality parameter.

The CDF function for the *T* distribution returns the probability that an observation from a *T* distribution, with degrees of freedom *df* and non-centrality parameter *nc*, is less than or equal to *x*. This function accepts non-integer degrees of freedom. If *nc* is omitted or equal to zero, the value returned is from the central *T* distribution. In the following equation, let ν = *df* and let δ = *nc*. The equation follows:

$$CDF\left('T',t,v,\delta\right)=\frac{1}{2^{\left(\frac{1}{2}v-1\right)}\Gamma\left(\frac{1}{2}v\right)}\int_{0}^{\infty}x^{v-1}e^{-\frac{1}{2}x^{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{tx}{\sqrt{v}}}e^{-\frac{1}{2}(u-\delta)^{2}}dudx$$

Note: There are no *location* or *scale* parameters for the *T* distribution. △

Uniform Distribution

CDF('UNIFORM',*x*<,*l*,*r*>)

where

x
is a numeric random variable.

l
is the numeric left location parameter.
Default: 0

r
is the numeric right location parameter.
Default: 1
Range: $r > l$

The CDF function for the uniform distribution returns the probability that an observation from a uniform distribution, with the left location parameter *l* and the right location parameter *r*, is less than or equal to *x*. The equation follows:

$$CDF\left('UNIFORM',x,l,r\right)=\begin{cases}0 & x<l\\ \frac{x-l}{r-l} & l\leq x<r\\ 1 & x\geq r\end{cases}$$

Note: The default values for *l* and *r* are 0 and 1, respectively. △

Wald (Inverse Gaussian) Distribution

CDF('WALD',*x*,*d*)

CDF('IGAUSS',*x*,*d*)

where

x
is a numeric random variable.

d
is a numeric shape parameter.
Range: $d > 0$

The CDF function for the Wald distribution returns the probability that an observation from a Wald distribution, with shape parameter d, is less than or equal to x. The equation follows:

$$CDF\left('WALD', x, d\right) = \begin{cases} 0 & x \le 0 \\ \Phi\left((x-1)\sqrt{\frac{d}{x}}\right) + e^{2d}\Phi\left(-(x+1)\sqrt{\frac{d}{x}}\right) & x > 0 \end{cases}$$

where $\Phi(.)$ denotes the probability from the standard normal distribution.

Note: There are no *location* or *scale* parameters for the Wald distribution. △

Weibull Distribution

CDF('WEIBULL',*x*,*a*<,λ>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.
Range: $a > 0$

λ
is a numeric scale parameter.
Default: 1
Range: $\lambda > 0$

The CDF function for the Weibull distribution returns the probability that an observation from a Weibull distribution, with the shape parameter a and the scale parameter λ is less than or equal to x. The equation follows:

$$CDF\left('WEIBULL', x, a, \lambda\right) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\left(\frac{x}{\lambda}\right)^a} & x \ge 0 \end{cases}$$

Examples

SAS Statements	Results
`y=cdf('BERN',0,.25);`	0.75
`y=cdf('BETA',0.2,3,4);`	0.09888
`y=cdf('BINOM',4,.5,10);`	0.37695
`y=cdf('CAUCHY',2);`	0.85242
`y=cdf('CHISQ',11.264,11);`	0.57858
`y=cdf('EXPO',1);`	0.63212
`y=cdf('F',3.32,2,3);`	0.82639
`y=cdf('GAMMA',1,3);`	0.080301
`y=cdf('HYPER',2,200,50,10);`	0.52367
`y=cdf('LAPLACE',1);`	0.81606

SAS Statements	Results
`y=cdf('LOGISTIC',1);`	`0.73106`
`y=cdf('LOGNORMAL',1);`	`0.5`
`y=cdf('NEGB',1,.5,2);`	`0.5`
`y=cdf('NORMAL',1.96);`	`0.97500`
`y=cdf('NORMALMIX',2.3,3,.33,.33,.34, .5,1.5,2.5,.79,1.6,4.3);`	`0.7181`
`y=cdf('PARETO',1,1);`	`0`
`y=cdf('POISSON',2,1);`	`0.91970`
`y=cdf('T',.9,5);`	`0.79531`
`y=cdf('UNIFORM',0.25);`	`0.25`
`y=cdf('WALD',1,2);`	`0.62770`
`y=cdf('WEIBULL',1,2);`	`0.63212`

See Also

Functions:

"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120
"QUANTILE Function" on page 1064

CEIL Function

Returns the smallest integer that is greater than or equal to the argument, fuzzed to avoid unexpected floating-point results.

Category: Truncation

Syntax

CEIL (*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

If the argument is within 1E-12 of an integer, the function returns that integer.

Comparisons

Unlike the CEILZ function, the CEIL function fuzzes the result. If the argument is within 1E-12 of an integer, the CEIL function fuzzes the result to be equal to that integer. The CEILZ function does not fuzz the result. Therefore, with the CEILZ function you might get unexpected results.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`var1=2.1;`
`a=ceil(var1);`
`put a;`
	`3`

`b=ceil(-2.4);` `put b;`	`-2`
`c=ceil(1+1.e-11);` `put c;`	`2`
`d=ceil(-1+1.e-11);` `put d;`	`0`
`e=ceil(1+1.e-13);` `put e;`	`1`
`f=ceil(223.456);` `put f;`	`224`
`g=ceil(763);` `put g;`	`763`
`h=ceil(-223.456);` `put h;`	`-223`

See Also

Functions:

"CEILZ Function" on page 575

CEILZ Function

Returns the smallest integer that is greater than or equal to the argument, using zero fuzzing.

Category: Truncation

Syntax

CEILZ (*argument*)

Arguments

argument
is a numeric constant, variable, or expression.

Comparisons

Unlike the CEIL function, the CEILZ function uses zero fuzzing. If the argument is within 1E-12 of an integer, the CEIL function fuzzes the result to be equal to that integer. The CEILZ function does not fuzz the result. Therefore, with the CEILZ function you might get unexpected results.

Examples

The following SAS statements produce these results.

SAS Statements	Results
`a=ceilz(2.1);` `put a;`	`3`
`b=ceilz(-2.4);` `put b;`	`-2`
`c=ceilz(1+1.e-11);` `put c;`	`2`
`d=ceilz(-1+1.e-11);` `put d;`	`0`
`e=ceilz(1+1.e-13);` `put e;`	`2`
`f=ceilz(223.456);` `put f;`	`224`
`g=ceilz(763);` `put g;`	`763`
`h=ceilz(-223.456);` `put h;`	`-223`

See Also

Functions:

“CEIL Function” on page 573
“FLOOR Function” on page 757
“FLOORZ Function” on page 758
“INT Function” on page 829
“INTZ Function” on page 861
“ROUND Function” on page 1099
“ROUNDE Function” on page 1106
“ROUNDZ Function” on page 1108

CEXIST Function

Verifies the existence of a SAS catalog or SAS catalog entry.

Category: SAS File I/O

Syntax

CEXIST(*entry*<,'U'>)

Arguments

entry
is a character constant, variable, or expression that specifies a SAS catalog, or the name of an entry in a catalog. If the *entry* value is a one- or two-level name, then it is assumed to be the name of a catalog. Use a three- or four-level name to test for the existence of an entry within a catalog.

'U'
tests whether the catalog can be opened for updating.

Details

CEXIST returns 1 if the SAS catalog or catalog entry exists, or 0 if the SAS catalog or catalog entry does not exist.

Examples

Example 1: Verifying the Existence of an Entry in a Catalog This example verifies the existence of the entry X.PROGRAM in LIB.CAT1:

```
data _null_;
   if cexist("lib.cat1.x.program") then
   put "Entry X.PROGRAM exists";
run;
```

Example 2: Determining if a Catalog Can Be Opened for Update This example tests whether the catalog LIB.CAT1 exists and can be opened for update. If the catalog does not exist, a message is written to the SAS log. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%if %sysfunc(cexist(lib.cat1,u)) %then
   %put The catalog LIB.CAT1 exists and can be opened for update.;
%else
   %put %sysfunc(sysmsg());
```

See Also

Functions:

"EXIST Function" on page 675

CHAR Function

Returns a single character from a specified position in a character string.

Category: Character

Syntax

CHAR(*string*, *position*)

Arguments

string
specifies a character constant, variable, or expression.

position
is an integer that specifies the position of the character to be returned.

Details

In a DATA step, the default length of the target variable for the CHAR function is 1.

If *position* has a missing value, then CHAR returns a string with a length of 0. Otherwise, CHAR returns a string with a length of 1.

If *position* is less than or equal to 0, or greater than the length of the string, then CHAR returns a blank. Otherwise, CHAR returns the character at the specified position in the string.

Comparisons

The CHAR function returns the same result as SUBPAD(*string*, *position*, 1). While the results are the same, the default length of the target variable is different.

Examples

The following example shows the results of using the CHAR function.

```
options pageno=1 ps=64 ls=80 nodate;

data test;
   retain string "abc";
   do position = -1 to 4;
      result=char(string, position);
      output;
   end;
run;

proc print noobs data=test;
run;
```

Output 4.37 Output from the CHAR Function

```
                              The SAS System                              1

                      string    position    result

                       abc          -1
                       abc           0
                       abc           1          a
                       abc           2          b
                       abc           3          c
                       abc           4
```

See Also

Functions:

"FIRST Function" on page 755

CHOOSEC Function

Returns a character value that represents the results of choosing from a list of arguments.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

CHOOSEC (*index-expression*, *selection-1 <,…selection-n>*)

Arguments

index-expression
specifies a numeric constant, variable, or expression.

selection
specifies a character constant, variable, or expression. The value of this argument is returned by the CHOOSEC function.

Details

Length of Returned Variable In a DATA step, if the CHOOSEC function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The CHOOSEC function uses the value of *index-expression* to select from the arguments that follow. For example, if *index-expression* is three, CHOOSEC returns the value of *selection-3*. If the first argument is negative, the function counts backwards from the list of arguments, and returns that value.

Comparisons

The CHOOSEC function is similar to the CHOOSEN function except that CHOOSEC returns a character value while CHOOSEN returns a numeric value.

Examples

The following example shows how CHOOSEC chooses from a series of values:

```
data _null_;
   Fruit=choosec(1,'apple','orange','pear','fig');
   Color=choosec(3,'red','blue','green','yellow');
   Planet=choosec(2,'Mars','Mercury','Uranus');
   Sport=choosec(-3,'soccer','baseball','gymnastics','skiing');
   put Fruit= Color= Planet= Sport=;
run;
```

SAS writes the following line to the log:

```
Fruit=apple Color=green Planet=Mercury Sport=baseball
```

See Also

Functions:

"CHOOSEN Function" on page 581

CHOOSEN Function

Returns a numeric value that represents the results of choosing from a list of arguments.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

CHOOSEN (*index-expression*, *selection-1* <,...*selection-n*>)

Arguments

index-expression
specifies a numeric constant, variable, or expression.

selection
specifies a numeric constant, variable, or expression. The value of this argument is returned by the CHOOSEN function.

Details

The CHOOSEN function uses the value of *index-expression* to select from the arguments that follow. For example, if *index-expression* is 3, CHOOSEN returns the value of *selection-3*. If the first argument is negative, the function counts backwards from the list of arguments, and returns that value.

Comparisons

The CHOOSEN function is similar to the CHOOSEC function except that CHOOSEN returns a numeric value while CHOOSEC returns a character value.

Examples

The following example shows how CHOOSEN chooses from a series of values:

```
data _null_;
   ItemNumber=choosen(5,100,50,3784,498,679);
   Rank=choosen(-2,1,2,3,4,5);
   Score=choosen(3,193,627,33,290,5);
   Value=choosen(-5,-37,82985,-991,3,1014,-325,3,54,-618);
   put ItemNumber= Rank= Score= Value=;
run;
```

SAS writes the following line to the log:

```
ItemNumber=679 Rank=4 Score=33 Value=1014
```

See Also

Functions:

"CHOOSEC Function" on page 579

CINV Function

Returns a quantile from the chi-square distribution.

Category: Quantile

Syntax

CINV (*p*,*df*<,*nc*>)

Arguments

p
is a numeric probability.

Range: $0 \leq p < 1$

df
is a numeric degrees of freedom parameter.

Range: $df > 0$

nc
is a numeric noncentrality parameter.

Range: $nc \geq 0$

Details

The CINV function returns the p^{th} quantile from the chi-square distribution with degrees of freedom *df* and a noncentrality parameter *nc*. The probability that an observation from a chi-square distribution is less than or equal to the returned quantile is *p*. This function accepts a noninteger degrees of freedom parameter *df*.

If the optional parameter *nc* is not specified or has the value 0, the quantile from the central chi-square distribution is returned. The noncentrality parameter *nc* is defined such that if X is a normal random variable with mean μ and variance 1, X^2 has a noncentral chi-square distribution with *df*=1 and $nc = \mu^2$.

CAUTION:
For large values of *nc*, the algorithm could fail. In that case, a missing value is returned. △

Note: CINV is the inverse of the PROBCHI function. △

Examples

The first statement following shows how to find the 95th percentile from a central chi-square distribution with 3 degrees of freedom. The second statement shows how to find the 95th percentile from a noncentral chi-square distribution with 3.5 degrees of freedom and a noncentrality parameter equal to 4.5.

SAS Statements	Results
`q1=cinv(.95,3);`	`7.8147279033`
`a2=cinv(.95,3.5,4.5);`	`7.504582117`

See Also

Functions:

"QUANTILE Function" on page 1064

CLOSE Function

Closes a SAS data set.

Category: SAS File I/O

Syntax

CLOSE(*data-set-id*)

Arguments

data-set-id
is a numeric variable that specifies the data set identifier that the OPEN function returns.

Details

CLOSE returns zero if the operation was successful, or returns a non-zero value if it was not successful. Close all SAS data sets as soon as they are no longer needed by the application.

Note: All data sets opened within a DATA step are closed automatically at the end of the DATA step. △

Examples

This example uses OPEN to open the SAS data set PAYROLL. If the data set opens successfully, indicated by a positive value for the variable PAYID, the example uses CLOSE to close the data set.

```
%let payid=%sysfunc(open(payroll,is));
   macro statements
%if &payid > 0 %then
   %let rc=%sysfunc(close(&payid));
```

See Also

Function:

"OPEN Function" on page 980

CMISS Function

Counts the number of missing arguments.

Category: Descriptive Statistics

Syntax

CMISS(*argument-1* <, *argument-2,...*>)

Arguments

argument
specifies a constant, variable, or expression. *Argument* can be either a character value or a numeric value.

Details

A character expression is counted as missing if it evaluates to a string that contains all blanks or has a length of zero.

A numeric expression is counted as missing if it evaluates to a numeric missing value: ., ._, .A, … , .Z.

Comparisons

The CMISS function does not convert any argument. The NMISS function converts all arguments to numeric values.

See Also

Functions:

"NMISS Function" on page 947

"MISSING Function" on page 929

CNONCT Function

Returns the noncentrality parameter from a chi-square distribution.

Category: Mathematical

Syntax

CNONCT(*x*,*df*,*prob*)

Arguments

x

is a numeric random variable.

Range: $x \geq 0$

df

is a numeric degrees of freedom parameter.

Range: $df > 0$

prob

is a probability.

Range: $0 < prob < 1$

Details

The CNONCT function returns the nonnegative noncentrality parameter from a noncentral chi-square distribution whose parameters are *x*, *df*, and *nc*. If *prob* is greater than the probability from the central chi-square distribution with the parameters *x* and *df*, a root to this problem does not exist. In this case a missing value is returned. A Newton-type algorithm is used to find a nonnegative root *nc* of the equation

$$P_c\left(x|df,nc\right)-prob=0$$

where

$$P_c\left(x|df,nc\right)=e^{\frac{-nc}{2}}\sum_{j=0}^{\infty}\frac{\left(\frac{nc}{2}\right)^j}{j!}P_g\left(\frac{x}{2}|\frac{df}{2}+j\right)$$

where $P_g\left(x|a\right)$ is the probability from the gamma distribution given by

$$P_g\left(x|a\right)=\frac{1}{\Gamma\left(a\right)}\int_0^x t^{a-1}e^{-t}dt$$

If the algorithm fails to converge to a fixed point, a missing value is returned.

Examples

```
data work;
   x=2;
   df=4;
   do nc=1 to 3 by .5;
      prob=probchi(x,df,nc);
      ncc=cnonct(x,df,prob);
      output;
   end;
run;
proc print;
run;
```

Output 4.38 Computations of the Noncentrality Parameters from the Chi-squared Distribution

OBS	x	df	nc	prob	ncc
1	2	4	1.0	0.18611	1.0
2	2	4	1.5	0.15592	1.5
3	2	4	2.0	0.13048	2.0
4	2	4	2.5	0.10907	2.5
5	2	4	3.0	0.09109	3.0

COALESCE Function

Returns the first non-missing value from a list of numeric arguments.

Category: Mathematical

Syntax

COALESCE(*argument-1*<*...*, *argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The Basics COALESCE accepts one or more numeric arguments. The COALESCE function checks the value of each argument in the order in which they are listed and returns the first non-missing value. If only one value is listed, then the COALESCE function returns the value of that argument. If all the values of all arguments are missing, then the COALESCE function returns a missing value.

Comparisons

The COALESCE function searches numeric arguments, whereas the COALESCEC function searches character arguments.

Examples

SAS Statements	Results
`x = COALESCE(42, .);`	`42`
`y = COALESCE(.A, .B, .C);`	`.`
`z = COALESCE(., 7, ., ., 42);`	`7`

See Also

Function:

"COALESCEC Function" on page 588

COALESCEC Function

Returns the first non-missing value from a list of character arguments.

Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

COALESCEC(*argument-1*<*...*, *argument-n*>)

Arguments

argument
specifies a character constant, variable, or expression.

Details

Length of Returned Variable In a DATA step, if the COALESCEC function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics COALESCEC accepts one or more character arguments. The COALESCEC function checks the value of each argument in the order in which they are listed and returns the first non-missing value. If only one value is listed, then the COALESCEC function returns the value of that argument. A character value is considered missing if it has a length of zero or if all the characters are blank. If all the values of all arguments are missing, then the COALESCEC function returns a string with a length of zero.

Comparisons

The COALESCEC function searches character arguments, whereas the COALESCE function searches numeric arguments.

Examples

SAS Statements	Results
`COALESCEC('', 'Hello')`	`Hello`
`COALESCEC ('', 'Goodbye', 'Hello')`	`Goodbye`

See Also

Function:
"COALESCE Function" on page 587

COLLATE Function

Returns a character string in ASCII or EBCDIC collating sequence.

Category: Character

Restriction: "I18N Level 0" on page 313

See: COLLATE Function in the documentation for your operating environment.

Syntax

COLLATE (*start-position*<,*end-position*>) | (*start-position*<,,*length*>)

Arguments

start-position
specifies the numeric position in the collating sequence of the first character to be returned.

Interaction: If you specify only *start-position*, COLLATE returns consecutive characters from that position to the end of the collating sequence or up to 255 characters, whichever comes first.

end-position
specifies the numeric position in the collating sequence of the last character to be returned.

The maximum *end-position* for the EBCDIC collating sequence is 255. For ASCII collating sequences, the characters that correspond to *end-position* values between 0 and 127 represent the standard character set. Other ASCII characters that correspond to *end-position* values between 128 and 255 are available on certain ASCII operating environments, but the information that those characters represent varies with the operating environment.

Tip: *end-position* must be larger than *start-position*

Tip: If you specify *end-position*, COLLATE returns all character values in the collating sequence between *start-position* and *end-position,* inclusive.

Tip: If you omit *end-position* and use *length*, mark the *end-position* place with a comma.

length
specifies the number of characters in the collating sequence.

Default: 200

Tip: If you omit *end-position*, use *length* to specify the length of the result explicitly.

Details

Length of Returned Variable In a DATA step, if the COLLATE function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics If you specify both *end-position* and *length*, COLLATE ignores *length*. If you request a string longer than the remainder of the sequence, COLLATE returns a string through the end of the sequence.

Examples

The following SAS statements produce these results.

SAS Statements	Results
ASCII	`----+----1----+-----2--`

`x=collate(48,,10);`
`y=collate(48,57);`
`put @1 x @14 y;`
	`0123456789  0123456789`

EBCDIC	

`x=collate(240,,10);`
`y=collate(240,249);`
`put @1 x @14 y;`
	`0123456789  0123456789`

See Also

Functions:

"BYTE Function" on page 431
"RANK Function" on page 1085

COMB Function

Computes the number of combinations of *n* elements taken *r* at a time.

Category: Combinatorial

Syntax

COMB(*n, r*)

Arguments

n
is a nonnegative integer that represents the total number of elements from which the sample is chosen.

r
is a nonnegative integer that represents the number of chosen elements.

Restriction: $r \leq n$

Details

The mathematical representation of the COMB function is given by the following equation:

$$COMB(n,r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

with $n \geq 0$, $r \geq 0$, and $n \geq r$.

If the expression cannot be computed, a missing value is returned. For moderately large values, it is sometimes not possible to compute the COMB function.

Examples

SAS Statements	Results
`x=comb(5,1);`	`5`

See Also

Functions:

"FACT Function" on page 678
"PERM Function" on page 1008
"LCOMB Function" on page 880

COMPARE Function

Returns the position of the leftmost character by which two strings differ, or returns 0 if there is no difference.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KCOMPARE in *SAS National Language Support (NLS): Reference Guide*. See also "DBCS Compatibility" on page 592.

Syntax

COMPARE(*string–1*, *string–2*<,*modifiers*>)

Arguments

string–1
specifies a character constant, variable, or expression.

string–2
specifies a character constant, variable, or expression.

modifier
specifies a character string that can modify the action of the COMPARE function. You can use one or more of the following characters as a valid modifier:

i or I	ignores the case in *string–1* and *string–2*.
l or L	removes leading blanks in *string–1* and *string–2* before comparing the values.
n or N	removes quotation marks from any argument that is a name literal and ignores the case of *string–1* and *string–2*.
	Tip: A name literal is a name token that is expressed as a string within quotation marks, followed by the uppercase or lowercase letter *n*. Name literals enable you to use special characters (including blanks) that are not otherwise allowed in SAS data set or variable names. For COMPARE to recognize a string as a name literal, the first character must be a quotation mark.
: (colon)	truncates the longer of *string–1* or *string–2* to the length of the shorter string, or to one, whichever is greater. If you do not specify this modifier, the shorter string is padded with blanks to the same length as the longer string.

Tip: COMPARE ignores blanks that are used as modifiers.

Details

The Basics The order in which the modifiers appear in the COMPARE function is relevant.

- "LN" first removes leading blanks from each string, and then removes quotation marks from name literals.
- "NL" first removes quotation marks from name literals, and then removes leading blanks from each string.

In the COMPARE function, if *string–1* and *string–2* do not differ, COMPARE returns a value of zero. If the arguments differ, then the following apply:

- The sign of the result is negative if *string–1* precedes *string–2* in a sort sequence, and positive if *string–1* follows *string–2* in a sort sequence.
- The magnitude of the result is equal to the position of the leftmost character at which the strings differ.

DBCS Compatibility

The DBCS equivalent function is KCOMPARE, which is documented in *SAS National Language Support (NLS): Reference Guide*. There are minor differences between the COMPARE and KCOMPARE functions. While both functions accept varying numbers of arguments, usage of the third argument is not compatible. The following example shows the differences in the syntax:

COMPARE(*string-1*, *string-2* <, *modifiers*>)

KCOMPARE(*string-1* <, *position* <, *count*>>, *string-2*)

Examples

Example 1: Understanding the Order of Comparisons When Comparing Two Strings The following example compares two strings by using the COMPARE function.

```
options pageno=1 nodate ls=80 ps=60;

data test;
   infile datalines missover;
   input string1 $char8. string2 $char8. modifiers $char8.;
   result=compare(string1, string2, modifiers);
   datalines;
1234567812345678
123     abc
abc     abx
xyz     abcdef
aBc     abc
aBc     AbC     i
   abc  abc
   abc  abc     l
 abc       abx
 abc       abx  l
ABC     'abc'n
ABC     'abc'n  n
 '$12'n $12     n
 '$12'n $12     nl
 '$12'n $12     ln
;

proc print data=test;
run;
```

The following output shows the results.

Output 4.39 Results of Comparing Two Strings by Using the COMPARE Function

The SAS System 1

Obs	string1	string2	modifiers	result
1	12345678	12345678		0
2	123	abc		-1
3	abc	abx		-3
4	xyz	abcdef		1
5	aBc	abc		-2
6	aBc	AbC	i	0
7	abc	abc		-1
8	abc	abc	l	0
9	abc	abx		2
10	abc	abx	l	-3
11	ABC	'abc'n		1
12	ABC	'abc'n	n	0
13	'$12'n	$12	n	-1
14	'$12'n	$12	nl	1
15	'$12'n	$12	ln	0

Example 2: Truncating Strings Using the COMPARE Function The following example uses the : (colon) modifier to truncate strings.

```
options pageno=1 nodate ls=80 pagesize=60;

data test2;
   pad1=compare('abc','abc            ');
   pad2=compare('abc','abcdef         ');
   truncate1=compare('abc','abcdef',':');
   truncate2=compare('abcdef','abc',':');
   blank=compare('','abc',          ':');
run;

proc print data=test2 noobs;
run;
```

The following output shows the results.

Output 4.40 Results of Using the Truncation Modifier

```
                              The SAS System                                1

    pad1    pad2    truncate1    truncate2    blank

      0      -4          0            0          -1
```

See Also

Functions and CALL Routines:

"COMPGED Function" on page 596
"COMPLEV Function" on page 601
"CALL COMPCOST Routine" on page 447

COMPBL Function

Removes multiple blanks from a character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

COMPBL(*source*)

Arguments

source
specifies a character constant, variable, or expression to compress.

Details

Length of Returned Variable In a DATA step, if the COMPBL function returns a value to a variable that has not previously been assigned a length, then the length of that variable defaults to the length of the first argument.

The Basics The COMPBL function removes multiple blanks in a character string by translating each occurrence of two or more consecutive blanks into a single blank.

Comparisons

The COMPRESS function removes every occurrence of the specific character from a string. If you specify a blank as the character to remove from the source string, the COMPRESS function removes all blanks from the source string, while the COMPBL function compresses multiple blanks to a single blank and has no effect on a single blank.

Examples

The following SAS statements produce these results.

SAS Statements	Results
	`----+----1----+-----2--`

`string='Hey`
`Diddle  Diddle';`
`string=compbl(string);`
`put string;`
	`Hey Diddle Diddle`

`string='125    E Main St';`
`length address $10;`
`address=compbl(string);`
`put address;`
	`125 E Main`

See Also

Function:

"COMPRESS Function" on page 604

COMPGED Function

Returns the generalized edit distance between two strings.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

COMPGED(*string-1*, *string-2* <,*cutoff*> <,*modifiers*>)

Arguments

string–1
specifies a character constant, variable, or expression.

string-2
specifies a character constant, variable, or expression.

cutoff
is a numeric constant, variable, or expression. If the actual generalized edit distance is greater than the value of *cutoff*, the value that is returned is equal to the value of *cutoff*.

Tip: Using a small value of *cutoff* improves the efficiency of COMPGED if the values of *string–1* and *string–2* are long.

modifiers
specifies a character string that can modify the action of the COMPGED function. You can use one or more of the following characters as a valid modifier:

i or I	ignores the case in *string–1* and *string–2*.
l or L	removes leading blanks in *string–1* and *string–2* before comparing the values.
n or N	removes quotation marks from any argument that is an n-literal and ignores the case of *string–1* and *string–2*.
: (colon)	truncates the longer of *string–1* or *string–2* to the length of the shorter string, or to one, whichever is greater.

Tip: COMPGED ignores blanks that are used as modifiers.

Details

The Order in Which Modifiers Appear The order in which the modifiers appear in the COMPGED function is relevant.

- "LN" first removes leading blanks from each string and then removes quotation marks from n-literals.
- "NL" first removes quotation marks from n-literals and then removes leading blanks from each string.

Definition of Generalized Edit Distance Generalized edit distance is a generalization of Levenshtein edit distance, which is a measure of dissimilarity between two strings.

The Levenshtein edit distance is the number of deletions, insertions, or replacements of single characters that are required to transform *string-1* into *string-2*.

Computing the Generalized Edit Distance The COMPGED function returns the generalized edit distance between *string-1* and *string-2*. The generalized edit distance is the minimum-cost sequence of operations for constructing *string-1* from *string-2*.

The algorithm for computing the sum of the costs involves a pointer that points to a character in *string-2* (the input string). An output string is constructed by a sequence of operations that might advance the pointer, add one or more characters to the output string, or both. Initially, the pointer points to the first character in the input string, and the output string is empty.

The operations and their costs are described in the following table.

Operation	Default Cost in Units	Description of Operation
APPEND	50	When the output string is longer than the input string, add any one character to the end of the output string without moving the pointer.
BLANK	10	Do any of the following: □ Add one space character to the end of the output string without moving the pointer. □ When the character at the pointer is a space character, advance the pointer by one position without changing the output string. □ When the character at the pointer is a space character, add one space character to the end of the output string, and advance the pointer by one position. If the cost for BLANK is set to zero by the COMPCOST function, the COMPGED function removes all space characters from both strings before doing the comparison.
DELETE	100	Advance the pointer by one position without changing the output string.
DOUBLE	20	Add the character at the pointer to the end of the output string without moving the pointer.
FDELETE	200	When the output string is empty, advance the pointer by one position without changing the output string.
FINSERT	200	When the pointer is in position one, add any one character to the end of the output string without moving the pointer.

Operation	Default Cost in Units	Description of Operation
FREPLACE	200	When the pointer is in position one and the output string is empty, add any one character to the end of the output string, and advance the pointer by one position.
INSERT	100	Add any one character to the end of the output string without moving the pointer.
MATCH	0	Copy the character at the pointer from the input string to the end of the output string, and advance the pointer by one position.
PUNCTUATION	30	Do any of the following: □ Add one punctuation character to the end of the output string without moving the pointer. □ When the character at the pointer is a punctuation character, advance the pointer by one position without changing the output string. □ When the character at the pointer is a punctuation character, add one punctuation character to the end of the output string, and advance the pointer by one position. If the cost for PUNCTUATION is set to zero by the COMPCOST function, the COMPGED function removes all punctuation characters from both strings before doing the comparison.
REPLACE	100	Add any one character to the end of the output string, and advance the pointer by one position.
SINGLE	20	When the character at the pointer is the same as the character that follows in the input string, advance the pointer by one position without changing the output string.
SWAP	20	Copy the character that follows the pointer from the input string to the output string. Then copy the character at the pointer from the input string to the output string. Advance the pointer two positions.
TRUNCATE	10	When the output string is shorter than the input string, advance the pointer by one position without changing the output string.

To set the cost of the string operations, you can use the CALL COMPCOST routine or use default costs. If you use the default costs, the values that are returned by COMPGED are approximately 100 times greater than the values that are returned by COMPLEV.

Examples of Errors The rationale for determining the generalized edit distance is based on the number and types of typographical errors that can occur. COMPGED assigns a cost to each error and determines the minimum sum of these costs that could be incurred. Some types of errors can be more serious than others. For example, inserting an extra letter at the beginning of a string might be more serious than omitting a letter from the end of a string. For another example, if you type a word or phrase that exists in *string-2* and introduce a typographical error, you might produce *string-1* instead of *string-2*.

Making the Generalized Edit Distance Symmetric Generalized edit distance is not necessarily symmetric. That is, the value that is returned by **`COMPGED(string1, string2)`** is not always equal to the value that is returned by **`COMPGED(string2, string1)`**. To make the generalized edit distance symmetric, use the CALL COMPCOST routine to assign equal costs to the operations within each of the following pairs:

- INSERT, DELETE
- FINSERT, FDELETE
- APPEND, TRUNCATE
- DOUBLE, SINGLE

Comparisons

You can compute the Levenshtein edit distance by using the COMPLEV function. You can compute the generalized edit distance by using the CALL COMPCOST routine and the COMPGED function. Computing generalized edit distance requires considerably more computer time than does computing Levenshtein edit distance. But generalized edit distance usually provides a more useful measure than Levenshtein edit distance for applications such as fuzzy file merging and text mining.

Examples

The following example uses the default costs to calculate the generalized edit distance.

```
options nodate pageno=1 linesize=70 pagesize=60;

data test;
   infile datalines missover;
   input String1 $char8. +1 String2 $char8. +1 Operation $40.;
   GED=compged(string1, string2);
   datalines;
baboon   baboon   match
baXboon  baboon   insert
baoon    baboon   delete
baXoon   baboon   replace
baboonX  baboon   append
baboo    baboon   truncate
babboon  baboon   double
babon    baboon   single
baobon   baboon   swap
bab oon  baboon   blank
```

```
bab,oon  baboon   punctuation
bXaoon   baboon   insert+delete
bXaYoon  baboon   insert+replace
bXoon    baboon   delete+replace
Xbaboon  baboon   finsert
aboon    baboon   trick question: swap+delete
Xaboon   baboon   freplace
axoon    baboon   fdelete+replace
axoo     baboon   fdelete+replace+truncate
axon     baboon   fdelete+replace+single
baby     baboon   replace+truncate*2
balloon  baboon   replace+insert
;

proc print data=test label;
   label GED='Generalized Edit Distance';
   var String1 String2 GED Operation;
run;
```

The following output shows the results.

Output 4.41 Generalized Edit Distance Based on Operation

```
                          The SAS System                          1

                         Generalized
                             Edit
Obs   String1   String2    Distance    Operation

  1   baboon    baboon          0       match
  2   baXboon   baboon        100       insert
  3   baoon     baboon        100       delete
  4   baXoon    baboon        100       replace
  5   baboonX   baboon         50       append
  6   baboo     baboon         10       truncate
  7   babboon   baboon         20       double
  8   babon     baboon         20       single
  9   baobon    baboon         20       swap
 10   bab oon   baboon         10       blank
 11   bab,oon   baboon         30       punctuation
 12   bXaoon    baboon        200       insert+delete
 13   bXaYoon   baboon        200       insert+replace
 14   bXoon     baboon        200       delete+replace
 15   Xbaboon   baboon        200       finsert
 16   aboon     baboon        200       trick question: swap+delete
 17   Xaboon    baboon        200       freplace
 18   axoon     baboon        300       fdelete+replace
 19   axoo      baboon        310       fdelete+replace+truncate
 20   axon      baboon        320       fdelete+replace+single
 21   baby      baboon        120       replace+truncate*2
 22   balloon   baboon        200       replace+insert
```

See Also

Functions:

"COMPARE Function" on page 591
"CALL COMPCOST Routine" on page 447
"COMPLEV Function" on page 601

COMPLEV Function

Returns the Levenshtein edit distance between two strings.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

COMPLEV(*string-1*, *string-2* <,*cutoff*> <,*modifiers*>)

Arguments

string–1
specifies a character constant, variable, or expression.

string–2
specifies a character constant, variable, or expression.

cutoff
specifies a numeric constant, variable, or expression. If the actual Levenshtein edit distance is greater than the value of *cutoff*, the value that is returned is equal to the value of *cutoff*.

Tip: Using a small value of *cutoff* improves the efficiency of COMPLEV if the values of *string–1* and *string–2* are long.

modifiers
specifies a character string that can modify the action of the COMPLEV function. You can use one or more of the following characters as a valid modifier:

i or I	ignores the case in *string–1* and *string–2*.
l or L	removes leading blanks in *string–1* and *string–2* before comparing the values.
n or N	removes quotation marks from any argument that is an n-literal and ignores the case of *string–1* and *string–2*.
: (colon)	truncates the longer of *string–1* or *string–2* to the length of the shorter string, or to one, whichever is greater.

TIP: COMPLEV ignores blanks that are used as modifiers.

Details

The order in which the modifiers appear in the COMPLEV function is relevant.

- "LN" first removes leading blanks from each string and then removes quotation marks from n-literals.
- "NL" first removes quotation marks from n-literals and then removes leading blanks from each string.

The COMPLEV function ignores trailing blanks.

COMPLEV returns the Levenshtein edit distance between *string-1* and *string-2*. Levenshtein edit distance is the number of insertions, deletions, or replacements of single characters that are required to convert one string to the other. Levenshtein edit distance is symmetric. That is, **COMPLEV(string-1,string-2)** is the same as **COMPLEV(string-2,string-1)**.

Comparisons

The Levenshtein edit distance that is computed by COMPLEV is a special case of the generalized edit distance that is computed by COMPGED.

COMPLEV executes much more quickly than COMPGED.

Examples

The following example compares two strings by computing the Levenshtein edit distance.

```
options pageno=1 nodate ls=80 ps=60;

data test;
   infile datalines missover;
   input string1 $char8. string2 $char8. modifiers $char8.;
   result=complev(string1, string2, modifiers);
   datalines;
12345678 12345678
abc      abxc
ac       abc
aXc      abc
aXbZc    abc
aXYZc    abc
WaXbYcZ  abc
XYZ      abcdef
aBc      abc
aBc      AbC      i
  abc    abc
  abc    abc      l
AxC      'abc'n
AxC      'abc'n   n
;

proc print data=test;
run;
```

The following output shows the results.

Output 4.42 Results of Comparing Two Strings by Computing the Levenshtein Edit Distance

The SAS System 1

Obs	string1	string2	modifiers	result
1	12345678	12345678		0
2	abc	abxc		1
3	ac	abc		1
4	aXc	abc		1
5	aXbZc	abc		2
6	aXYZc	abc		3
7	WaXbYcZ	abc		4
8	XYZ	abcdef		6
9	aBc	abc		1
10	aBc	AbC	i	0
11	abc	abc		2
12	abc	abc	l	0
13	AxC	'abc'n		6
14	AxC	'abc'n	n	1

See Also

Functions and CALL Routines:

"COMPARE Function" on page 591
"COMPGED Function" on page 596
"CALL COMPCOST Routine" on page 447

COMPOUND Function

Returns compound interest parameters.

Category: Financial

Syntax

COMPOUND(*a,f,r,n*)

Arguments

a
is numeric, and specifies the initial amount.
Range: $a \geq 0$

f
is numeric, and specifies the future amount (at the end of *n* periods).
Range: $f \geq 0$

r
is numeric, and specifies the periodic interest rate expressed as a fraction.

Range: $r \geq 0$

n
is an integer, and specifies the number of compounding periods.

Range: $n \geq 0$

Details

The COMPOUND function returns the missing argument in the list of four arguments from a compound interest calculation. The arguments are related by the following equation:

$$f = a(1 + r)^n$$

One missing argument must be provided. A compound interest parameter is then calculated from the remaining three values. No adjustment is made to convert the results to round numbers.

If n=0, then $f = a$ and $(1 + r)^n$ are equal to 1.

Note: If you choose *r* as your missing value, then COMPOUND returns an error. △

Examples

The accumulated value of an investment of $2000 at a nominal annual interest rate of 9 percent, compounded monthly after 30 months, can be expressed as

```
future=compound(2000,.,0.09/12,30);
```

The value returned is 2502.54. The second argument has been set to missing, indicating that the future amount is to be calculated. The 9 percent nominal annual rate has been converted to a monthly rate of 0.09/12. The rate argument is the fractional (not the percentage) interest rate per compounding period.

COMPRESS Function

Returns a character string with specified characters removed from the original string.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KCOMPRESS in *SAS National Language Support (NLS): Reference Guide*.

Syntax

COMPRESS(<*source*><, *chars*><, *modifiers*>)

Arguments

source
specifies a character constant, variable, or expression from which specified characters will be removed.

chars
specifies a character constant, variable, or expression that initializes a list of characters.

By default, the characters in this list are removed from the *source* argument. If you specify the K modifier in the third argument, then only the characters in this list are kept in the result.

Tip: You can add more characters to this list by using other modifiers in the third argument.

Tip: Enclose a literal string of characters in quotation marks.

modifier
specifies a character constant, variable, or expression in which each non-blank character modifies the action of the COMPRESS function. Blanks are ignored. The following characters can be used as modifiers:

a or A	adds alphabetic characters to the list of characters.
c or C	adds control characters to the list of characters.
d or D	adds digits to the list of characters.
f or F	adds the underscore character and English letters to the list of characters.
g or G	adds graphic characters to the list of characters.
h or H	adds a horizontal tab to the list of characters.
i or I	ignores the case of the characters to be kept or removed.
k or K	keeps the characters in the list instead of removing them.
l or L	adds lowercase letters to the list of characters.
n or N	adds digits, the underscore character, and English letters to the list of characters.
o or O	processes the second and third arguments once rather than every time the COMPRESS function is called. Using the O modifier in the DATA step (excluding WHERE clauses), or in the SQL procedure, can make COMPRESS run much faster when you call it in a loop where the second and third arguments do not change.
p or P	adds punctuation marks to the list of characters.
s or S	adds space characters (blank, horizontal tab, vertical tab, carriage return, line feed, and form feed) to the list of characters.
t or T	trims trailing blanks from the first and second arguments.
u or U	adds uppercase letters to the list of characters.
w or W	adds printable characters to the list of characters.
x or X	adds hexadecimal characters to the list of characters.

Tip: If the *modifier* is a constant, enclose it in quotation marks. Specify multiple constants in a single set of quotation marks. *Modifier* can also be expressed as a variable or an expression.

Details

Length of Returned Variable In a DATA step, if the COMPRESS function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the first argument.

The Basics The COMPRESS function allows null arguments. A null argument is treated as a string that has a length of zero.

Based on the number of arguments, the COMPRESS functions works as follows:

Number of Arguments	Result
only the first argument, *source*	The argument has all blanks removed. If the argument is completely blank, then the result is a string with a length of zero. If you assign the result to a character variable with a fixed length, then the value of that variable will be padded with blanks to fill its defined length.
the first two arguments, *source* and *chars*	All characters that appear in the second argument are removed from the result.
three arguments, *source*, *chars*, and *modifier(s)*	The K modifier (specified in the third argument) determines whether the characters in the second argument are kept or removed from the result.

The COMPRESS function compiles a list of characters to keep or remove, comprising the characters in the second argument plus any types of characters that are specified by the modifiers. For example, the D modifier specifies digits. Both of the following function calls remove digits from the result:

```
COMPRESS(source, "1234567890");
COMPRESS(source, , "d");
```

To remove digits and plus or minus signs, you can use either of the following function calls:

```
COMPRESS(source, "1234567890+-");
COMPRESS(source, "+-", "d");
```

Examples

Example 1: Compressing Blanks

SAS Statements	Results
	`----+----1`

`a='AB C D ';`
`b=compress(a);`
`put b;`
	`ABCD`

Example 2: Compressing Lowercase Letters

SAS Statements	Results
	`----+----1----+----2----+----3`

`x='123-4567-8901 B 234-5678-9012 c';`
`y=compress(x,'ABCD','l');`
`put y;`
	`123-4567-8901 234-5678-9012`

Example 3: Compressing Space Characters

SAS Statements	Results
	`----+----1`

`x='1  2  3  4  5';`
`y=compress(x,,'s');`
`put y;`
	`12345`

Example 4: Keeping Characters in the List

SAS Statements	Results
	`----+----1`

`x='Math A English B Physics A';`
`y=compress(x,'ABCD','k');`
`put y;`
	`ABA`

Example 5: Compressing a String and Returning a Length of 0

SAS Statements	Results
	`----+----1`

`x=' ';`
`l=lengthn(compress(x));`
`put l;`
	`0`

See Also

Functions:

"COMPBL Function" on page 594
"LEFT Function" on page 881
"TRIM Function" on page 1173

CONSTANT Function

Computes machine and mathematical constants.

Category: Mathematical

Syntax

CONSTANT(*constant*<, *parameter*>)

Arguments

constant
is a character constant, variable, or expression that identifies the constant to be returned. Valid constants are as follows:

Description	Constant
The natural base	`'E'`
Euler constant	`'EULER'`
Pi	`'PI'`
Exact integer	`'EXACTINT' <,nbytes>`
The largest double-precision number	`'BIG'`
The log with respect to *base* of BIG	`'LOGBIG' <,base>`
The square root of BIG	`'SQRTBIG'`

Description	Constant
The smallest double-precision number	**'SMALL'**
The log with respect to *base* of SMALL	**'LOGSMALL' <,base>**
The square root of SMALL	**'SQRTSMALL'**
Machine precision constant	**'MACEPS'**
The log with respect to *base* of MACEPS	**'LOGMACEPS' <,base>**
The square root of MACEPS	**'SQRTMACEPS'**

parameter
is an optional numeric parameter. Some of the constants specified in *constant* have an optional argument that alters the functionality of the CONSTANT function.

Details

CAUTION:
In some operating environments, the run-time library might have limitations that prevent the use of the full range of floating-point numbers that the hardware provides. In such cases, the CONSTANT function attempts to return values that are compatible with the limitations of the run-time library.

For example, if the run-time library cannot compute **EXP(LOG(CONSTANT('BIG')))**, then **CONSTANT('LOGBIG')** will not return the same value as **LOG(CONSTANT('BIG'))**, but will return a value such that **EXP(CONSTANT('LOGBIG'))** can be computed. △

The natural base

CONSTANT('E')

The natural base is described by the following equation:

$$\lim_{x \to 0} (1 + x)^{\frac{1}{x}} \approx 2.718281828459045$$

Euler constant

CONSTANT('EULER')

Euler's constant is described by the following equation:

$$\lim_{n \to \infty} \left\{ \sum_{j=1}^{j=n} \frac{1}{j} - \log(n) \right\} \approx 0.577215664901532860$$

Pi

CONSTANT('PI')

Pi is the ratio between the circumference and the diameter of a circle. Many expressions exist for computing this constant. One such expression for the series is described by the following equation:

$$4\sum_{j=0}^{j=\infty}\frac{(-1)^j}{2j+1} \approx 3.14159265358979323846$$

Exact integer

CONSTANT('EXACTINT' <, *nbytes*>)

where

nbytes
is a numeric value that is the number of bytes.

Range: $2 \le nbytes \le 8$

Default: 8

The exact integer is the largest integer k such that all integers less than or equal to k in absolute value have an exact representation in a SAS numeric variable of length *nbytes*. This information can be useful to know before you trim a SAS numeric variable from the default 8 bytes of storage to a lower number of bytes to save storage.

The largest double-precision number

CONSTANT('BIG')

This case returns the largest double-precision floating-point number (8-bytes) that is representable on your computer.

The logarithm of BIG

CONSTANT('LOGBIG' <, *base*>)

where

base
is a numeric value that is the base of the logarithm.

Restriction: The *base* that you specify must be greater than the value of 1+SQRTMACEPS.

Default: the natural base, E.

This case returns the logarithm with respect to *base* of the largest double-precision floating-point number (8-bytes) that is representable on your computer.

It is safe to exponentiate the given *base* raised to a power less than or equal to CONSTANT('LOGBIG', *base*) by using the power operation (**) without causing any overflows.

It is safe to exponentiate any floating-point number less than or equal to CONSTANT('LOGBIG') by using the exponential function, EXP, without causing any overflows.

The square root of BIG

CONSTANT('SQRTBIG')

This case returns the square root of the largest double-precision floating-point number (8-bytes) that is representable on your computer.

It is safe to square any floating-point number less than or equal to CONSTANT('SQRTBIG') without causing any overflows.

The smallest double-precision number

CONSTANT('SMALL')

This case returns the smallest double-precision floating-point number (8-bytes) that is representable on your computer.

The logarithm of SMALL

CONSTANT('LOGSMALL' <, *base*>)

where

base
is a numeric value that is the base of the logarithm.

Restriction: The *base* that you specify must be greater than the value of 1+SQRTMACEPS.

Default: the natural base, E.

This case returns the logarithm with respect to *base* of the smallest double-precision floating-point number (8-bytes) that is representable on your computer.

It is safe to exponentiate the given *base* raised to a power greater than or equal to CONSTANT('LOGSMALL', *base*) by using the power operation (**) without causing any underflows or 0.

It is safe to exponentiate any floating-point number greater than or equal to CONSTANT('LOGSMALL') by using the exponential function, EXP, without causing any underflows or 0.

The square root of SMALL

CONSTANT('SQRTSMALL')

This case returns the square root of the smallest double-precision floating-point number (8-bytes) that is representable on the computer.

It is safe to square any floating-point number greater than or equal to CONSTANT('SQRTBIG') without causing any underflows or 0.

Machine precision

CONSTANT('MACEPS')

This case returns the smallest double-precision floating-point number (8-bytes) $\epsilon = 2^{-j}$ for some integer *j*, such that $1 + \epsilon > 1$.

This constant is important in finite precision computations.

The logarithm of MACEPS

CONSTANT('LOGMACEPS' <, *base*>)

where

base
is a numeric value that is the base of the logarithm.

Restriction: The *base* that you specify must be greater than the value of 1+SQRTMACEPS.

Default: the natural base, E.

This case returns the logarithm with respect to *base* of CONSTANT('MACEPS').

The square root of MACEPS

CONSTANT('SQRTMACEPS')

This case returns the square root of CONSTANT('MACEPS').

CONVX Function

Returns the convexity for an enumerated cash flow.

Category: Financial

Syntax

CONVX(*y,f,c(1), ... ,c(k)*)

Arguments

y
specifies the effective per-period yield-to-maturity, expressed as a fraction.

Range: $0 < y < 1$

f
specifies the frequency of cash flows per period.

Range: $f > 0$

c(1), ... ,c(k)
specifies a list of cash flows.

Details

The CONVX function returns the value

$$C = \sum_{k=1}^{K} \frac{k\left(k+f\right)\ \frac{c(k)}{(1+y)^{\frac{k}{f}}}}{P\left(\left(1+y\right)^{2}\right) f^{2}}$$

where

$$P = \sum_{k=1}^{K} \frac{c\left(k\right)}{\left(1+y\right)^{\frac{k}{f}}}$$

Examples

```
data _null_;
   c=convx(1/20,1,.33,.44,.55,.49,.50,.22,.4,.8,.01,.36,.2,.4);
   put c;
run;
```

The value returned is 42.3778.

CONVXP Function

Returns the convexity for a periodic cash flow stream, such as a bond.

Category: Financial

Syntax

CONVXP(*A,c,n,K,k*$_0$*,y*)

Arguments

A
specifies the par value.
Range: $A > 0$

c
specifies the nominal per-period coupon rate, expressed as a fraction.
Range: $0 \leq c < 1$

n
specifies the number of coupons per period.
Range: $n > 0$ and is an integer

K
specifies the number of remaining coupons.
Range: $K > 0$ and is an integer

k_0
specifies the time from the present date to the first coupon date, expressed in terms of the number of periods.
Range: $0 < k_0 \leq \frac{1}{n}$

y
specifies the nominal per-period yield-to-maturity, expressed as a fraction.
Range: $y > 0$

Details

The CONVXP function returns the value

$$C = \frac{1}{n^2} \left(\frac{\sum_{k=1}^{K} t_k \left(t_k + 1\right) \frac{c(k)}{\left(1+\frac{y}{n}\right)^{t_k}}}{P\left(1 + \frac{y}{n}\right)^2} \right)$$

where

$$t_k = nk_0 + k - 1$$
$$c(k) = \frac{c}{n}A \quad for \;\; k = 1, \ldots, K-1$$
$$c(K) = \left(1 + \frac{c}{n}\right) A$$

and where

$$P = \sum_{k=1}^{K} \frac{c(k)}{\left(1 + \frac{y}{n}\right)^{t_k}}$$

Examples

In the following example, the CONVXP function returns the convexity of a bond that has a face value of 1000, an annual coupon rate of 0.01, 4 coupons per year, and 14 remaining coupons. The time from settlement date to next coupon date is 0.165, and the annual yield-to-maturity is 0.08.

```
data _null_;
   y=convxp(1000,.01,4,14,.33/2,.08);
   put y;
run;
```

The value that is returned is 11.729001987.

COS Function

Returns the cosine.

Category: Trigonometric

Syntax

COS (*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression and is expressed in radians. If the magnitude of *argument* is so great that **mod(argument,pi)** is accurate to less than about three decimal places, COS returns a missing value.

Examples

SAS Statements	Results
`x=cos(0.5);`	`0.8775825619`
`x=cos(0);`	`1`
`x=cos(3.14159/3);`	`0.500000766`

COSH Function

Returns the hyperbolic cosine.

Category: Hyperbolic

Syntax

COSH(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The COSH function returns the hyperbolic cosine of the argument, given by

$$\left(e^{argument} + e^{-argument}\right)/2$$

Examples

SAS Statements	Results
`x=cosh(0);`	`1`
`x=cosh(-5.0);`	`74.209948525`
`x=cosh(0.5);`	`1.1276259652`

COUNT Function

Counts the number of times that a specified substring appears within a character string.

Category: Character

Restriction: "I18N Level 1" on page 314

Tip: You can use the KCOUNT function in *SAS National Language Support (NLS): Reference Guide* for DBCS processing, but the functionality is different. See "DBCS Compatibility" on page 617.

Syntax

COUNT(*string, substring* <,*modifiers*>)

Arguments

string
specifies a character constant, variable, or expression in which substrings are to be counted.

Tip: Enclose a literal string of characters in quotation marks.

substring
is a character constant, variable, or expression that specifies the substring of characters to count in *string*.

Tip: Enclose a literal string of characters in quotation marks.

modifiers
is a character constant, variable, or expression that specifies one or more modifiers. The following *modifiers* can be in uppercase or lowercase:

i ignores character case during the count. If this modifier is not specified, COUNT only counts character substrings with the same case as the characters in *substring*.

t trims trailing blanks from *string* and *substring*.

Tip: If the *modifier* is a constant, enclose it in quotation marks. Specify multiple constants in a single set of quotation marks. *Modifier* can also be expressed as a variable or an expression.

Details

The Basics

The COUNT function searches *string*, from left to right, for the number of occurrences of the specified *substring*, and returns that number of occurrences. If the substring is not found in *string*, COUNT returns a value of 0.

CAUTION:
If two occurrences of the specified substring overlap in the string, the result is undefined. For example, COUNT('boobooboo', 'booboo') might return either a 1 or a 2. △

DBCS Compatibility

You can use the KCOUNT function, which is documented in *SAS National Language Support (NLS): Reference Guide*, for DBCS processing, but the functionality is different.

If the value of *substring* in the COUNT function is longer than two bytes, then the COUNT function can handle DBCS strings. The following examples show the differences in syntax:

COUNT(*string*, *substring* <,*modifiers*>

KCOUNT(*string)*

Comparisons

The COUNT function counts substrings of characters in a character string, whereas the COUNTC function counts individual characters in a character string.

Examples

The following SAS statements produce these results:

SAS Statements	Results

`xyz='This is a thistle? Yes, this is a thistle.';`
`howmanythis=count(xyz,'this');`
`put howmanythis;`
	`3`

`xyz='This is a thistle? Yes, this is a thistle.';`
`howmanyis=count(xyz,'is');`
`put howmanyis;`
	`6`

`howmanythis_i=count('This is a thistle? Yes, this is a thistle.'`
`,'this','i');`
`put howmanythis_i;`
	`4`

SAS Statements	Results

`variable1='This is a thistle? Yes, this is a thistle.';`
`variable2='is ';`
`variable3='i';`
`howmanyis_i=count(variable1,variable2,variable3);`
`put howmanyis_i;`
	4

`expression1='This is a thistle? '||'Yes, this is a thistle.';`
`expression2=kscan('This is',2)||'     ';`
`expression3=compress('i     '||'     t');`
`howmanyis_it=count(expression1,expression2,expression3);`
`put howmanyis_it;`
	6

See Also

Functions:

"COUNTC Function" on page 618
"COUNTW Function" on page 621
"FIND Function" on page 735
"INDEX Function" on page 817

COUNTC Function

Counts the number of characters in a string that appear or do not appear in a list of characters.

Category: Character
Restriction: "I18N Level 1" on page 314

Syntax

COUNTC(*string*, *charlist* <,*modifiers*>)

Arguments

string
specifies a character constant, variable, or expression in which characters are counted.

Tip: Enclose a literal string of characters in quotation marks.

charlist
specifies a character constant, variable, or expression that initializes a list of characters. COUNTC counts characters in this list, provided that you do not specify the V modifier in the *modifier* argument. If you specify the V modifier, then all characters that are not in this list are counted. You can add more characters to the list by using other modifiers.

Tip: Enclose a literal string of characters in quotation marks.

Tip: If there are no characters in the list after processing the modifiers, COUNTC returns 0.

modifier
specifies a character constant, variable, or expression in which each non-blank character modifies the action of the COUNTC function. Blanks are ignored. The following characters, in uppercase or lowercase, can be used as modifiers:

blank	is ignored.
a or A	adds alphabetic characters to the list of characters.
b or B	scans *string* from right to left, instead of from left to right.
c or C	adds control characters to the list of characters.
d or D	adds digits to the list of characters.
f or F	adds an underscore and English letters (that is, the characters that can begin a SAS variable name using VALIDVARNAME=V7) to the list of characters.
g or G	adds graphic characters to the list of characters.
h or H	adds a horizontal tab to the list of characters.
i or I	ignores case.
l or L	adds lowercase letters to the list of characters.
n or N	adds digits, an underscore, and English letters (that is, the characters that can appear in a SAS variable name using VALIDVARNAME=V7) to the list of characters.
o or O	processes the *charlist* and *modifier* arguments only once, at the first call to this instance of COUNTC. If you change the value of *charlist* or *modifier* in subsequent calls, the change might be ignored by COUNTC.
p or P	adds punctuation marks to the list of characters.
s or S	adds space characters to the list of characters (blank, horizontal tab, vertical tab, carriage return, line feed, and form feed).
t or T	trims trailing blanks from *string* and *chars*. **Tip:** If you want to remove trailing blanks from only one character argument instead of both (or all) character arguments, use the TRIM function instead of the COUNTC function with the T modifier.
u or U	adds uppercase letters to the list of characters.
v or V	counts characters that do not appear in the list of characters. If you do not specify this modifier, then COUNTC counts characters that do appear in the list of characters.
w or W	adds printable characters to the list of characters.
x or X	adds hexadecimal characters to the list of characters.

Tip: If *modifier* is a constant, enclose it in quotation marks. Specify multiple constants in a single set of quotation marks.

Details

The COUNTC function allows character arguments to be null. Null arguments are treated as character strings with a length of zero. If there are no characters in the list of characters to be counted, COUNTC returns zero.

Comparisons

The COUNTC function counts individual characters in a character string, whereas the COUNT function counts substrings of characters in a character string.

Examples

The following example uses the COUNTC function with and without modifiers to count the number of characters in a string.

```
data test;
   string  = 'Baboons Eat Bananas     ';
   a       = countc(string, 'a');
   b       = countc(string,'b');
   b_i     = countc(string,'b','i');
   abc_i   = countc(string,'abc','i');
      /* Scan string for characters that are not "a", "b", */
      /* and "c", ignore case, (and include blanks).       */
   abc_iv  = countc(string,'abc','iv');
      /* Scan string for characters that are not "a", "b", */
      /* and "c", ignore case, and trim trailing blanks.   */
   abc_ivt = countc(string,'abc','ivt');
run;

options pageno=1 ls=80 nodate;
proc print data=test noobs;
run;
```

Output 4.43 Output from Using the COUNTC Functions with and without Modifiers

```
                              The SAS System                                1

            string           a    b    b_i    abc_i    abc_iv    abc_ivt

      Baboons Eat Bananas    5    1     3       8        16         11
```

See Also

Functions:

"ANYALNUM Function" on page 379
"ANYALPHA Function" on page 381
"ANYCNTRL Function" on page 383
"ANYDIGIT Function" on page 384
"ANYGRAPH Function" on page 388
"ANYLOWER Function" on page 390
"ANYPRINT Function" on page 394
"ANYPUNCT Function" on page 396
"ANYSPACE Function" on page 397
"ANYUPPER Function" on page 399
"ANYXDIGIT Function" on page 401

"NOTALNUM Function" on page 948
"NOTALPHA Function" on page 950
"NOTCNTRL Function" on page 952
"NOTDIGIT Function" on page 954
"NOTGRAPH Function" on page 959
"NOTLOWER Function" on page 961
"NOTPRINT Function" on page 965
"NOTPUNCT Function" on page 967
"NOTSPACE Function" on page 969
"NOTUPPER Function" on page 971
"NOTXDIGIT Function" on page 973
"FINDC Function" on page 737
"INDEXC Function" on page 819
"VERIFY Function" on page 1193

COUNTW Function

Counts the number of words in a character string.

Category: Character

Syntax

COUNTW(<*string*><, *chars*><, *modifiers*>)

Arguments

string
specifies a character constant, variable, or expression in which words are counted.

chars
specifies an optional character constant, variable, or expression that initializes a list of characters. The characters in this list are the delimiters that separate words, provided that you do not use the K modifier in the *modifier* argument. If you specify the K modifier, then all characters that are not in this list are delimiters. You can add more characters to the list by using other modifiers.

modifier
specifies a character constant, variable, or expression in which each non-blank character modifies the action of the COUNTW function. The following characters, in uppercase or lowercase, can be used as modifiers:

blank	is ignored.
a or A	adds alphabetic characters to the list of characters.
b or B	counts from right to left instead of from left to right. Right-to-left counting makes a difference only when you use the Q modifier and the string contains unbalanced quotation marks.

c or C	adds control characters to the list of characters.
d or D	adds digits to the list of characters.
f or F	adds an underscore and English letters (that is, the characters that can begin a SAS variable name using VALIDVARNAME=V7) to the list of characters.
g or G	adds graphic characters to the list of characters.
h or H	adds a horizontal tab to the list of characters.
i or I	ignores the case of the characters.
k or K	causes all characters that are not in the list of characters to be treated as delimiters. If K is not specified, then all characters that are in the list of characters are treated as delimiters.
l or L	adds lowercase letters to the list of characters.
m or M	specifies that multiple consecutive delimiters, and delimiters at the beginning or end of the *string* argument, refer to words that have a length of zero. If the M modifier is not specified, then multiple consecutive delimiters are treated as one delimiter, and delimiters at the beginning or end of the *string* argument are ignored.
n or N	adds digits, an underscore, and English letters (that is, the characters that can appear after the first character in a SAS variable name using VALIDVARNAME=V7) to the list of characters.
o or O	processes the *chars* and *modifier* arguments only once, rather than every time the COUNTW function is called. Using the O modifier in the DATA step (excluding WHERE clauses), or in the SQL procedure, can make COUNTW run faster when you call it in a loop where *chars* and *modifier* arguments do not change.
p or P	adds punctuation marks to the list of characters.
q or Q	ignores delimiters that are inside of substrings that are enclosed in quotation marks. If the value of *string* contains unmatched quotation marks, then scanning from left to right will produce different words than scanning from right to left.
s or S	adds space characters (blank, horizontal tab, vertical tab, carriage return, line feed, and form feed) to the list of characters.
t or T	trims trailing blanks from the *string* and *chars* arguments.
u or U	adds uppercase letters to the list of characters.
w or W	adds printable characters to the list of characters.
x or X	adds hexadecimal characters to the list of characters.

Details

Definition of "Word" In the COUNTW function, "word" refers to a substring that has one of the following characteristics:

- is bounded on the left by a delimiter or the beginning of the string
- is bounded on the right by a delimiter or the end of the string

- contains no delimiters, except if you use the Q modifier and the delimiters are within substrings that have quotation marks

Note: The definition of "word" is the same in both the SCAN function and the COUNTW.sgml function. △

Delimiter refers to any of several characters that you can specify to separate words.

Using the COUNTW Function in ASCII and EBCDIC Environments If you use the COUNTW function with only two arguments, the default delimiters depend on whether your computer uses ASCII or EBCDIC characters.

- If your computer uses ASCII characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ^ |

  In ASCII environments that do not contain the ^ character, the SCAN function uses the ~ character instead.

- If your computer uses EBCDIC characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ¬ | ¢

Using Null Arguments The COUNTW function allows character arguments to be null. Null arguments are treated as character strings with a length of zero. Numeric arguments cannot be null.

Using the M Modifier

If you do not use the M modifier, then a word must contain at least one character. If you use the M modifier, then a word can have a length of zero. In this case, the number of words is one plus the number of delimiters in the string, not counting delimiters inside of strings that are enclosed in quotation marks when you use the Q modifier.

Examples

The following example shows how to use the COUNTW function with the M and P modifiers.

```
options ls=64 pageno=1 nodate;
data test;
   length default blanks mp 8;
   input string $char60.;
   default = countw(string);
   blanks = countw(string, ' ');
   mp = countw(string, 'mp');
   datalines;
The quick brown fox jumps over the lazy dog.
        Leading blanks
2+2=4
/unix/path/names/use/slashes
\Windows\Path\Names\Use\Backslashes
;
run;

proc print noobs data=test;
run;
```

Output 4.44 Output from the COUNTW Function

```
                          The SAS System                          1

default blanks mp                    string

   9       9    2 The quick brown fox jumps over the lazy dog.
   2       2    1         Leading blanks
   2       1    1 2+2=4
   5       1    3 /unix/path/names/use/slashes
   1       1    2 \Windows\Path\Names\Use\Backslashes
```

See Also

Functions and CALL Routines:

"COUNT Function" on page 616
"COUNTC Function" on page 618
"FINDW Function" on page 743
"SCAN Function" on page 1111
"CALL SCAN Routine" on page 516

CSS Function

Returns the corrected sum of squares.

Category: Descriptive Statistics

Syntax

CSS(*argument-1*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least one nonmissing argument is required. Otherwise, the function returns a missing value. If you have more than one argument, the argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=css(5,9,3,6);`	`18.75`
`x2=css(5,8,9,6,.);`	`10`
`x3=css(8,9,6,.);`	`4.6666666667`
`x4=css(of x1-x3);`	`101.11574074`

CUROBS Function

Returns the observation number of the current observation.

Category: SAS File I/O

Requirement: Use this function only with an uncompressed SAS data set that is accessed using a native library engine.

Syntax

CUROBS(*data-set-id*)

Arguments

data-set-id
is a numeric value that specifies the data set identifier that the OPEN function returns.

Details

If the engine being used does not support observation numbers, the function returns a missing value.

With a SAS view, the function returns the relative observation number, that is, the number of the observation within the SAS view (as opposed to the number of the observation within any related SAS data set).

Examples

This example uses the FETCHOBS function to fetch the tenth observation in the data set MYDATA. The value of OBSNUM returned by CUROBS is 10.

```
%let dsid=%sysfunc(open(mydata,i));
%let rc=%sysfunc(fetchobs(&dsid,10));
%let obsnum=%sysfunc(curobs(&dsid));
```

See Also

Functions:

"FETCHOBS Function" on page 685
"OPEN Function" on page 980

CV Function

Returns the coefficient of variation.

Category: Descriptive Statistics

Syntax

CV(*argument-1,argument-2<,...argument-n>*)

Arguments

argument
specifies a numeric constant, variable, or expression. At least two arguments are required. The argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=cv(5,9,3,6);`	`43.47826087`
`x2=cv(5,8,9,6,.);`	`26.082026548`
`x3=cv(8,9,6,.);`	`19.924242152`
`x4=cv(of x1-x3);`	`40.953539216`

DACCDB Function

Returns the accumulated declining balance depreciation.

Category: Financial

Syntax

DACCDB(*p,v,y,r*)

Arguments

p
is numeric, the period for which the calculation is to be done. For noninteger *p* arguments, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

y
is numeric, the lifetime of the asset.
Range: $y > 0$

r
is numeric, the rate of depreciation expressed as a decimal.
Range: $r > 0$

Details

The DACCDB function returns the accumulated depreciation by using a declining balance method. The formula is

$$\mathrm{DACCDB}(p, v, y, r) = \begin{cases} 0 & p \le 0 \\ v\left(1 - \left(1 - \frac{r}{y}\right)^{int(p)}\right)\left(1 - (p - int\,(p))\,\frac{r}{y}\right) & p > 0 \end{cases}$$

Note that int(*p*) is the integer part of *p*. The *p* and *y* arguments must be expressed by using the same units of time. A double-declining balance is obtained by setting *r* equal to 2.

Examples

An asset has a depreciable initial value of $1000 and a fifteen-year lifetime. Using a 200 percent declining balance, the depreciation throughout the first 10 years can be expressed as

```
a=daccdb(10,1000,15,2);
```

The value returned is 760.93. The first and the third arguments are expressed in years.

DACCDBSL Function

Returns the accumulated declining balance with conversion to a straight-line depreciation.

Category: Financial

Syntax

DACCDBSL(*p,v,y,r*)

Arguments

p
is numeric, the period for which the calculation is to be done.

v
is numeric, the depreciable initial value of the asset.

y
is an integer, the lifetime of the asset.
Range: $y > 0$

r
is numeric, the rate of depreciation that is expressed as a fraction.
Range: $r > 0$

Details

The DACCDBSL function returns the accumulated depreciation by using a declining balance method, with conversion to a straight-line depreciation function that is defined by

$$\mathrm{DACCDBSL}(p, v, y, r) = \sum_{i=1}^{p} \mathrm{DEPDBSL}(i, v, y, r)$$

The declining balance with conversion to a straight-line depreciation chooses for each time period the method of depreciation (declining balance or straight-line on the remaining balance) that gives the larger depreciation. The *p* and *y* arguments must be expressed by using the same units of time.

Examples

An asset has a depreciable initial value of $1,000 and a ten-year lifetime. Using a declining balance rate of 150 percent, the accumulated depreciation of that asset in its fifth year can be expressed as

```
y5=daccdbsl(5,1000,10,1.5);
```

The value returned is 564.99. The first and the third arguments are expressed in years.

DACCSL Function

Returns the accumulated straight-line depreciation.

Category: Financial

Syntax

DACCSL(*p,v,y*)

Arguments

p
is numeric, the period for which the calculation is to be done. For fractional p, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

y
is numeric, the lifetime of the asset.

Range: $y > 0$

Details

The DACCSL function returns the accumulated depreciation by using the straight-line method, which is given by

$$\mathrm{DACCSL}(p,v,y) = \begin{cases} 0 & p < 0 \\ v\left(\frac{p}{y}\right) & 0 \le p \le y \\ v & p > y \end{cases}$$

The p and y arguments must be expressed by using the same units of time.

Examples

An asset, acquired on 01APR86, has a depreciable initial value of $1000 and a ten-year lifetime. The accumulated depreciation in the value of the asset through 31DEC87 can be expressed as

```
a=daccsl(1.75,1000,10);
```

The value returned is 175.00. The first and the third arguments are expressed in years.

DACCSYD Function

Returns the accumulated sum-of-years-digits depreciation.

Category: Financial

Syntax

DACCSYD(*p,v,y*)

Arguments

p
is numeric, the period for which the calculation is to be done. For noninteger *p* arguments, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

y
is numeric, the lifetime of the asset.

Range: $y > 0$

Details

The DACCSYD function returns the accumulated depreciation by using the sum-of-years-digits method. The formula is

$$\text{DACCSYD}\,(p,v,y) = \begin{cases} 0 & p < 0 \\ v\dfrac{int(p)\left(y-\frac{int(p)-1}{2}\right)+(p-int(p))(y-int(p))}{int(y)\left(y-\frac{int(y)-1}{2}\right)+(y-int(y))^2} & 0 \le p \le y \\ v & p > y \end{cases}$$

Note that int(*y*) indicates the integer part of *y*. The *p* and *y* arguments must be expressed by using the same units of time.

Examples

An asset, acquired on 01OCT86, has a depreciable initial value of $1,000 and a five-year lifetime. The accumulated depreciation of the asset throughout 01JAN88 can be expressed as

```
y2=daccsyd(15/12,1000,5);
```

The value returned is 400.00. The first and the third arguments are expressed in years.

DACCTAB Function

Returns the accumulated depreciation from specified tables.

Category: Financial

Syntax

DACCTAB(*p,v,t1, . . . ,tn*)

Arguments

p
is numeric, the period for which the calculation is to be done. For noninteger *p* arguments, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

t1,t2, . . . ,tn
are numeric, the fractions of depreciation for each time period with *t1*+*t2*+...*tn*≤1.

Details

The DACCTAB function returns the accumulated depreciation by using user-specified tables. The formula for this function is

$$DACCTAB\left(p,v,t_1,t_2,...,t_n\right)=\begin{cases}0 & p\leq 0\\ v\left(t_1+t_2+...+t_{int(p)}+\left(p-int\left(p\right)\right)t_{int(p)+1}\right) & 0<p<n\\ v & p\geq n\end{cases}$$

For a given *p*, only the arguments $t_1, t_2, \ldots, t_k$ need to be specified with **k=ceil(p)**.

Examples

An asset has a depreciable initial value of $1000 and a five-year lifetime. Using a table of the annual depreciation rates of .15, .22, .21, .21, and .21 during the first, second, third, fourth, and fifth years, respectively, the accumulated depreciation throughout the third year can be expressed as

```
y3=dacctab(3,1000,.15,.22,.21,.21,.21);
```

The value that is returned is 580.00. The fourth rate, .21, and the fifth rate, .21, can be omitted because they are not needed in the calculation.

DAIRY Function

Returns the derivative of the AIRY function.

Category: Mathematical

Syntax

DAIRY(*x*)

Arguments

x
specifies a numeric constant, variable, or expression.

Details

The DAIRY function returns the value of the derivative of the AIRY function (Abramowitz and Stegun 1964; Amos, Daniel, and Weston 1977).

Examples

SAS Statements	Results
`x=dairy(2.0);`	`-0.053090384`
`x=dairy(-2.0);`	`0.6182590207`

DATDIF Function

Returns the number of days between two dates after computing the difference between the dates according to specified day count conventions.

Category: Date and Time

Syntax

DATDIF(*sdate*,*edate*,*basis*)

Arguments

sdate
specifies a SAS date value that identifies the starting date.

Tip: If *sdate* falls at the end of a month, then SAS treats the date as if it were the last day of a 30-day month.

edate
specifies a SAS date value that identifies the ending date.

Tip: If *edate* falls at the end of a month, then SAS treats the date as if it were the last day of a 30-day month.

basis
specifies a character string that represents the day count basis. The following values for *basis* are valid:

'30/360'
specifies a 30-day month and a 360-day year, regardless of the actual number of calendar days in a month or year.

A security that pays interest on the last day of a month will either always make its interest payments on the last day of the month, or it will always make its payments on the numerically same day of a month, unless that day is not a valid day of the month, such as February 30. For more information, see "Method of Calculation for Day Count Basis (30/360)" on page 633.

Alias: '360'

'ACT/ACT'
uses the actual number of days between dates. Each month is considered to have the actual number of calendar days in that month, and each year is considered to have the actual number of calendar days in that year.

Alias: *'Actual'*

'ACT/360'
uses the actual number of calendar days in a particular month, and 360 days as the number of days in a year, regardless of the actual number of days in a year.

Tip: *ACT/360* is used for short-term securities.

'ACT/365'
uses the actual number of calendar days in a particular month, and 365 days as the number of days in a year, regardless of the actual number of days in a year.

Tip: *ACT/365* is used for short-term securities.

Details

The Basics

The DATDIF function has a specific meaning in the securities industry, and the method of calculation is not the same as the actual day count method. Calculations can use months and years that contain the actual number of days. Calculations can also be based on a 30-day month or a 360-day year. For more information about standard securities calculation methods, see "References" on page 634.

Note: When counting the number of days in a month, DATDIF *always* includes the starting date and excludes the ending date. △

Method of Calculation for Day Count Basis (30/360) To calculate the number of days between two dates, use the following formula:

$$Numberofdays = [(Y2 - Y1) * 360] + [(M2 - M1) * 30] + (D2 - D1)$$

where

Y2	specifies the year of the later date.
Y1	specifies the year of the earllier date.
M2	specifies the month of the later date.
M1	specifies the month of the earlier date.
D2	specifies the day of the later date.
D1	specifies the day of the earlier date.

Because all months can contain only 30 days, you must adjust for the months that do not contain 30 days. Do this before you calculate the number of days between the two dates.

The following rules apply:

- If the security follows the End-of-Month rule, and D2 is the last day of February (28 days in a non-leap year, 29 days in a leap year), and D1 is the last day of February, then change D2 to 30.
- If the security follows the End-of-Month rule, and D1 is the last day of February, then change D1 to 30.
- If the value of D2 is 31 and the value of D1 is 30 or 31, then change D2 to 30.
- If the value of D1 is 31, then change D1 to 30.

Examples

In the following example, DATDIF returns the actual number of days between two dates, as well as the number of days based on a 30-day month and a 360-day year.

```
data _null;
   sdate='16oct78'd;
   edate='16feb96'd;
   actual=datdif(sdate, edate, 'act/act');
   days360=datdif(sdate, edate, '30/360');
   put actual= days360=;
run;
```

SAS Statements	Results
put actual=;	**6332**
put days360=;	**6240**

See Also

Functions:

"YRDIF Function" on page 1235

References

Securities Industry Association. 1994. *Standard Securities Calculation Methods - Fixed Income Securities Formulas for Analytic Measures, Volume 2*. New York: Securities Industry Association.

DATE Function

Returns the current date as a SAS date value.

Category: Date and Time

Alias: TODAY

See: "TODAY Function" on page 1165

Syntax

DATE()

DATEJUL Function

Converts a Julian date to a SAS date value.

Category: Date and Time

Syntax

DATEJUL(*julian-date*)

Arguments

julian-date
specifies a SAS numeric expression that represents a Julian date. A Julian date in SAS is a date in the form *yyddd* or *yyyyddd*, where *yy* or *yyyy* is a two-digit or four-digit integer that represents the year and *ddd* is the number of the day of the year. The value of *ddd* must be between 1 and 365 (or 366 for a leap year).

Examples

The following SAS statements produce these results:

SAS Statements	Results
`Xstart=datejul(94365);` `put Xstart / Xstart date9.;`	`12783` `31DEC1994`
`Xend=datejul(2001001);` `put Xend / Xend date9.;`	`14976` `01JAN2001`

See Also

Function:

"JULDATE Function" on page 866

DATEPART Function

Extracts the date from a SAS datetime value.

Category: Date and Time

Syntax

DATEPART(*datetime*)

Arguments

datetime
specifies a SAS expression that represents a SAS datetime value.

Examples

The following SAS statements produce this result:

SAS Statements	Results

`conn='01feb94:8:45'dt;`
`servdate=datepart(conn);`
`put servdate worddate.;`
	`February 1, 1994`

See Also

Functions:

"DATETIME Function" on page 637
"TIMEPART Function" on page 1162

DATETIME Function

Returns the current date and time of day as a SAS datetime value.

Category: Date and Time

Syntax

DATETIME()

Examples

This example returns a SAS value that represents the number of seconds between January 1, 1960 and the current time:

```
when=datetime();
put when=;
```

See Also

Functions:

"DATE Function" on page 634
"TIME Function" on page 1161

DAY Function

Returns the day of the month from a SAS date value.

Category: Date and Time

Syntax

DAY(*date*)

Arguments

date
specifies a SAS expression that represents a SAS date value.

Details

The DAY function produces an integer from 1 to 31 that represents the day of the month.

Examples

The following SAS statements produce this result:

SAS Statements	Results

`now='05may97'd;`
`d=day(now);`
`put d;`
	`5`

See Also

Functions:

"MONTH Function" on page 936
"YEAR Function" on page 1233

DCLOSE Function

Closes a directory that was opened by the DOPEN function.

Category: External Files

Syntax

DCLOSE(*directory-id*)

Argument

directory-id
is a numeric variable that specifies the identifier that was assigned when the directory was opened by the DOPEN function.

Details

DCLOSE returns 0 if the operation was successful, ≠0 if it was not successful. The DCLOSE function closes a directory that was previously opened by the DOPEN function. DCLOSE also closes any open members.

Note: All directories or members opened within a DATA step are closed automatically when the DATA step ends. △

Examples

Example 1: Using DCLOSE to Close a Directory This example opens the directory to which the fileref MYDIR has previously been assigned, returns the number of members, and then closes the directory:

```
%macro memnum(filrf,path);
%let rc=%sysfunc(filename(filrf,&path));
%if %sysfunc(fileref(&filrf)) = 0 %then
   %do;
         /* Open the directory. */
      %let did=%sysfunc(dopen(&filrf));
      %put did=&did;
         /* Get the member count. */
      %let memcount=%sysfunc(dnum(&did));
      %put &memcount members in &filrf.;
         /* Close the directory. */
      %let rc= %sysfunc(dclose(&did));
   %end;
%else %put Invalid FILEREF;
%mend;
%memnum(MYDIR,physical-filename)
```

Example 2: Using DCLOSE within a DATA Step This example uses the DCLOSE function within a DATA step:

```
%let filrf=MYDIR;
data _null_;
  rc=filename("&filrf","physical-filename");
  if fileref("&filrf") = 0 then
     do;
           /* Open the directory. */
        did=dopen("&filrf");
           /* Get the member count. */
        memcount=dnum(did);
        put memcount "members in &filrf";
           /* Close the directory. */
        rc=dclose(did);
     end;
  else put "Invalid FILEREF";
run;
```

See Also

Functions:

"DOPEN Function" on page 661
"FCLOSE Function" on page 680
"FOPEN Function" on page 762
"MOPEN Function" on page 936

DCREATE Function

Returns the complete pathname of a new, external directory.

Category: External Files

Syntax

DCREATE(*directory-name*<,*parent-directory*>)

Arguments

directory-name
is a character constant, variable, or expression that specifies the name of the directory to create. This value cannot include a pathname.

parent-directory
is a character constant, variable, or expression that contains the complete pathname of the directory in which to create the new directory. If you do not supply a value for *parent-directory*, then the current directory is the parent directory.

Details

The DCREATE function enables you to create a directory in your operating environment. If the directory cannot be created, then DCREATE returns an empty string.

Examples

To create a new directory in the UNIX operating environment, using the name that is stored in the variable DirectoryName, follow this form:

```
NewDirectory=dcreate(DirectoryName,'/local/u/abcdef/');
```

To create a new directory in the Windows operating environment, using the name that is stored in the variable DirectoryName, follow this form:

```
NewDirectory=dcreate(DirectoryName,'d:\testdir\');
```

DEPDB Function

Returns the declining balance depreciation.

Category: Financial

Syntax

DEPDB(*p,v,y,r*)

Arguments

p
is numeric, the period for which the calculation is to be done. For noninteger *p* arguments, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

y
is numeric, the lifetime of the asset.

Range: $y > 0$

r
is numeric, the rate of depreciation that is expressed as a fraction.

Range: $r \geq 0$

Details

The DEPDB function returns the depreciation by using the declining balance method, which is given by

$$\begin{aligned}\mathrm{DEPDB}\,(p,v,y,r) = {} & \mathrm{DACCDB}\,(p,v,y,r) \\ & - \mathrm{DACCDB}\,(p-1,v,y,r)\end{aligned}$$

The *p* and *y* arguments must be expressed by using the same units of time. A double-declining balance is obtained by setting *r* equal to 2.

Examples

An asset has an initial value of $1,000 and a fifteen-year lifetime. Using a declining balance rate of 200 percent, the depreciation of the value of the asset for the tenth year can be expressed as

```
y10=depdb(10,1000,15,2);
```

The value returned is 36.78. The first and the third arguments are expressed in years.

DEPDBSL Function

Returns the declining balance with conversion to a straight-line depreciation.

Category: Financial

Syntax

DEPDBSL(*p,v,y,r*)

Arguments

p
is an integer, the period for which the calculation is to be done.

v
is numeric, the depreciable initial value of the asset.

y
is an integer, the lifetime of the asset.

Range: $y > 0$

r
is numeric, the rate of depreciation that is expressed as a fraction.

Range: $r \geq 0$

Details

The DEPDBSL function returns the depreciation by using the declining balance method with conversion to a straight-line depreciation, which is given by the following equation:

$$\mathrm{DEPDBSL}\,(p,v,y,r) = \begin{cases} 0 & p \leq 0 \\ v\frac{r}{y}\left(1-\frac{r}{y}\right)^{p-1} & 0 < p \leq t \\ \frac{v\left(1-\frac{r}{y}\right)^{t}}{(y-t)} & t < p \leq y \\ 0 & p > y \end{cases}$$

where

$$t = int\left(y - \frac{y}{r} + 1\right)$$

and *int*( )denotes the integer part of a numeric argument.

The *p* and *y* arguments must be expressed by using the same units of time. The declining balance that changes to a straight-line depreciation chooses for each time period the method of depreciation (declining balance or straight-line on the remaining balance) that gives the larger depreciation.

Examples

An asset has a depreciable initial value of $1,000 and a ten-year lifetime. Using a declining balance rate of 150 percent, the depreciation of the value of the asset in the fifth year can be expressed as

```
y5=depdbsl(5,1000,10,1.5);
```

The value 87.001041667 is returned. The first and the third arguments are expressed in years.

DEPSL Function

Returns the straight-line depreciation.

Category: Financial

Syntax

DEPSL(*p,v,y*)

Arguments

p
is numeric, the period for which the calculation is to be done. For fractional *p*, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

y
is numeric, the lifetime of the asset.

Range: $y > 0$

Details

The DEPSL function returns the straight-line depreciation, which is given by

$$\begin{aligned}\mathrm{DEPSL}(p,v,y) = {} & \mathrm{DACCSL}(p,v,y) \\ & - \mathrm{DACCSL}(p-1,v,y)\end{aligned}$$

The *p* and *y* arguments must be expressed by using the same units of time.

Examples

An asset, acquired on 01APR86, has a depreciable initial value of $1,000 and a ten-year lifetime. The depreciation in the value of the asset for the year 1986 can be expressed as

```
d=depsl(9/12,1000,10);
```

The value returned is 75.00. The first and the third arguments are expressed in years.

DEPSYD Function

Returns the sum-of-years-digits depreciation.

Category: Financial

Syntax

DEPSYD(*p,v,y*)

Arguments

p
is numeric, the period for which the calculation is to be done. For noninteger *p* arguments, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

y
is numeric, the lifetime of the asset in number of depreciation periods.
Range: $y > 0$

Details

The DEPSYD function returns the sum-of-years-digits depreciation, which is given by

$$\begin{aligned}\mathrm{DEPSYD}\,(p,v,y) = {} & \mathrm{DACCSYD}\,(p,v,y) \\ & - \mathrm{DACCSYD}\,(p-1,v,y)\end{aligned}$$

The *p* and *y* arguments must be expressed by using the same units of time.

Examples

An asset, acquired on 01OCT86, has a depreciable initial value of $1,000 and a five-year lifetime. The depreciations in the value of the asset for the years 1986 and 1987 can be expressed as

```
y1=depsyd(3/12,1000,5);
y2=depsyd(15/12,1000,5);
```

The values returned are 83.33 and 316.67, respectively. The first and the third arguments are expressed in years.

DEPTAB Function

Returns the depreciation from specified tables.

Category: Financial

Syntax

DEPTAB(*p,v,t1,...,tn*)

Arguments

p
is numeric, the period for which the calculation is to be done. For noninteger *p* arguments, the depreciation is prorated between the two consecutive time periods that precede and follow the fractional period.

v
is numeric, the depreciable initial value of the asset.

t1,t2, . . . ,tn
are numeric, the fractions of depreciation for each time period with *t1+t2+...tn*$\leq$1.

Details

The DEPTAB function returns the depreciation by using specified tables. The formula is

$$DEPTAB\,(p,v,t_1,t_2,...,t_n) = DACCTAB\,(p,v,t_1,t_2,...,t_n) - DACCTAB\,(p-1,v,t_1,t_2,...,t_n)$$

For a given p, only the arguments $t_1, t_2, \ldots, t_k$ need to be specified with **`k=ceil(p)`**.

Examples

An asset has a depreciable initial value of $1,000 and a five-year lifetime. Using a table of the annual depreciation rates of .15, .22, .21, .21, and .21 during the first, second, third, fourth, and fifth years, respectively, the depreciation in the third year can be expressed as

$$y3 = deptab\,(3,1000,.15,.22,.21,.21,.21\,);$$

The value that is returned is 210.00. The fourth rate, .21, and the fifth rate, .21, can be omitted because they are not needed in the calculation.

DEQUOTE Function

Removes matching quotation marks from a character string that begins with a quotation mark, and deletes all characters to the right of the closing quotation mark.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

DEQUOTE(*string*)

Arguments

string
specifies a character constant, variable, or expression.

Details

Length of Returned Variable In a DATA step, if the DEQUOTE function returns a value to a variable that has not been previously assigned a length, then that variable is given the length of the argument.

The Basics The value that is returned by the DEQUOTE function is determined as follows:

- If the first character of *string* is not a single or double quotation mark, DEQUOTE returns *string* unchanged.
- If the first two characters of *string* are both single quotation marks or both double quotation marks, and the third character is not the same type of quotation mark, then DEQUOTE returns a result with a length of zero.
- If the first character of *string* is a single quotation mark, the DEQUOTE function removes that single quotation mark from the result. DEQUOTE then scans *string* from left to right, looking for more single quotation marks. Each pair of consecutive, single quotation marks is reduced to one single quotation mark. The first single quotation mark that does not have an ending quotation mark in *string* is removed and all characters to the right of that quotation mark are also removed.
- If the first character of *string* is a double quotation mark, the DEQUOTE function removes that double quotation mark from the result. DEQUOTE then scans *string* from left to right, looking for more double quotation marks. Each pair of consecutive, double quotation marks is reduced to one double quotation mark. The first double quotation mark that does not have an ending quotation mark in *string* is removed and all characters to the right of that quotation mark are also removed.

Note: If *string* is a constant enclosed by quotation marks, those quotation marks are not part of the value of *string*. Therefore, you do not need to use DEQUOTE to remove the quotation marks that denote a constant. △

Examples

This example demonstrates the use of DEQUOTE within a DATA step.

```
options pageno=1 nodate ls=80 ps=64;

data test;
   input string $60.;
   result = dequote(string);
   datalines;
No quotation marks, no change
No "leading" quotation marks, no change
"Matching double quotation marks are removed"
'Matching single quotation marks are removed'
"Paired ""quotation marks"" are reduced"
'Paired '' quotation marks '' are reduced'
"Single 'quotation marks' inside '' double'' quotation marks are unchanged"
'Double "quotation marks" inside ""single"" quotation marks are unchanged'
"No matching quotation mark, no problem
Don't remove this apostrophe
"Text after the matching quotation mark" is "deleted"
;

proc print noobs;
title 'Input Strings and Output Results from DEQUOTE';
run;
```

Output 4.45 Removing Matching Quotation Marks with the DEQUOTE Function

```
                Input Strings and Output Results from DEQUOTE                1

          string

          No quotation marks, no change
          No "leading" quotation marks, no change
          "Matching double quotation marks are removed"
          'Matching single quotation marks are removed'
          "Paired ""quotation marks"" are reduced"
          'Paired '' quotation marks '' are reduced'
          "Single 'quotation marks' inside '' double'' quotation marks
          'Double "quotation marks" inside ""single"" quotation marks
          "No matching quotation mark, no problem
          Don't remove this apostrophe
          "Text after the matching quotation mark" is "deleted"

          result

          No quotation marks, no change
          No "leading" quotation marks, no change
          Matching double quotation marks are removed
          Matching single quotation marks are removed
          Paired "quotation marks" are reduced
          Paired ' quotation marks ' are reduced
          Single 'quotation marks' inside '' double'' quotation marks
          Double "quotation marks" inside ""single"" quotation marks
          No matching quotation mark, no problem
          Don't remove this apostrophe
          Text after the matching quotation mark
```

DEVIANCE Function

Returns the deviance based on a probability distribution.

Category: Mathematical

Syntax

DEVIANCE(*distribution*, *variable*, *shape-parameters*<,ε>)

Arguments

distribution
is a character constant, variable, or expression that identifies the distribution. Valid distributions are listed in the following table:

Distribution	Argument
Bernoulli	**'BERNOULLI' | 'BERN'**
Binomial	**'BINOMIAL' | 'BINO'**
Gamma	**'GAMMA'**
Inverse Gauss (Wald)	**'IGAUSS' | 'WALD'**
Normal	**'NORMAL' | 'GAUSSIAN'**
Poisson	**'POISSON' | 'POIS'**

variable
is a numeric constant, variable, or expression.

shape-parameter
are one or more distribution-specific numeric parameters that characterize the shape of the distribution.

ε
is an optional numeric small value used for all of the distributions, except for the normal distribution.

Details

The Bernoulli Distribution

DEVIANCE('BERNOULLI', *variable*, *p*<, ε>)

where

variable
is a binary numeric random variable that has the value of 1 for success and 0 for failure.

p
is a numeric probability of success with $\varepsilon \le p \le 1-\varepsilon$.

ε
is an optional positive numeric value that is used to bound *p*. Any value of *p* in the interval $0 \le p \le \varepsilon$ is replaced by ε. Any value of *p* in the interval $1-\varepsilon \le p \le 1$ is replaced by $1-\varepsilon$.

The DEVIANCE function returns the deviance from a Bernoulli distribution with a probability of success *p*, where success is defined as a random variable value of 1. The equation follows:

$$\text{DEVIANCE}\left('BERN', variable, p, \epsilon\right) = \begin{cases} -2\log\left(1-p\right) & x = 0 \\ -2\log\left(p\right) & x = 1 \\ . & otherwise \end{cases}$$

The Binomial Distribution

DEVIANCE('BINO', *variable*, μ, *n*<, ε>)

where

variable
is a numeric random variable that contains the number of successes.
Range: $0 \le variable \le 1$

μ
is a numeric mean parameter.
Range: $n\varepsilon \le \mu \le n(1-\varepsilon)$

n
is an integer number of Bernoulli trials parameter
Range: $n \ge 0$

ε
is an optional positive numeric value that is used to bound μ. Any value of μ in the interval $0 \le \mu \le n\varepsilon$ is replaced by $n\varepsilon$. Any value of μ in the interval $n(1-\varepsilon) \le \mu \le n$ is replaced by $n(1-\varepsilon)$.

The DEVIANCE function returns the deviance from a binomial distribution, with a probability of success *p*, and a number of independent Bernoulli trials *n*. The following equation describes the DEVIANCE function for the Binomial distribution, where *x* is the random variable.

$$\text{DEVIANCE}\left('BINO', x, \mu, n\right) = \begin{cases} . & x < 0 \\ 2\left(x\log\left(\frac{x}{\mu}\right) + (n-x)\log\left(\frac{n-x}{n-\mu}\right)\right) & 0 \le x \le n \\ . & x > n \end{cases}$$

The Gamma Distribution

DEVIANCE('GAMMA', *variable*, μ <, ε>)

where

variable
is a numeric random variable.

Range: *variable* ≥ ε

μ
is a numeric mean parameter.

Range: $\mu \geq \varepsilon$

ε
is an optional positive numeric value that is used to bound *variable* and μ. Any value of *variable* in the interval 0 ≤ *variable* ≤ ε is replaced by ε. Any value of μ in the interval 0 ≤ μ ≤ ε is replaced by ε.

The DEVIANCE function returns the deviance from a gamma distribution with a mean parameter μ. The following equation describes the DEVIANCE function for the gamma distribution, where *x* is the random variable:

$$\text{DEVIANCE}\left('GAMMA', x, \mu\right) = \begin{cases} . & x < 0 \\ 2\left(-\log\left(\frac{x}{\mu}\right) + \frac{x-\mu}{\mu}\right) & x \geq \epsilon, \mu \geq \epsilon \end{cases}$$

The Inverse Gauss (Wald) Distribution

DEVIANCE('IGAUSS' | 'WALD', *variable*, μ<, ε>)

where

variable
is a numeric random variable.

Range: *variable* ≥ ε

μ
is a numeric mean parameter.

Range: $\mu \geq \varepsilon$

ε
is an optional positive numeric value that is used to bound *variable* and μ. Any value of *variable* in the interval 0 ≤ *variable* ≤ ε is replaced by ε. Any value of μ in the interval 0 ≤ μ ≤ ε is replaced by ε.

The DEVIANCE function returns the deviance from an inverse Gaussian distribution with a mean parameter μ. The following equation describes the DEVIANCE function for the inverse Gaussian distribution, where *x* is the random variable:

$$\text{DEVIANCE}\left('IGAUSS', x, \mu\right) = \begin{cases} . & x < 0 \\ \frac{(x-\mu)^2}{\mu^2 x} & x \geq \epsilon, \mu \geq \epsilon \end{cases}$$

The Normal Distribution

DEVIANCE('NORMAL' | 'GAUSSIAN', *variable*, μ)

where

variable
is a numeric random variable.

μ
is a numeric mean parameter.

The DEVIANCE function returns the deviance from a normal distribution with a mean parameter μ. The following equation describes the DEVIANCE function for the normal distribution, where x is the random variable:

$$\text{DEVIANCE}\left('NORMAL', x, \mu\right) = (x - \mu)^2$$

The Poisson Distribution

DEVIANCE('POISSON', *variable*, μ<, ε>)

where

variable
is a numeric random variable.
Range: *variable* ≥ 0

μ
is a numeric mean parameter.
Range: $\mu \geq \varepsilon$

ε
is an optional positive numeric value that is used to bound μ. Any value of μ in the interval $0 \leq \mu \leq \varepsilon$ is replaced by ε.

The DEVIANCE function returns the deviance from a Poisson distribution with a mean parameter μ. The following equation describes the DEVIANCE function for the Poisson distribution, where x is the random variable:

$$\text{DEVIANCE}\left('POISSON', x, \mu\right) = \begin{cases} . & x < 0 \\ 2\left(x \log\left(\frac{x}{\mu}\right) - (x - \mu)\right) & x \geq 0, \mu \geq \epsilon \end{cases}$$

DHMS Function

Returns a SAS datetime value from date, hour, minute, and second values.

Category: Date and Time

Syntax

DHMS(*date,hour,minute,second*)

Arguments

date
specifies a SAS expression that represents a SAS date value.

hour
is numeric.

minute
is numeric.

second
is numeric.

Details

The DHMS function returns a numeric value that represents a SAS datetime value. This numeric value can be either positive or negative.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`dtid=dhms('01jan03'd,15,30,15);`	
`put dtid;`	`1357054215`
`put dtid datetime.;`	`01JAN03:15:30:15`
`dtid2=dhms('01jan03'd,15,30,61);`	
`put dtid2;`	`1357054261`
`put dtid2 datetime.;`	`01JAN03:15:31:01`
`dtid3=dhms('01jan03'd,15,.5,15);`	
`put dtid3;`	`1357052445`
`put dtid3 datetime.;`	`01JAN03:15:00:45`

The following SAS statements show how to combine a SAS date value with a SAS time value into a SAS datetime value. If you execute these statements on April 2, 2003 at the time of 15:05:02, it produces these results:

SAS Statements	Result
`day=date();`	
`time=time();`	
`sasdt=dhms(day,0,0,time);`	
`put sasdt datetime.;`	`02APR03:15:05:02`

See Also

Function:

"HMS Function" on page 803

DIF Function

Returns differences between an argument and its *n*th lag.

Category: Special

Syntax

DIF<*n*>(*argument*)

Arguments

n
specifies the number of lags.

argument
specifies a numeric constant, variable, or expression.

Details

The DIF functions, DIF1, DIF2, ..., DIF100, return the first differences between the argument and its *n*th lag. DIF1 can also be written as DIF. DIF*n* is defined as DIF*n*(*x*)=*x*-LAG*n*(*x*).

For details on storing and returning values from the LAG*n* queue, see the LAG function.

Comparisons

The function DIF2(X) is not equivalent to the second difference DIF(DIF(X)).

Examples

This example demonstrates the difference between the LAG and DIF functions.

```
data two;
   input X @@;
   Z=lag(x);
   D=dif(x);
   datalines;
1 2 6 4 7
;
proc print data=two;
run;
```

Results of the PROC PRINT step follow:

```
OBS         X              Z                 D
 1          1              .                 .
 2          2              1                 1
 3          6              2                 4
 4          4              6               - 2
 5          7              4                 3
```

See Also

Function:

"LAG Function" on page 870

DIGAMMA Function

Returns the value of the digamma function.

Category: Mathematical

Syntax

DIGAMMA(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Restriction: Nonpositive integers are invalid.

Details

The DIGAMMA function returns the ratio that is given by

$$\Psi(x) = \Gamma'(x) / \Gamma(x)$$

where $\Gamma(.)$ and $\Gamma'(.)$ denote the Gamma function and its derivative, respectively. For *argument*>0, the DIGAMMA function is the derivative of the LGAMMA function.

Examples

SAS Statements	Results
`x=digamma(1.0);`	`-0.577215665`

DIM Function

Returns the number of elements in an array.

Category: Array

Syntax

DIM<*n*>(*array-name*)

DIM(*array-name*,*bound-n*)

Arguments

n
specifies the dimension, in a multidimensional array, for which you want to know the number of elements. If no *n* value is specified, the DIM function returns the number of elements in the first dimension of the array.

array-name
specifies the name of an array that was previously defined in the same DATA step. This argument cannot be a constant, variable, or expression.

bound-n
is a numeric constant, variable, or expression that specifies the dimension, in a multidimensional array, for which you want to know the number of elements. Use *bound-n* only when *n* is not specified.

Details

The DIM function returns the number of elements in a one-dimensional array or the number of elements in a specified dimension of a multidimensional array when the lower bound of the dimension is 1. Use DIM in array processing to avoid changing the upper bound of an iterative DO group each time you change the number of array elements.

Comparisons

- DIM always returns a total count of the number of elements in an array dimension.
- HBOUND returns the literal value of the upper bound of an array dimension.

Note: This distinction is important when the lower bound of an array dimension has a value other than 1 and the upper bound has a value other than the total number of elements in the array dimension. △

Examples

Example 1: One-dimensional Array In this example, DIM returns a value of 5. Therefore, SAS repeats the statements in the DO loop five times.

```
array big{5} weight sex height state city;
do i=1 to dim(big);
  more SAS statements;
end;
```

Example 2: Multidimensional Array This example shows two ways of specifying the DIM function for multidimensional arrays. Both methods return the same value for DIM, as shown in the table that follows the SAS code example.

```
array mult{5,10,2} mult1-mult100;
```

Syntax	Alternative Syntax	Value
DIM(MULT)	DIM(MULT,1)	5
DIM2(MULT)	DIM(MULT,2)	10
DIM3(MULT)	DIM(MULT,3)	2

See Also

Functions:
- "HBOUND Function" on page 802
- "LBOUND Function" on page 878

Statements:
- "ARRAY Statement" on page 1440
- "Array Reference Statement" on page 1445

"Array Processing" in *SAS Language Reference: Concepts*

DINFO Function

Returns information about a directory.

Category: External Files
See: DINFO Function in the documentation for your operating environment.

Syntax

DINFO(*directory-id*,*info-item*)

Arguments

directory-id
is a numeric variable that specifies the identifier that was assigned when the directory was opened by the DOPEN function.

info-item
is a character constant, variable, or expression that specifies the information item to be retrieved. DINFO returns a blank if the value of the *info-item* argument is invalid. The information available varies according to the operating environment.

Details

Use the DOPTNAME function to determine the names of the available system-dependent directory information items. Use the DOPTNUM function to determine the number of directory information items that are available.

Operating Environment Information: DINFO returns the value of a system-dependent directory parameter. See the SAS documentation for your operating environment for information about system-dependent directory parameters. △

Examples

Example 1: Using DINFO to Return Information about a Directory This example opens the directory MYDIR, determines the number of directory information items available, and retrieves the value of the last one:

```
%let filrf=MYDIR;
%let rc=%sysfunc(filename(filrf,physical-name));
%let did=%sysfunc(dopen(&filrf));
%let numopts=%sysfunc(doptnum(&did));
%let foption=%sysfunc(doptname(&did,&numopts));
%let charval=%sysfunc(dinfo(&did,&foption));
%let rc=%sysfunc(dclose(&did));
```

Example 2: Using DINFO within a DATA Step This example creates a data set that contains the name and value of each directory information item:

```
data diropts;
   length foption $ 12 charval $ 40;
   keep foption charval;
   rc=filename("mydir","physical-name");
   did=dopen("mydir");
   numopts=doptnum(did);
   do i=1 to numopts;
      foption=doptname(did,i);
      charval=dinfo(did,foption);
      output;
   end;
run;
```

See Also

Functions:

"DOPEN Function" on page 661
"DOPTNAME Function" on page 662
"DOPTNUM Function" on page 664
"FINFO Function" on page 749
"FOPTNAME Function" on page 764
"FOPTNUM Function" on page 766

DIVIDE Function

Returns the result of a division that handles special missing values for ODS output.

Category: Arithmetic

Syntax

DIVIDE(*x*, *y*)

Arguments

x
is a numeric constant, variable, or expression.

y
is a numeric constant, variable, or expression.

Details

The DIVIDE function divides two numbers and returns a result that is compatible with ODS conventions. The function handles special missing values for ODS output. The following list shows how certain special missing values are interpreted in ODS:

- .I as infinity
- .M as minus infinity
- ._ as a blank

The following table shows the values that are returned by the DIVIDE function, based on the values of *x* and *y*.

Display 4.7 Values That Are Returned by the DIVIDE Function

		x						
		positive	zero	negative	.I	.M	.__	other
	positive	x/y or .I	0	x/y or .M	.I	.M	.__	x
	zero	.I	.	.M	.I	.M	.__	x
	negative	x/y or .M	0	x/y or .I	.M	.I	.__	x
y	.I	0	0	0	.	.	.__	x
	.M	0	0	0	.	.	.__	x
	.__	.__	.__	.__	.__	.__	.__	.__
	other	y	y	y	y	y	.__	x

Note: The DIVIDE function never writes a note to the SAS log regarding missing values, division by zero, or overflow. △

Examples

The following example shows the results of using the DIVIDE function.

```
data _null_;
   a = divide(1, 0);
   put +3 a= '(infinity)';
   b = divide(2, .I);
   put +3 b=;
   c = divide(.I, -1);
   put +3 c= '(minus infinity)';
   d = divide(constant('big'), constant('small'));
   put +3 d= '(infinity because of overflow)';
run;
```

SAS writes the following output to the log:

```
a=I (infinity)
b=0
c=M (minus infinity)
d=I (infinity because of overflow)
```

DNUM Function

Returns the number of members in a directory.

Category: External Files

Syntax

DNUM(*directory-id*)

Argument

directory-id
is a numeric variable that specifies the identifier that was assigned when the directory was opened by the DOPEN function.

Details

You can use DNUM to determine the highest possible member number that can be passed to DREAD.

Examples

Example 1: Using DNUM to Return the Number of Members This example opens the directory MYDIR, determines the number of members, and closes the directory:

```
%let filrf=MYDIR;
%let rc=%sysfunc(filename(filrf,physical-name));
%let did=%sysfunc(dopen(&filrf));
%let memcount=%sysfunc(dnum(&did));
%let rc=%sysfunc(dclose(&did));
```

Example 2: Using DNUM within a DATA Step This example creates a DATA step that returns the number of members in a directory called MYDIR:

```
data _null_;
   rc=filename("mydir","physical-name");
   did=dopen("mydir");
   memcount=dnum(did);
   rc=dclose(did);
run;
```

See Also

Functions:

"DOPEN Function" on page 661
"DREAD Function" on page 665

DOPEN Function

Opens a directory, and returns a directory identifier value.

Category: External Files

See: DOPEN Function in the documentation for your operating environment.

Syntax

DOPEN(*fileref*)

Argument

fileref
is a character constant, variable, or expression that specifies the fileref assigned to the directory.

Restriction: You must associate a fileref with the directory before calling DOPEN.

Details

DOPEN opens a directory and returns a directory identifier value (a number greater than 0) that is used to identify the open directory in other SAS external file access functions. If the directory could not be opened, DOPEN returns 0, and you can obtain the error message by calling the SYSMSG function. The directory to be opened must be identified by a fileref. You can assign filerefs using the FILENAME statement or the FILENAME external file access function. Under some operating environments, you can also assign filerefs using system commands.

If you call the DOPEN function from a macro, then the result of the call is valid only when the result is passed to functions in a macro. If you call the DOPEN function from the DATA step, then the result is valid only when the result is passed to functions in the same DATA step.

Operating Environment Information: The term *directory* that is used in the description of this function and related SAS external file access functions refers to an aggregate grouping of files managed by the operating environment. Different operating environments identify such groupings with different names, such as directory, subdirectory, folder, MACLIB, or partitioned data set. For details, see the SAS documentation for your operating environment. △

Examples

Example 1: Using DOPEN to Open a Directory This example assigns the fileref MYDIR to a directory. It uses DOPEN to open the directory. DOPTNUM determines the number of system-dependent directory information items available, and DCLOSE closes the directory:

```
%let filrf=MYDIR;
%let rc=%sysfunc(filename(filrf,physical-name));
%let did=%sysfunc(dopen(&filrf));
%let infocnt=%sysfunc(doptnum(&did));
%let rc=%sysfunc(dclose(&did));
```

Example 2: Using DOPEN within a DATA Step This example opens a directory for processing within a DATA step.

```
data _null_;
   drop rc did;
   rc=filename("mydir","physical-name");
   did=dopen("mydir");
   if did > 0 then do;
      ...more statements...
   end;
   else do;
      msg=sysmsg();
      put msg;
   end;
run;
```

See Also

Functions:

"DCLOSE Function" on page 638
"DOPTNUM Function" on page 664
"FOPEN Function" on page 762
"MOPEN Function" on page 936
"SYSMSG Function" on page 1154

DOPTNAME Function

Returns directory attribute information.

Category: External Files

See: DOPTNAME Function in the documentation for your operating environment.

Syntax

DOPTNAME(*directory-id*,*nval*)

Arguments

directory-id
is a numeric variable that specifies the identifier that was assigned when the directory was opened by the DOPEN function.

nval
is a numeric constant, variable, or expression that specifies the sequence number of the option.

Details

Operating Environment Information: The number, names, and nature of the directory information varies between operating environments. The number of options that are available for a directory varies depending on the operating environment. For details, see the SAS documentation for your operating environment. △

Examples

Example 1: Using DOPTNAME to Retrieve Directory Attribute Information This example opens the directory with the fileref MYDIR, retrieves all system-dependent directory information items, writes them to the SAS log, and closes the directory:

```
%let filrf=mydir;
%let rc=%sysfunc(filename(filrf,physical-name));
%let did=%sysfunc(dopen(&filrf));
%let infocnt=%sysfunc(doptnum(&did));
%do j=1 %to &infocnt;
   %let opt=%sysfunc(doptname(&did,&j));
   %put Directory information=&opt;
%end;
%let rc=%sysfunc(dclose(&did));
```

Example 2: Using DOPTNAME within a DATA Step This example creates a data set that contains the name and value of each directory information item:

```
data diropts;
   length optname $ 12 optval $ 40;
   keep optname optval;
   rc=filename("mydir","physical-name");
   did=dopen("mydir");
   numopts=doptnum(did);
   do i=1 to numopts;
      optname=doptname(did,i);
      optval=dinfo(did,optname);
      output;
   end;
   run;
```

See Also

Functions:

"DINFO Function" on page 657
"DOPEN Function" on page 661
"DOPTNUM Function" on page 664

DOPTNUM Function

Returns the number of information items that are available for a directory.

Category: External Files

See: DOPTNUM Function in the documentation for your operating environment.

Syntax

DOPTNUM(*directory-id*)

Argument

directory-id
is a numeric variable that specifies the identifier that was assigned when the directory was opened by the DOPEN function.

Details

Operating Environment Information: The number, names, and nature of the directory information varies between operating environments. The number of options that are available for a directory varies depending on the operating environment. For details, see the SAS documentation for your operating environment. △

Examples

Example 1: Retrieving the Number of Information Items This example retrieves the number of system-dependent directory information items that are available for the directory MYDIR and closes the directory:

```
%let filrf=mydir;
%let rc=%sysfunc(filename(filrf,physical-name));
%let did=%sysfunc(dopen(&filrf));
%let infocnt=%sysfunc(doptnum(&did));
%let rc=%sysfunc(dclose(&did));
```

Example 2: Using DOPTNUM within a DATA Step This example creates a data set that retrieves the number of system-dependent information items that are available for the MYDIR directory:

```
data _null_;
   rc=filename("mydir","physical-name");
   did=dopen("mydir");
   infocnt=doptnum(did);
   rc=dclose(did);
run;
```

See Also

Functions:

"DINFO Function" on page 657

"DOPEN Function" on page 661

"DOPTNAME Function" on page 662

DREAD Function

Returns the name of a directory member.

Category: External Files

Syntax

DREAD(*directory-id*,*nval*)

Arguments

directory-id
is a numeric value that specifies the identifier that was assigned when the directory was opened by the DOPEN function.

nval
is a numeric constant, variable, or expression that specifies the sequence number of the member within the directory.

Details

DREAD returns a blank if an error occurs (such as when *nval* is out-of-range). Use DNUM to determine the highest possible member number that can be passed to DREAD.

Examples

This example opens the directory identified by the fileref MYDIR, retrieves the number of members, and places the number in the variable MEMCOUNT. It then retrieves the name of the last member, places the name in the variable LSTNAME , and closes the directory:

```
%let filrf=mydir;
%let rc=%sysfunc(filename(filrf,physical-name));
%let did=%sysfunc(dopen(&filrf));
%let lstname=;
%let memcount=%sysfunc(dnum(&did));
%if &memcount > 0 %then
   %let lstname=%sysfunc(dread(&did,&memcount));
%let rc=%sysfunc(dclose(&did));
```

See Also

Functions:

"DNUM Function" on page 660
"DOPEN Function" on page 661

DROPNOTE Function

Deletes a note marker from a SAS data set or an external file.

Category: SAS File I/O
Category: External Files

Syntax

DROPNOTE(*data-set-id* | *file-id*,*note-id*)

Arguments

data-set-id* | *file-id
is a numeric variable that specifies the identifier that was assigned when the data set or external file was opened, generally by the OPEN function or the FOPEN function.

note-id
is a numeric value that specifies the identifier that was assigned by the NOTE or FNOTE function.

Details

DROPNOTE deletes a marker set by NOTE or FNOTE. It returns a 0 if successful and ≠0 if not successful.

Examples

This example opens the SAS data set MYDATA, fetches the first observation, and sets a note ID at the beginning of the data set. It uses POINT to return to the first observation, and then uses DROPNOTE to delete the note ID:

```
%let dsid=%sysfunc(open(mydata,i));
%let rc=%sysfunc(fetch(&dsid));
%let noteid=%sysfunc(note(&dsid));
   more macro statements
%let rc=%sysfunc(point(&dsid,&noteid));
%let rc=%sysfunc(fetch(&dsid));
%let rc=%sysfunc(dropnote(&dsid,&noteid));
```

See Also

Functions:

"FETCH Function" on page 684
"FNOTE Function" on page 760
"FOPEN Function" on page 762
"FPOINT Function" on page 767
"NOTE Function" on page 956
"OPEN Function" on page 980
"POINT Function" on page 1010

DSNAME Function

Returns the SAS data set name that is associated with a data set identifier.

Category: SAS File I/O

Syntax

DSNAME(*data-set-id*)

Arguments

data-set-id
is a numeric variable that specifies the data set identifier that is returned by the OPEN function.

Details

DSNAME returns the data set name that is associated with a data set identifier, or a blank if the data set identifier is not valid.

Examples

This example determines the name of the SAS data set that is associated with the variable DSID and displays the name in the SAS log.

```
%let dsid=%sysfunc(open(sasuser.houses,i));
%put The current open data set
is %sysfunc(dsname(&dsid)).;
```

See Also

Function:

"OPEN Function" on page 980

DUR Function

Returns the modified duration for an enumerated cash flow.

Category: Financial

Syntax

DUR(*y,f,c(1), ... ,c(k)*)

Arguments

y
specifies the effective per-period yield-to-maturity, expressed as a fraction.
Range: $y > 0$

f
specifies the frequency of cash flows per period.
Range: $f > 0$

c(1), ... ,c(k)
specifies a list of cash flows.

Details

The DUR function returns the value

$$C = \sum_{k=1}^{K} \frac{k\left(\frac{c(k)}{(1+y)^{\frac{k}{f}}}\right)}{\left(P\left(1+y\right)f\right)}$$

where

$$P = \sum_{k=1}^{K} \frac{c\left(k\right)}{\left(1+y\right)^{\frac{k}{f}}}$$

Examples

```
data _null_;
   d=dur(1/20,1,.33,.44,.55,.49,.50,.22,.4,.8,.01,.36,.2,.4);
   put d;
run;
```

The value that is returned is 5.28402.

DURP Function

Returns the modified duration for a periodic cash flow stream, such as a bond.

Category: Financial

Syntax

DURP(*A,c,n,K,k*$_0$*,y*)

Arguments

A
specifies the par value.
Range: $A > 0$

c
specifies the nominal per-period coupon rate, expressed as a fraction.
Range: $0 \leq c < 1$

n
specifies the number of coupons per period.
Range: $n > 0$ and is an integer

K
specifies the number of remaining coupons.
Range: $K > 0$ and is an integer

***k*$_0$**
specifies the time from the present date to the first coupon date, expressed in terms of the number of periods.
Range: $0 < k_0 \leq \frac{1}{n}$

y
specifies the nominal per-period yield-to-maturity, expressed as a fraction.
Range: $y > 0$

Details

The DURP function returns the value

$$D = \frac{1}{n} \frac{\sum_{k=1}^{K} t_k \frac{c(k)}{\left(1+\frac{y}{n}\right)^{t_k}}}{P\left(1+\frac{y}{n}\right)}$$

where

$t_k = nk_0 + k - 1$
$c(k) = \frac{c}{n}A \quad for \;\; k = 1, \ldots, K-1$
$c(K) = \left(1+\frac{c}{n}\right)A$

and where

$$P = \sum_{k=1}^{K} \frac{c(k)}{\left(1 + \frac{y}{n}\right)^{t_k}}$$

Examples

```
data _null_;
d=durp(1000,1/100,4,14,.33/2,.10);
put d;
run;
```

The value returned is 3.26496.

ENVLEN Function

Returns the length of an environment variable.

Category: SAS File I/O

Syntax

ENVLEN(*argument*)

Arguments

argument
specifies a character variable that is the name of an operating system environment variable. Enclose *argument* in quotation marks.

Details

The ENVLEN function returns the length of the value of an operating system environment variable. If the environment variable does not exist, SAS returns –1.

Operating Environment Information: The value of *argument* is specific to your operating environment. △

Examples

The following examples are for illustration purposes only. The actual value that is returned depends on where SAS is installed on your computer.

SAS Statements	Results

`/* Windows operating environment */`
`x=envlen("PATH");`
`put x;`
	`309`

`/* UNIX operating environment */`
`y=envlen("PATH");`
`put y;`
	`365`

`z=envlen("THIS IS NOT DEFINED");` `put z;`	`-1`

ERF Function

Returns the value of the (normal) error function.

Category: Mathematical

Syntax

ERF(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The ERF function returns the integral, given by

$$\mathrm{ERF}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz$$

Examples

You can use the ERF function to find the probability (p) that a normally distributed random variable with mean 0 and standard deviation will take on a value less than X. For example, the quantity that is given by the following statement is equivalent to PROBNORM(X):

```
p=.5+.5*erf(x/sqrt(2));
```

SAS Statements	Results
`y=erf(1.0);`	`0.8427007929`
`y=erf(-1.0);`	`-0.842700793`

ERFC Function

Returns the value of the complementary (normal) error function.

Category: Mathematical

Syntax

ERFC(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The ERFC function returns the complement to the ERF function (that is, 1 – ERF(*argument*)).

Examples

SAS Statements	Results
`x=erfc(1.0);`	`0.1572992071`
`x=erfc(-1.0);`	`.8427007929`

EUCLID Function

Returns the Euclidean norm of the non-missing arguments.

Category: Descriptive Statistics

Syntax

EUCLID(*value-1 <,value-2 ...>*)

Arguments

value
specifies a numeric constant, variable, or expression.

Details

If all arguments have missing values, then the result is a missing value. Otherwise, the result is the Euclidean norm of the non-missing values.

In the following example, $x1, x2, ..., xn$ are the values of the non-missing arguments.

$$EUCLID\,(x1, x2, ..., xn) = \sqrt{x1^2 + x2^2 + ... + xn^2}$$

Examples

Example 1: Calculating the Euclidean Norm of Non-missing Arguments The following example returns the Euclidean norm of the non-missing arguments.

```
data _null_;
   x=euclid(.,3,0,.q,-4);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=5
```

Example 2: Calculating the Euclidean Norm When You Use a Variable List The following example uses a variable list to calculate the Euclidean norm.

```
data _null_;
   x1 = 1;
   x2 = 3;
   x3 = 4;
   x4 = 3;
   x5 = 1;
   x = euclid(of x1-x5);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=6
```

See Also

Functions:

"RMS Function" on page 1098

"LPNORM Function" on page 914

EXIST Function

Verifies the existence of a SAS library member.

Category: SAS File I/O

Syntax

EXIST(*member-name*<,*member-type*<, *generation*>>)

Arguments

member-name
is a character constant, variable, or expression that specifies the SAS library member. If *member-name* is blank or a null string, then EXIST uses the value of the _LAST_ system variable as the member name.

member-type
is a character constant, variable, or expression that specifies the type of SAS library member. A few common member types include ACCESS, CATALOG, DATA, and VIEW. If you do not specify a *member-type*, then the member type DATA is assumed.

generation
is a numeric constant, variable, or expression that specifies the generation number of the SAS data set whose existence you are checking. If *member-type* is not DATA, *generation* is ignored.

Positive numbers are absolute references to a historical version by its generation number. Negative numbers are relative references to a historical version in relation to the base version, from the youngest predecessor to the oldest. For example, –1 refers to the youngest version or, one version back from the base version. Zero is treated as a relative generation number.

Details

If you use a sequential library, then the results of the EXIST function are undefined. If you do *not* use a sequential library, then EXIST returns 1 if the library member exists, or 0 if *member-name* does not exist or *member-type* is invalid.

Use the CEXIST function to verify the existence of an entry in a catalog.

Examples

Example 1: Verifying the Existence of a Data Set This example verifies the existence of a data set. If the data set does not exist, then the example displays a message in the log:

```
%let dsname=sasuser.houses;
%macro opends(name);
%if %sysfunc(exist(&name)) %then
   %let dsid=%sysfunc(open(&name,i));
%else %put Data set &name does not exist.;
%mend opends;
```

```
%opends(&dsname);
```

Example 2: Verifying the Existence of a Data View This example verifies the existence of the SAS view TEST.MYVIEW. If the view does not exist, then the example displays a message in the log:

```
data _null_;
dsname="test.myview";
   if (exist(dsname,"VIEW")) then
      dsid=open(dsname,"i");
   else put dsname 'does not exist.';
run;
```

Example 3: Determining If a Generation Data Set Exists This example verifies the existence of a generation data set by using positive generation numbers (absolute reference):

```
data new(genmax=3);
   x=1;
run;
data new;
   x=99;
run;
data new;
   x=100;
run;
data new;
   x=101;
run;
data _null_;
   test=exist('new', 'DATA', 4);
   put test=;
   test=exist('new', 'DATA', 3);
   put test=;
   test=exist('new', 'DATA', 2);
   put test=;
   test=exist('new', 'DATA', 1);
   put test=;
run;
```

These lines are written to the SAS log:

```
test=1
test=1
test=1
test=0
```

You can change this example to verify the existence of the generation data set by using negative numbers (relative reference):

```
data new2(genmax=3);
   x=1;
run;
data new2;
   x=99;
run;
data new2;
   x=100;
```

```
run;
data new2;
   x=101;
run;
data _null_;
   test=exist('new2', 'DATA', 0);
   put test=;
   test=exist('new2', 'DATA', -1);
   put test=;
   test=exist('new2', 'DATA', -2);
   put test=;
   test=exist('new2', 'DATA', -3);
   put test=;
   test=exist('new2', 'DATA', -4);
   put test=;
run;
```

These lines are written to the SAS log:

```
test=1
test=1
test=1
test=0
test=0
```

See Also

Functions:

"CEXIST Function" on page 577

"FEXIST Function" on page 686

"FILEEXIST Function" on page 689

EXP Function

Returns the value of the exponential function.

Category: Mathematical

Syntax

EXP(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The EXP function raises the constant *e*, which is approximately 2.71828, to the power that is supplied by the argument. The result is limited by the maximum value of a floating-point value on the computer.

Examples

SAS Statements	Results
`x=exp(1.0);`	`2.7182818285`
`x=exp(0);`	`1`

FACT Function

Computes a factorial.

Category: Mathematical

Syntax

FACT(*n*)

Arguments

n
is a numeric constant, variable, or expression.

Details

The mathematical representation of the FACT function is given by the following equation:

$$FACT(n) = n!$$

with $n \geq 0$.

If the expression cannot be computed, a missing value is returned. For moderately large values, it is sometimes not possible to compute the FACT function.

Examples

SAS Statements	Results
`x=fact(5);`	`120`

See Also

Functions:

"COMB Function" on page 590
"PERM Function" on page 1008
"LFACT Function" on page 899

FAPPEND Function

Appends the current record to the end of an external file.

Category: External Files

Syntax

FAPPEND(*file-id*<,*cc*>)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

cc
is a character constant, variable, or expression that specifies a carriage-control character:

blank	indicates that the record starts a new line.
0	skips one blank line before this new line.
-	skips two blank lines before this new line.
1	specifies that the line starts a new page.
+	specifies that the line overstrikes a previous line.
P	specifies that the line is a computer prompt.
=	specifies that the line contains carriage control information.
all else	specifies that the line record starts a new line.

Details

FAPPEND adds the record that is currently contained in the File Data Buffer (FDB) to the end of an external file. FAPPEND returns a 0 if the operation was successful and ≠0 if it was not successful.

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file. If the file is opened successfully, it moves data into the File Data Buffer, appends a

record, and then closes the file. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf,a));
%if &fid > 0 %then
   %do;
      %let rc=%sysfunc(fput(&fid,
                       Data for the new record));
      %let rc=%sysfunc(fappend(&fid));
      %let rc=%sysfunc(fclose(&fid));
   %end;
%else
   %do;
      /* unsuccessful open processing */
   %end;
```

See Also

Functions:

"DOPEN Function" on page 661
"FCLOSE Function" on page 680
"FGET Function" on page 687
"FOPEN Function" on page 762
"FPUT Function" on page 771
"FWRITE Function" on page 778
"MOPEN Function" on page 936

FCLOSE Function

Closes an external file, directory, or directory member.

Category: External Files

See: FCLOSE Function in the documentation for your operating environment.

Syntax

FCLOSE(*file-id*)

Argument

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

Details

FCLOSE returns a 0 if the operation was successful and ≠0 if it was not successful. If you open a file within a DATA step, it is closed automatically when the DATA step ends.

Operating Environment Information: In some operating environments you must close the file with the FCLOSE function at the end of the DATA step. For more information, see the SAS documentation for your operating environment. △

Examples

This example assigns the fileref MYFILE to an external file, and attempts to open the file. If the file is opened successfully, indicated by a positive value in the variable FID, the program reads the first record, closes the file, and deassigns the fileref:

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf));
%if &fid > 0 %then
   %do;
      %let rc=%sysfunc(fread(&fid));
      %let rc=%sysfunc(fclose(&fid));
   %end;
%else
   %do;
      %put %sysfunc(sysmsg());
%end;
%let rc=%sysfunc(filename(filrf));
```

See Also

Functions:

"DCLOSE Function" on page 638
"DOPEN Function" on page 661
"FOPEN Function" on page 762
"FREAD Function" on page 772
"MOPEN Function" on page 936

FCOL Function

Returns the current column position in the File Data Buffer (FDB).

Category: External Files

Syntax

FCOL(*file-id*)

Argument

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

Details

Use FCOL combined with FPOS to manipulate data in the File Data Buffer (FDB).

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file. If the file is successfully opened, indicated by a positive value in the variable FID, it puts more data into the FDB relative to position POS, writes the record, and closes the file:

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf,o));
%if (&fid > 0) %then
   %do;
      %let record=This is data for the record.;
      %let rc=%sysfunc(fput(&fid,&record));
      %let pos=%sysfunc(fcol(&fid));
      %let rc=%sysfunc(fpos(&fid,%eval(&pos+1)));
      %let rc=%sysfunc(fput(&fid,more data));
      %let rc=%sysfunc(fwrite(&fid));
      %let rc=%sysfunc(fclose(&fid));
   %end;
%let rc=%sysfunc(filename(filrf));
```

The new record written to the external file is

```
This is data for the record. more data
```

See Also

Functions:

“FCLOSE Function” on page 680
“FOPEN Function” on page 762
“FPOS Function” on page 769
“FPUT Function” on page 771
“FWRITE Function” on page 778
“MOPEN Function” on page 936

FDELETE Function

Deletes an external file or an empty directory.

Category: External Files

See: FDELETE Function in the documentation for your operating environment.

Syntax

FDELETE(*fileref* | *directory*)

Argument

fileref
is a character constant, variable, or expression that specifies the fileref that you assigned to the external file. You can assign filerefs by using the FILENAME statement or the FILENAME external file access function.

Restriction: The fileref that you use with FDELETE cannot be a concatenation.

Operating Environment Information: In some operating environments, you can specify a fileref that was assigned with an environment variable. You can also assign filerefs using system commands. For details, see the SAS documentation for your operating environment. △

directory
is a character constant, variable, or expression that specifies an empty directory that you want to delete.

Restriction: You must have authorization to delete the directory.

Details

FDELETE returns 0 if the operation was successful or ≠0 if it was not successful.

Examples

Example 1: Deleting an External File This example generates a fileref for an external file in the variable FNAME. Then it calls FDELETE to delete the file and calls the FILENAME function again to deassign the fileref.

```
data _null_;
    fname="tempfile";
    rc=filename(fname,"physical-filename");
    if rc = 0 and fexist(fname) then
       rc=fdelete(fname);
    rc=filename(fname);
run;
```

Example 2: Deleting a Directory This example uses FDELETE to delete an empty directory to which you have write access. If the directory is not empty, the optional SYSMSG function returns an error message stating that SAS is unable to delete the file.

```
filename testdir 'physical-filename';
data _null_;
   rc=fdelete('testdir');
   put rc=;
```

```
   msg=sysmsg();
   put msg=;
run;
```

See Also

Functions:

"FEXIST Function" on page 686

"FILENAME Function" on page 690

Statement:

"FILENAME Statement" on page 1520

FETCH Function

Reads the next non-deleted observation from a SAS data set into the Data Set Data Vector (DDV).

Category: SAS File I/O

Syntax

FETCH(*data-set-id* <,'NOSET'>)

Arguments

data-set-id
is a numeric variable that specifies the data set identifier that is returned by the OPEN function.

'NOSET'
prevents the automatic passing of SAS data set variable values to macro or DATA step variables even if the SET routine has been called.

Details

FETCH returns a 0 if the operation is successful, ≠0 if it is not successful, and – 1 if the end of the data set is reached. FETCH skips observations marked for deletion.

If the SET routine has been called previously, the values for any data set variables are automatically passed from the DDV to the corresponding DATA step or macro variables. To override this behavior temporarily so that fetched values are not automatically copied to the DATA step or macro variables, use the NOSET option.

Examples

This example fetches the next observation from the SAS data set MYDATA. If the end of the data set is reached or if an error occurs, SYSMSG retrieves the appropriate message and writes it to the SAS log. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let dsid=%sysfunc(open(mydata,i));
%let rc=%sysfunc(fetch(&dsid));
%if &rc ne 0 %then
  %put %sysfunc(sysmsg());
%else
   %do;
      ...more macro statements...
   %end;
%let rc=%sysfunc(close(&dsid));
```

See Also

CALL Routine:

"CALL SET Routine" on page 525

Functions:

"FETCHOBS Function" on page 685
"GETVARC Function" on page 794
"GETVARN Function" on page 795

FETCHOBS Function

Reads a specified observation from a SAS data set into the Data Set Data Vector (DDV).

Category: SAS File I/O

Syntax

FETCHOBS(*data-set-id*,*obs-number*<,*options*>)

Arguments

data-set-id
is a numeric variable that specifies the data set identifier that is returned by the OPEN function.

obs-number
is a numeric constant, variable, or expression that specifies the number of the observation to read. FETCHOBS treats the observation value as a relative observation number unless you specify the ABS option. The relative observation number might not coincide with the physical observation number on disk, because the function skips observations marked for deletion. When a WHERE clause is active, the function counts only observations that meet the WHERE condition.

Default: FETCHOBS skips deleted observations.

options
is a character constant, variable, or expression that names one or more options, separated by blanks:

ABS	specifies that the value of *obs-number* is absolute. That is, deleted observations are counted.
NOSET	prevents the automatic passing of SAS data set variable values to DATA step or macro variables even if the SET routine has been called.

Details

FETCHOBS returns 0 if the operation was successful, ≠0 if it was not successful, and –1 if the end of the data set is reached. To retrieve the error message that is associated with a non-zero return code, use the SYSMSG function. If the SET routine has been called previously, the values for any data set variables are automatically passed from the DDV to the corresponding DATA step or macro variables. To override this behavior temporarily, use the NOSET option.

If *obs-number* is less than 1, the function returns an error condition. If *obs-number* is greater than the number of observations in the SAS data set, the function returns an end-of-file condition.

Examples

This example fetches the tenth observation from the SAS data set MYDATA. If an error occurs, the SYSMSG function retrieves the error message and writes it to the SAS log. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let rc = %sysfunc(fetchobs(&mydataid,10));
%if &rc = –1 %then
   %put End of data set has been reached.;
%if &rc > 0 %then %put %sysfunc(sysmsg());
```

See Also

CALL Routine:

"CALL SET Routine" on page 525

Functions:

"FETCH Function" on page 684
"GETVARC Function" on page 794
"GETVARN Function" on page 795

FEXIST Function

Verifies the existence of an external file that is associated with a fileref.

Category: External Files

See: FEXIST Function in the documentation for your operating environment.

Syntax

FEXIST(*fileref*)

Argument

fileref
is a character constant, variable, or expression that specifies the fileref that is assigned to an external file.

Restriction: The *fileref* must have been previously assigned.

Operating Environment Information: In some operating environments, you can specify a fileref that was assigned with an environment variable. For details, see the SAS documentation for your operating environment. △

Details

FEXIST returns 1 if the external file that is associated with *fileref* exists, and 0 if the file does not exist. You can assign filerefs by using the FILENAME statement or the FILENAME external file access function. In some operating environments, you can also assign filerefs by using system commands.

Comparison

FILEEXIST verifies the existence of a file based on its physical name.

Examples

This example verifies the existence of an external file and writes the result to the SAS log:

```
%if %sysfunc(fexist(&fref)) %then
   %put The file identified by the fileref
      &fref exists.;
%else
   %put %sysfunc(sysmsg());
```

See Also

Functions:
“EXIST Function” on page 675
“FILEEXIST Function” on page 689
“FILENAME Function” on page 690
“FILEREF Function” on page 692
Statement:
“FILENAME Statement” on page 1520

FGET Function

Copies data from the File Data Buffer (FDB) into a variable.

Category: External Files

Syntax

FGET(*file-id*,*variable*<,*length*>)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

variable
in a DATA step, specifies a character variable to hold the data. In a macro, specifies a macro variable to hold the data. If *variable* is a macro variable and it does not exist, it is created.

length
specifies the number of characters to retrieve from the FDB. If *length* is specified, only the specified number of characters is retrieved (or the number of characters remaining in the buffer if that number is less than length). If *length* is omitted, all characters in the FDB from the current column position to the next delimiter are returned. The default delimiter is a blank. The delimiter is not retrieved.

See: The "FSEP Function" on page 775 for more information about delimiters.

Details

FGET returns 0 if the operation was successful, or –1 if the end of the FDB was reached or no more tokens were available.

After FGET is executed, the column pointer moves to the next read position in the FDB.

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file. If the file is opened successfully, it reads the first record into the File Data Buffer, retrieves the first token of the record and stores i t in the variable MYSTRING, and then closes the file. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf));
%if &fid > 0 %then
   %do;
      %let rc=%sysfunc(fread(&fid));
      %let rc=%sysfunc(fget(&fid,mystring));
      %put &mystring;
      %let rc=%sysfunc(fclose(&fid));
   %end;
%let rc=%sysfunc(filename(filrf));
```

See Also

Functions:

"FCLOSE Function" on page 680
"FILENAME Function" on page 690
"FOPEN Function" on page 762
"FPOS Function" on page 769
"FREAD Function" on page 772
"FSEP Function" on page 775
"MOPEN Function" on page 936

FILEEXIST Function

Verifies the existence of an external file by its physical name.

Category: External Files

See: FILEEXIST Function in the documentation for your operating environment.

Syntax

FILEEXIST(*file-name*)

Argument

file-name
is a character constant, variable, or expression that specifies a fully qualified physical filename of the external file in the operating environment.

Details

FILEEXIST returns 1 if the external file exists and 0 if the external file does not exist. The specification of the physical name for *file-name* varies according to the operating environment.

Although your operating environment utilities might recognize partial physical filenames, you must always use fully qualified physical filenames with FILEEXIST.

Examples

This example verifies the existence of an external file. If the file exists, FILEEXIST opens the file. If the file does not exist, FILEEXIST displays a message in the SAS log. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%if %sysfunc(fileexist(&myfilerf)) %then
    %let fid=%sysfunc(fopen(&myfilerf));
%else
   %put The external file &myfilerf does not exist.;
```

See Also

Functions:

"EXIST Function" on page 675
"FEXIST Function" on page 686
"FILENAME Function" on page 690
"FILEREF Function" on page 692
"FOPEN Function" on page 762

FILENAME Function

Assigns or deassigns a fileref to an external file, directory, or output device.

Category: External Files

See: FILENAME Function in the documentation for your operating environment.

Syntax

FILENAME(*fileref* <,*file-name*> <,*device-type*> <,'*host-options*'> <,*dir-ref*>)

Arguments

fileref
specifies the fileref to assign to the external file. In a DATA step, *fileref* can be a character expression, a string enclosed in quotation marks that specifies the fileref, or a DATA step variable whose value contains the fileref. In a macro (for example, in the %SYSFUNC function), *fileref* is the name of a macro variable (without an ampersand) whose value contains the fileref to assign to the external file.

Requirement: If *fileref* is a DATA step variable, its length must be no longer than eight characters.

Tip: If a fileref is a DATA step character variable with a blank value and a maximum length of eight characters, or if a macro variable named in *fileref* has a null value, then a fileref is generated and assigned to the character variable or macro variable, respectively.

file-name
is a character constant, variable, or expression that specifies the external file. Specifying a blank *file-name* deassigns a fileref that was assigned previously.

device-type
is a character constant, variable, or expression that specifies the type of device or the access method that is used if the fileref points to an input or output device or location that is not a physical file:

DISK specifies that the device is a disk drive.

Tip: When you assign a fileref to a file on disk, you are not required to specify DISK.

Alias: BASE

DUMMY — specifies that the output to the file is discarded.

Tip: Specifying DUMMY can be useful for testing.

GTERM — indicates that the output device type is a graphics device that will be receiving graphics data.

PIPE — specifies an unnamed pipe.

Note: Some operating environments do not support pipes. △

PLOTTER — specifies an unbuffered graphics output device.

PRINTER — specifies a printer or printer spool file.

TAPE — specifies a tape drive.

TEMP — creates a temporary file that exists only as long as the filename is assigned. The temporary file can be accessed only through the logical name and is available only while the logical name exists.

Restriction: Do not specify a physical pathname. If you do, SAS returns an error.

Tip: Files that are manipulated by the TEMP device can have the same attributes and behave identically to DISK files.

TERMINAL — specifies the user's personal computer.

UPRINTER — specifies a Universal Printing printer definition name.

Operating Environment Information: The FILENAME function also supports operating environment-specific devices. For more information, see the SAS documentation for your operating environment. △

'host-options'
specifies host-specific details such as file attributes and processing attributes. For more information, see the SAS documentation for your operating environment.

dir-ref
specifies the fileref that was assigned to the directory or partitioned data set in which the external file resides.

Details

FILENAME returns 0 if the operation was successful; ≠0 if it was not successful. The name that is associated with the file or device is called a *fileref* (file reference name). Other system functions that manipulate external files and directories require that the files be identified by fileref rather than by physical filename. The association between a fileref and a physical file lasts only for the duration of the current SAS session or until you change or discontinue the association by using FILENAME. You can deassign filerefs by specifying a null string for the *file-name* argument in FILENAME.

Operating Environment Information: The term *directory* in this description refers to an aggregate grouping of files that are managed by the operating environment. Different operating environments identify these groupings with different names, such as directory, subdirectory, folder, MACLIB, or partitioned data set. For details, see the SAS documentation for your operating environment.

Under some operating environments, you can also assign filerefs by using system commands. Depending on the operating environment, FILENAME might be unable to change or deassign filerefs that are assigned outside of SAS. △

Examples

Example 1: Assigning a Fileref to an External File This example assigns the fileref MYFILE to an external file. Next, it deassigns the fileref. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf, physical-filename));
%if &rc ne 0 %then
   %put %sysfunc(sysmsg());
%let rc=%sysfunc(filename(filrf));
```

Example 2: Assigning a System-Generated Fileref This example assigns a system-generated fileref to an external file. The fileref is stored in the variable FNAME. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let rc=%sysfunc(filename(fname, physical-filename));
%if &rc %then
   %put %sysfunc(sysmsg());
%else
   %do;
      more macro statements
   %end;
```

Example 3: Assigning a Fileref to a Pipe File This example assigns the fileref MYPIPE to a pipe file with the output from the UNIX command LS, which lists the files in the directory /u/myid. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let filrf=mypipe;
%let rc=%sysfunc(filename(filrf, %str(ls /u/myid), pipe));
```

See Also

Functions:

"FEXIST Function" on page 686
"FILEEXIST Function" on page 689
"FILEREF Function" on page 692
"SYSMSG Function" on page 1154

FILEREF Function

Verifies whether a fileref has been assigned for the current SAS session.

Category: External Files
See: FILEREF Function in the documentation for your operating environment.

Syntax

FILEREF(*fileref*)

Argument

fileref
is a character constant, variable, or expression that specifies the fileref to be validated.

Range: 1 to 8 characters

Details

A negative return code indicates that the fileref exists but the physical file associated with the fileref does not exist. A positive value indicates that the fileref is not assigned. A value of zero indicates that the fileref and external file both exist.

A fileref can be assigned to an external file by using the FILENAME statement or the FILENAME function.

Operating Environment Information: Under some operating environments, filerefs can also be assigned by using system commands. For details, see the SAS documentation for your operating environment. △

Examples

Example 1: Verifying That a Fileref Is Assigned This example tests whether the fileref MYFILE is currently assigned to an external file. A system error message is issued if the fileref is not currently assigned:

```
%if %sysfunc(fileref(myfile))>0 %then
   %put MYFILE is not assigned;
```

Example 2: Verifying That Both a Fileref and a File Exist This example tests for a zero value to determine whether both the fileref and the file exist:

```
%if %sysfunc(fileref(myfile)) ne 0 %then
   %put %sysfunc(sysmsg());
```

See Also

Functions:

"FEXIST Function" on page 686
"FILEEXIST Function" on page 689
"FILENAME Function" on page 690
"SYSMSG Function" on page 1154

Statement:

"FILENAME Statement" on page 1520

FINANCE Function

Computes financial calculations such as depreciation, maturation, accrued interest, net present value, periodic savings, and internal rates of return.

Category: Financial

Syntax

FINANCE(*string-identifier*, *parm1*, *parm2,...*)

Arguments

string-identifier
specifies a character constant, variable, or expression. Valid values for *string-identifier* are listed in the following table.

string-identifier	Description
"ACCRINT" on page 697	computes the accrued interest for a security that pays periodic interest.
"ACCRINTM" on page 697	computes the accrued interest for a security that pays interest at maturity.
"AMORDEGRC" on page 698	computes the depreciation for each accounting period by using a depreciation coefficient.
"AMORLINC" on page 698	computes the depreciation for each accounting period.
"COUPDAYBS" on page 699	computes the number of days from the beginning of the coupon period to the settlement date.
"COUPDAYS" on page 699	computes the number of days in the coupon period that contains the settlement date.
"COUPDAYSNC" on page 699	computes the number of days from the settlement date to the next coupon date.
"COUPNCD" on page 700	computes the next coupon date after the settlement date.
"COUPNUM" on page 700	computes the number of coupons that are payable between the settlement date and the maturity date.
"COUPPCD" on page 701	computes the previous coupon date before the settlement date.
"CUMIPMT" on page 701	computes the cumulative interest that is paid between two periods.
"CUMPRINC" on page 701	computes the cumulative principal that is paid on a loan between two periods.
"DB" on page 702	computes the depreciation of an asset for a specified period by using the fixed-declining balance method.
"DDB" on page 702	computes the depreciation of an asset for a specified period by using the double-declining balance method or some other method that you specify.
"DISC" on page 703	computes the discount rate for a security.

string-identifier	Description
"DOLLARDE" on page 703	converts a dollar price, expressed as a fraction, to a dollar price, expressed as a decimal number.
"DOLLARFR" on page 703	converts a dollar price, expressed as a decimal number, to a dollar price, expressed as a fraction.
"DURATION" on page 704	computes the annual duration of a security with periodic interest payments.
"EFFECT" on page 704	computes the effective annual interest rate.
"FV" on page 704	computes the future value of an investment.
"FVSCHEDULE" on page 705	computes the future value of an initial principal after applying a series of compound interest rates.
"INTRATE" on page 705	computes the interest rate for a fully invested security.
"IPMT" on page 705	computes the interest payment for an investment for a given period.
"IRR" on page 706	computes the internal rate of return for a series of cash flows.
"MDURATION" on page 706	computes the Macaulay modified duration for a security with an assumed face value of $100.
"MIRR" on page 707	computes the internal rate of return where positive and negative cash flows are financed at different rates.
"NOMINAL" on page 707	computes the annual nominal interest rate.
"NPER" on page 707	computes the number of periods for an investment.
"NPV" on page 708	computes the net present value of an investment based on a series of periodic cash flows and a discount rate.
"ODDFPRICE" on page 708	computes the price per $100 face value of a security with an odd first period.
"ODDFYIELD" on page 709	computes the yield of a security with an odd first period.
"ODDLPRICE" on page 709	computes the price per $100 face value of a security with an odd last period.
"ODDLYIELD" on page 710	computes the yield of a security with an odd last period.
"PMT" on page 711	computes the periodic payment for an annuity.
"PPMT" on page 711	computes the payment on the principal for an investment for a given period.
"PRICE" on page 712	computes the price per $100 face value of a security that pays periodic interest.
"PRICEDISC" on page 712	computes the price per $100 face value of a discounted security.
"PRICEMAT" on page 713	computes the price per $100 face value of a security that pays interest at maturity.

string-identifier	Description
"PV" on page 713	computes the present value of an investment.
"RATE" on page 713	computes the interest rate per period of an annuity.
"RECEIVED" on page 714	computes the amount received at maturity for a fully invested security.
"SLN" on page 715	computes the straight-line depreciation of an asset for one period.
"SYD" on page 715	computes the sum-of-years digits depreciation of an asset for a specified period.
"TBILLEQ" on page 716	computes the bond-equivalent yield for a treasury bill.
"TBILLPRICE" on page 716	computes the price per $100 face value for a treasury bill.
"TBILLYIELD" on page 716	computes the yield for a treasury bill.
"VDB" on page 717	computes the depreciation of an asset for a specified or partial period by using a declining balance method.
"XIRR" on page 717	computes the internal rate of return for a schedule of cash flows that is not necessarily periodic.
"XNPV" on page 717	computes the net present value for a schedule of cash flows that is not necessarily periodic.
"YIELD" on page 718	computes the yield on a security that pays periodic interest.
"YIELDDISC" on page 719	computes the annual yield for a discounted security (for example, a treasury bill).
"YIELDMAT" on page 719	computes the annual yield of a security that pays interest at maturity.

parm
specifies a parameter that is associated with each *string-identifier*. The following parameters are available:

basis
is an optional parameter that specifies a character or numeric value that indicates the type of day count basis to use.

Numeric Value	String Value	Day Count Method
0	"30/360"	US (NASD) 30/360
1	"ACTUAL"	Actual/actual
2	"ACT/360"	Actual/360
3	"ACT/365"	Actual/365
4	"EU30/360"	European 30/360

interest-rates
specifies rates that are provided as numeric values and not as percentages.

dates
specifies that all dates in the financial functions are SAS dates.

sign-of-cash-values
for all the arguments, specifies that the cash you pay out, such as deposits to savings or other withdrawals, is represented by negative numbers. It also specifies that the cash you receive, such as dividend checks and other deposits, is represented by positive numbers.

Details

ACCRINT Computes the accrued interest for a security that pays periodic interest.

FINANCE('ACCRINT', *issue*, *first-interest*, *settlement*, *rate*, *par*, *frequency*, <*basis*>);

where

issue
specifies the issue date of the security.

first-interest
specifies the first interest date of the security.

settlement
specifies the settlement date.

rate
specifies the interest rate.

par
specifies the par value of the security. If you omit *par*, SAS uses the value **$1000**.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 1 on page 720

ACCRINTM Computes the accrued interest for a security that pays interest at maturity.

FINANCE('ACCRINTM', *issue*, *settlement*, *rate*, *par*, <*basis*>);

where

issue
specifies the issue date of the security.

settlement
specifies the settlement date.

rate
specifies the interest rate.

par
specifies the par value of the security. If you omit *par*, SAS uses the value **$1000**.

basis
specifies the optional day count value.

Featured in: Example 2 on page 720

AMORDEGRC Computes the depreciation for each accounting period by using a depreciation coefficient.

FINANCE('AMORDEGRC', *cost*, *date-purchased*, *first-period*, *salvage*, *period*, *rate*, <*basis*>);

where

cost
specifies the initial cost of the asset.

date-purchased
specifies the date of the purchase of the asset.

first-period
specifies the date of the end of the first period.

salvage
specifies the value at the end of the depreciation (also called the salvage value of the asset).

period
specifies the depreciation period.

rate
specifies the rate of depreciation.

basis
specifies the optional day count value.

Tip: When the first argument of the FINANCE function is AMORDEGRC and the value of *basis* is 2, the function returns a missing value.

Featured in: Example 3 on page 721

AMORLINC Computes the depreciation for each accounting period.

FINANCE('AMORLINC', *cost*, *date-purchased*, *first-period*, *salvage*, *period*, *rate*, <*basis*>);

where

cost
specifies the initial cost of the asset.

date-purchased
specifies the date of the purchase of the asset.

first-period
specifies the date of the end of the first period.

salvage
specifies the value at the end of the depreciation (also called the salvage value of the asset).

period
specifies the depreciation period.

rate
specifies the rate of depreciation.

basis
specifies the optional day count value.

Tip: When the first argument of the FINANCE function is AMORLINC and the value of *basis* is 2, the function returns a missing value.

Featured in: Example 4 on page 721

COUPDAYBS Computes the number of days from the beginning of the coupon period to the settlement date.

FINANCE('COUPDAYBS', *date-purchased*, *first-period*, *period*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 5 on page 721

COUPDAYS Computes the number of days in the coupon period that contains the settlement date.

FINANCE('COUPDAYS', *settlement*, *maturity*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 6 on page 722

COUPDAYSNC Computes the number of days from the settlement date to the next coupon date.

FINANCE('COUPDAYSNC', *settlement*, *maturity*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 7 on page 722

COUPNCD Computes the next coupon date after the settlement date.

FINANCE('COUPNCD', *settlement*, *maturity*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 8 on page 722

COUPNUM Computes the number of coupons that are payable between the settlement date and the maturity date.

FINANCE('COUPNUM', *settlement*, *maturity*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 9 on page 723

COUPPCD Computes the previous coupon date before the settlement date.

FINANCE('COUPPCD', *settlement*, *maturity*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 10 on page 723

CUMIPMT Computes the cumulative interest paid between two periods.

FINANCE('CUMIPMT', *rate*, *nper*, *pv*, *start-period*, *end-period*, <*type*>);

where

rate
specifies the interest rate.

nper
specifies the total number of payment periods.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently.

start-period
specifies the first period in the calculation. Payment periods are numbered beginning with 1.

end-period
specifies the last period in the calculation.

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 11 on page 723

CUMPRINC Computes the cumulative principal that is paid on a loan between two periods.

FINANCE('CUMPRINC', *rate*, *nper*, *pv*, *start-period*, *end-period*, <*type*>);

where

rate
specifies the interest rate.

nper
specifies the total number of payment periods.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently.

start-period
specifies the first period in the calculation. Payment periods are numbered beginning with 1.

end-period
specifies the last period in the calculation.

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 12 on page 724

DB Computes the depreciation of an asset for a specified period by using the fixed-declining balance method.

FINANCE('DB', *cost*, *salvage*, *life*, *period*, <*month*>);

where

cost
specifies the initial cost of the asset.

salvage
specifies the value at the end of the depreciation (also called the salvage value of the asset).

life
specifies the number of periods over which the asset is depreciated (also called the useful life of the asset).

period
specifies the period for which you want to calculate the depreciation. *Period* must use the same time units as *life*.

month
specifies the number of months (month is an optional numeric argument). If month is omitted, it defaults to a value of 12.

Featured in: Example 13 on page 724

DDB Computes the depreciation of an asset for a specified period by using the double-declining balance method or some other method that you specify.

FINANCE('DDB', *cost*, *salvage*, *life*, *period*, <*factor*>);

where

cost
specifies the initial cost of the asset.

salvage
specifies the value at the end of the depreciation (also called the salvage value of the asset).

life
specifies the number of periods over which the asset is depreciated (also called the useful life of the asset).

period
specifies the period for which you want to calculate the depreciation. *Period* must use the same time units as *life*.

factor
specifies the rate at which the balance declines. If *factor* is omitted, it is assumed to be 2 (the double-declining balance method).

Featured in: Example 14 on page 724

DISC Computes the discount rate for a security.

FINANCE('DISC', *settlement*, *maturity*, *pr*, *redemption*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

pr
specifies the price of security per $100 face value.

redemption
specifies the amount to be received at maturity.

basis
specifies the optional day count value.

Featured in: Example 15 on page 724

DOLLARDE Converts a dollar price, expressed as a fraction, to a dollar price, expressed as a decimal number.

FINANCE('DOLLARDE', *fractionaldollar*, *fraction*);

where

fractionaldollar
specifies the number expressed as a fraction.

fraction
specifies the integer to use in the denominator of a fraction.

Featured in: Example 16 on page 725

DOLLARFR Converts a dollar price, expressed as a decimal number, to a dollar price, expressed as a fraction.

FINANCE('DOLLARFR', *decimaldollar*, *fraction*);

where

decimaldollar
specifies a decimal number.

fraction
specifies the integer to use in the denominator of a fraction.

Featured in: Example 17 on page 725

DURATION Computes the annual duration of a security with periodic interest payments.

FINANCE('DURATION', *settlement*, *maturity*, *coupon*, *yld*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

coupon
specifies the annual coupon rate of the security.

yld
specifies the annual yield of the security.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 18 on page 725

EFFECT Computes the effective annual interest rate.

FINANCE('EFFECT', *nominalrate*, *npery*);

where

nominalrate
specifies the nominal interest rate.

npery
specifies the number of compounding periods per year.

Featured in: Example 19 on page 725

FV Computes the future value of an investment.

FINANCE('FV', *rate*, *nper*, <*pmt*>, <*pv*>, <*type*>);

where

rate
specifies the interest rate.

nper
specifies the total number of payment periods.

pmt
specifies the payment that is made each period; the payment cannot change over the life of the annuity. Typically, *pmt* contains principal and interest but no fees and taxes. If *pmt* is omitted, you must include the *pv* argument.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently. If *pv* is omitted, it is assumed to be 0 (zero), and you must include the *pmt* argument.

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 20 on page 726

FVSCHEDULE Computes the future value of the initial principal after applying a series of compound interest rates.

FINANCE('FVSCHEDULE', *principal*, *schedule1*, *schedule2*...);

where

principal
specifies the present value.

schedule
specifies the sequence of interest rates to apply.

Featured in: Example 21 on page 726

INTRATE Computes the interest rate for a fully invested security.

FINANCE('INTRATE', *settlement*, *maturity*, *investment*, *redemption*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

investment
specifies the amount that is invested in the security.

redemption
specifies the amount to be received at maturity.

basis
specifies the optional day count value.

Featured in: Example 22 on page 726

IPMT Computes the interest payment for an investment for a specified period.

FINANCE('IPMT', *rate*, *period*, *nper*, *pv*, <*fv*>, <*type*>);

where

rate
specifies the interest rate.

period
specifies the period for which you want to calculate the depreciation. *Period* must use the same units as *life*.

nper
specifies the total number of payment periods.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently. If *pv* is omitted, it is assumed to be 0 (zero), and you must include the *fv* argument.

fv
specifies the future value or a cash balance that you want to attain after the last payment is made. If *fv* is omitted, it is assumed to be 0 (for example, the future value of a loan is 0).

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 23 on page 726

IRR Computes the internal rate of return for a series of cash flows.

FINANCE('IRR', *value1*, *value2*, ..., *value_n*);

where

value
specifies a list of numeric arguments that contain numbers for which you want to calculate the internal rate of return.

Featured in: Example 24 on page 727

MDURATION Computes the Macaulay modified duration for a security with an assumed face value of $100.

FINANCE('MDURATION', *settlement*, *maturity*, *coupon*, *yld*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

coupon
specifies the annual coupon rate of the security.

yld
specifies the annual yield of the security.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 25 on page 727

MIRR Computes the internal rate of return where positive and negative cash flows are financed at different rates.

FINANCE('MIRR', *value1*, …, *value_n*, *financerate*, *reinvestrate*);

where

values
specifies a list of numeric arguments that contain numbers. These numbers represent a series of payments (negative values) and income (positive values) that occur at regular periods. *Values* must contain at least one positive value and one negative value to calculate the modified internal rate of return.

financerate
specifies the interest rate that you pay on the money that is used in the cash flows.

reinvestrate
specifies the interest rate that you receive on the cash flows as you reinvest them.

Featured in: Example 26 on page 727

NOMINAL Computes the annual nominal interest rates.

FINANCE('NOMINAL', *effectrate*, *npery*);

where

effectrate
specifies the effective interest rate.

npery
specifies the number of compounding periods per year.

Featured in: Example 27 on page 728

NPER Computes the number of periods for an investment.

FINANCE('NPER', *rate*, *pmt*, *pv*, <*fv*>, <*type*>);

where

rate
specifies the interest rate.

pmt
specifies the payment that is made each period; the payment cannot change over the life of the annuity. Typically, *pmt* contains principal and interest but no other fees or taxes. If *pmt* is omitted, you must include the *pv* argument.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently. If *pv* is omitted, it is assumed to be 0 (zero), and you must include the *pmt* argument.

fv
specifies the future value or a cash balance that you want to attain after the last payment is made. If *fv* is omitted, it is assumed to be 0 (for example, the future value of a loan is 0).

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 28 on page 728

NPV Computes the net present value of an investment based on a series of periodic cash flows and a discount rate.

FINANCE('NPV', *rate*, *value–1* <,...value-n>);

where

rate
specifies the interest rate.

value
represents the sequence of the cash flows.

Featured in: Example 29 on page 728

ODDFPRICE Computes the price of a security per $100 face value with an odd first period.

FINANCE('ODDFPRICE', *settlement*, *maturity*, *issue*, *first-coupon*, *rate*, *yld*, *redemption*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

issue
specifies the issue date of the security.

first-coupon
specifies the first coupon date of the security.

rate
specifies the interest rate.

yld
specifies the annual yield of the security.

redemption
specifies the amount to be received at maturity.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 30 on page 728

ODDFYIELD Computes the yield of a security with an odd first period.

FINANCE('ODDFYIELD', *settlement*, *maturity*, *issue*, *first-coupon*, *rate*, *pr*, *redemption*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

issue
specifies the issue date of the security.

first-coupon
specifies the first coupon date of the security.

rate
specifies the interest rate.

pr
specifies the price of the security per $100 face value.

redemption
specifies the amount to be received at maturity.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 31 on page 729

ODDLPRICE Computes the price of a security per $100 face value with an odd last period.

FINANCE('ODDLPRICE', *settlement*, *maturity*, *last_interest*, *rate*, *yld*, *redemption*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

last_interest
specifies the last coupon date of the security.

rate
specifies the interest rate.

yld
specifies the annual yield of the security.

redemption
specifies the amount to be received at maturity.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 32 on page 729

ODDLYIELD Computes the yield of a security with an odd last period.

FINANCE('ODDLYIELD', *settlement*, *maturity*, *last_interest*, *rate*, *pr*, *redemption*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

last_interest
specifies the last coupon date of the security.

rate
specifies the interest rate.

pr
specifies the price of the security per $100 face value.

redemption
specifies the amount to be received at maturity.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 33 on page 729

PMT Computes the periodic payment of an annuity.

FINANCE('PMT', *rate*, *nper*, *pv*, <*fv*>, <*type*>);

where

rate
specifies the interest rate.

nper
specifies the number of payment periods.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently. If *pv* is omitted, it is assumed to be 0 (zero), and you must include the *fv* argument.

fv
specifies the future value or a cash balance that you want to attain after the last payment is made. If *fv* is omitted, it is assumed to be 0 (for example, the future value of a loan is 0).

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 34 on page 730

PPMT Computes the payment on the principal for an investment for a specified period.

FINANCE('PPMT', *rate*, *per*, *nper*, *pv*, <*fv*>, <*type*>);

where

rate
specifies the interest rate.

per
specifies the period.

Range: 1–*nper*

nper
specifies the number of payment periods.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently. If *pv* is omitted, it is assumed to be 0 (zero), and you must include the *fv* argument.

fv
specifies the future value or a cash balance that you want to attain after the last payment is made. If *fv* is omitted, it is assumed to be 0 (for example, the future value of a loan is 0).

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 35 on page 730

PRICE Computes the price of a security per $100 face value that pays periodic interest.

FINANCE('PRICE', *settlement*, *maturity*, *rate*, *yld*, *redemption*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

rate
specifies the interest rate.

yld
specifies the annual yield of the security.

redemption
specifies the amount to be received at maturity.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 36 on page 730

PRICEDISC Computes the price of a discounted security per $100 face value.

FINANCE('PRICEDISC', *settlement*, *maturity*, *discount*, *redemption*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

discount
specifies the discount rate of the security.

redemption
specifies the amount to be received at maturity.

basis
specifies the optional day count value.

Featured in: Example 37 on page 730

PRICEMAT Computes the price of a security per $100 face value that pays interest at maturity.

FINANCE('PRICEMAT', *settlement*, *maturity*, *issue*, *rate*, *yld*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

issue
specifies the issue date of the security.

rate
specifies the interest rate.

yld
specifies the annual yield of the security.

basis
specifies the optional day count value.

Featured in: Example 38 on page 731

PV Computes the present value of an investment.

FINANCE('PV', *rate*, *nper*, *pmt*, <*fv*>, <*type*>);

where

rate
specifies the interest rate.

nper
specifies the total number of payment periods.

pmt
specifies the payment that is made each period; the payment cannot change over the life of the annuity. Typically, *pmt* contains principal and interest but no other fees or taxes.

fv
specifies the future value or a cash balance that you want to attain after the last payment is made. If *fv* is omitted, it is assumed to be 0 (for example, the future value of a loan is 0).

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 39 on page 731

RATE Computes the interest rate per period of an annuity.

FINANCE('RATE', *nper*, *pmt*, *pv*, <*fv*>, <*type*>);

where

nper
specifies the total number of payment periods.

pmt
specifies the payment that is made each period; the payment cannot change over the life of the annuity. Typically, *pmt* contains principal and interest but no other fees or taxes. If *pmt* is omitted, you must include the *pv* argument.

pv
specifies the present value or the lump-sum amount that a series of future payments is worth currently. If *pv* is omitted, it is assumed to be 0 (zero), and you must include the *fv* argument.

fv
specifies the future value or a cash balance you want to attain after the last payment is made. If *fv* is omitted, it is assumed to be 0 (for example, the future value of a loan is 0).

type
specifies the number 0 or 1 and indicates when payments are due. If *type* is omitted, it is assumed to be 0.

Featured in: Example 40 on page 731

RECEIVED Computes the amount that is received at maturity for a fully invested security.

FINANCE('RECEIVED', *settlement*, *maturity*, *investment*, *discount*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

investment
specifies the amount that is invested in the security.

discount
specifies the discount rate of the security.

basis
specifies the optional day count value.

Featured in: Example 41 on page 731

SLN Computes the straight-line depreciation of an asset for one period.

FINANCE('SLN', *cost*, *salvage*, *life*);

where

cost
specifies the initial cost of the asset.

salvage
specifies the value at the end of the depreciation (also called the salvage value of an asset).

life
specifies the number of periods over which the asset is depreciated (also called the useful life of the asset).

Featured in: Example 42 on page 732

SYD Computes the sum-of-years digits depreciation of an asset for a specified period.

FINANCE('SYD', *cost*, *salvage*, *life*, *period*);

where

cost
specifies the initial cost of the asset.

salvage
specifies the value at the end of the depreciation (also called the salvage value of the asset).

life
specifies the number of periods over which the asset is depreciated (also called the useful life of the asset).

period
specifies a period in the same time units that are used for the argument *life*.

Featured in: Example 43 on page 732

TBILLEQ Computes the bond-equivalent yield for a treasury bill.

FINANCE('TBILLEQ', *settlement*, *maturity*, *discount*);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

discount
specifies the discount rate of the security.

Featured in: Example 44 on page 732

TBILLPRICE Computes the price of a treasury bill per $100 face value.

FINANCE('TBILLPRICE', *settlement*, *maturity*, *discount*);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

discount
specifies the discount rate of the security.

See

Featured in: Example 45 on page 732

TBILLYIELD Computes the yield for a treasury bill.

FINANCE('TBILLYIELD', *settlement*, *maturity*, *pr*);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

pr
specifies the price of the security per $100 face value.

Featured in: Example 46 on page 733

VDB Computes the depreciation of an asset for a specified or partial period by using a declining balance method.

FINANCE('VDB', *cost*, *salvage*, *life*, *start-period*, *end-period*, <*factor*>, <*noswitch*>);

where

cost
specifies the initial cost of the asset.

salvage
specifies the value at the end of the depreciation (also called the salvage value of the asset).

life
specifies the number of periods over which the asset is depreciated (also called the useful life of the asset).

start-period
specifies the first period in the calculation. Payment periods are numbered beginning with 1.

end-period
specifies the last period in the calculation.

factor
specifies the rate at which the balance declines. If *factor* is omitted, it is assumed to be 2 (the double-declining balance method).

noswitch
specifies a logical value that determines whether to switch to straight-line depreciation when the depreciation is greater than the declining balance calculation. If *noswitch* is omitted, it is assumed to be 1.

Featured in: Example 47 on page 733

XIRR Computes the internal rate of return for a schedule of cash flows that is not necessarily periodic.

FINANCE('XIRR', *values*, *dates*, <*guess*>);

where

values
specifies a series of cash flows that corresponds to a schedule of payments in dates. The first payment is optional and corresponds to a cost or payment that occurs at the beginning of the investment. If the first value is a cost or payment, it must be a negative value. All succeeding payments are discounted based on a 365-day year. The series of values must contain at least one positive value and one negative value.

dates
specifies a schedule of payment dates that corresponds to the cash flow payments. The first payment date indicates the beginning of the schedule of payments. All other dates must be later than this date, but they can occur in any order.

guess
specifies an optional number that you guess is close to the result of XIRR.

Featured in: Example 48 on page 733

XNPV Computes the net present value for a schedule of cash flows that is not necessarily periodic.

FINANCE('XNPV', *rate*, *values*, *dates*);

where

rate
specifies the interest rate.

values
specifies a series of cash flows that corresponds to a schedule of payments in dates. The first payment is optional and corresponds to a cost or payment that occurs at the beginning of the investment. If the first value is a cost or payment, it must be a negative value. All succeeding payments are discounted based on a 365-day year. The series of values must contain at least one positive value and one negative value.

dates
specifies a schedule of payment dates that corresponds to the cash flow payments. The first payment date indicates the beginning of the schedule of payments. All other dates must be later than this date, but they can occur in any order.

Featured in: Example 49 on page 733

YIELD Computes the yield on a security that pays periodic interest.

FINANCE('YIELD', *settlement*, *maturity*, *rate*, *pr*, *redemption*, *frequency*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

rate
specifies the interest rate.

pr
specifies the price of the security per $100 face value.

redemption
specifies the amount to be received at maturity.

frequency
specifies the number of coupon payments per year. For annual payments, *frequency* = 1; for semiannual payments, *frequency* = 2; for quarterly payments, *frequency* = 4.

basis
specifies the optional day count value.

Featured in: Example 50 on page 734

YIELDDISC Computes the annual yield for a discounted security (for example, a treasury bill).

FINANCE('YIELDDISC', *settlement*, *maturity*, *rate*, *pr*, *redemption*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

rate
specifies the interest rate.

pr
specifies the price of the security per $100 face value.

redemption
specifies the amount to be received at maturity.

basis
specifies the optional day count value.

Featured in: Example 51 on page 734

YIELDMAT Computes the annual yield of a security that pays interest at maturity.

FINANCE('YIELDMAT', *settlement*, *maturity*, *issue*, *rate*, *pr*, <*basis*>);

where

settlement
specifies the settlement date.

maturity
specifies the maturity date.

issue
specifies the issue date of the security.

rate
specifies the interest rate.

pr
specifies the price of the security per $100 face value.

basis
specifies the optional day count value.

Featured in: Example 52 on page 734

Examples

Example 1: Computing Accrued Interest: ACCRINT The following example computes the accrued interest for a security that pays periodic interest.

```
data _null_;
   issue = mdy(2,27,1996);
   firstinterest = mdy(8,31,1998);
   settlement = mdy(5,1,1998);
   rate = 0.1;
   par = 1000;
   frequency = 2;
   basis = 1;
   r = finance('accrint', issue, firstinterest,
               settlement, rate, par, frequency, basis);
   put r=;
run;
```

The value of *r* that is returned is 217.39728.

Example 2: Computing Accrued Interest: ACCRINTM The following example computes the accrued interest for a security that pays interest at maturity.

```
data _null_;
   issue = mdy(2,28,1998);
   maturity = mdy(8,31,1998);
   rate = 0.1;
   par = 1000;
   basis = 0;
   r = finance('accrintm', issue, maturity, rate, par, basis);
   put r=;
run;
```

The value of *r* that is returned is 50.555555556.

Example 3: Computing Depreciation: AMORDEGRC The following example computes the depreciation for each accounting period by using a depreciation coefficient.

```
data _null_;
   cost = 2400;
   datepurchased = mdy(8,19,2008);
   firstperiod = mdy(12,31,2008);
   salvage = 300;
   period = 1;
   rate = 0.15;
   basis = 1;
   r = finance('amordegrc', cost, datepurchased,
               firstperiod, salvage, period, rate, basis);
   put r=;
run;
```

The value of *r* that is returned is 776.

Example 4: Computing Description: AMORLINC The following example computes the depreciation for each accounting period.

```
data _null_;
   cost = 2400;
   dp = mdy(9,30,1998);
   fp = mdy(12,31,1998);
   salvage = 245;
   period = 0;
   rate = 0.115;
   basis = 0;
   r = finance('amorlinc', cost, dp, fp, salvage,
                period, rate, basis);
   put r = ;
run;
```

The value of *r* that is returned is 69.

Example 5: Computing Description: COUPDAYBS The following example computes the number of days from the beginning of the coupon period to the settlement date.

```
data _null_;
   settlement = mdy(12,30,1994);
   maturity = mdy(11,29,1997);
   frequency = 4;
   basis = 2;
   r = finance('coupdaybs', settlement, maturity, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 31.

Example 6: Computing Description: COUPDAYS The following example computes the number of days in the coupon period that contains the settlement date.

```
data _null_;
   settlement = mdy(1,25,2007);
   maturity = mdy(11,15,2008);
   frequency = 2;
   basis = 1;
   r = finance('coupdays', settlement, maturity, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 181.

Example 7: Computing Description: COUPDAYSNC The following example computes the number of days from the settlement date to the next coupon date.

```
data _null_;
   settlement = mdy(1,25,2007);
   maturity = mdy(11,15,2008);
   frequency = 2;
   basis = 1;
   r = finance('coupdaysnc', settlement, maturity, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 110.

Example 8: Computing Description: COUPNCD The following example computes the next coupon date after the settlement date.

```
data _null_;
   settlement = mdy(1,25,2007);
   maturity = mdy(11,15,2008);
   frequency = 2;
   basis = 1;
   r = finance('coupncd', settlement, maturity, frequency, basis);
   put r = date7.;
run;
```

The value of *r* that is returned is 15MAY07.

Note: *r* is a numeric SAS value and can be printed using the DATE7 format. △

Example 9: Computing Description: COUPNUM The following example computes the number of coupons that are payable between the settlement date and the maturity date.

```
data _null_;
   settlement = mdy(1,25,2007);
   maturity = mdy(11,15,2008);
   frequency = 2;
   basis = 1;
   r = finance('coupnum', settlement, maturity, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 4.

Example 10: Computing Description: COUPPCD The following example computes the previous coupon date before the settlement date.

```
data _null_;
   settlement = mdy(1,25,2007);
   maturity = mdy(11,15,2008);
   frequency = 2;
   basis = 1;
   r = finance('couppcd', settlement, maturity, frequency, basis);
   put  settlement;
   put  maturity;
   put r date7.;
run;
```

The value of *r* that is returned is 11/15/2006.

Example 11: Computing Description: CUMIPMT The following example computes the cumulative interest that is paid between two periods.

```
data _null_;
   rate = 0.09;
   nper = 30;
   pv = 125000;
   startperiod = 13;
   endperiod = 24;
   type = 0;
   r = finance('cumipmt', rate, nper, pv,
               startperiod, endperiod, type);


   put r = ;
run;
```

The value of *r* that is returned is –94054.82033.

Example 12: Computing Description: CUMPRINC The following example computes the cumulative principal that is paid on a loan between two periods.

```
data _null_;
   rate = 0.09;
   nper = 30;
   pv = 125000;
   startperiod = 13;
   endperiod = 24;
   type = 0;
   r = finance('cumprinc', rate, nper, pv,
               startperiod, endperiod, type);
   put r = ;
run;
```

The value of *r* that is returned is –51949.70676.

Example 13: Computing Description: DB The following example computes the depreciation of an asset for a specified period by using the fixed-declining balance method.

```
data _null_;
   cost = 1000000;
   salvage = 100000;
   life = 6;
   period = 2;
   month = 7;
   r = finance('db', cost, salvage, life, period, month);
   put r = ;
 run;
```

The value of *r* that is returned is 259639.41667.

Example 14: Computing Description: DDB The following example computes the depreciation of an asset for a specified period by using the double-declining balance method or some other method that you specify.

```
data _null_;
   cost = 2400;
   salvage = 300;
   life = 10*365;
   period = 1;
   factor = .;
   r = finance('ddb', cost, salvage, life, period, factor);
   put r = ;
run;
```

The value of *r* that is returned is 1.3150684932.

Example 15: Computing Description: DISC The following example computes the discount rate for a security.

```
data _null_;
   settlement = mdy(1,25,2007);
   maturity = mdy(6,15,2007);
```

```
   pr = 97.975;
   redemption = 100;
   basis = 1;
   r = finance('disc', settlement, maturity, pr, redemption, basis);
   put r = ;
run;
```

The value of *r* that is returned is 0.052420213.

Example 16: Computing Description: DOLLARDE The following example converts a dollar price, expressed as a fraction, to a dollar price, expressed as a decimal number.

```
data _null_;
   fractionaldollar = 1.125;
   fraction = 16;
   r = finance('dollarde', fractionaldollar, fraction);
   put r = ;
run;
```

The value of *r* that is returned is 1.78125.

Example 17: Computing Description: DOLLARFR The following example converts a dollar price, expressed as a decimal number, to a dollar price, expressed as a fraction.

```
data _null_;
   decimaldollar = 1.125;
   fraction = 16;
   r = finance('dollarfr', decimaldollar, fraction);
   put r = ;
run;
```

The value of *r* that is returned is 1.02. In fraction form, the value of *r* is read as $1\ ^{2}/_{16}$.

Example 18: Computing Description: DURATION The following example computes the annual duration of a security with periodic interest payments.

```
data _null_;
   settlement = mdy(1,1,2008);
   maturity = mdy(1,1,2016);
   couponrate = 0.08;
   yield = 0.09;
   frequency = 2;
   basis = 1;
   r = finance('duration', settlement,
          maturity, couponrate, yield, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 5.993775.

Example 19: Computing Description: EFFECT The following example computes the effective annual interest rate.

```
data _null_;
   nominalrate = 0.0525;
   npery = 4;
   r = finance('effect', nominalrate, npery);
```

```
   put r = ;
run;
```

The value of *r* that is returned is 0.053543.

Example 20: Computing Description: FV The following example computes the future value of an investment.

```
data _null_;
   rate = 0.06/12;
   nper = 10;
   pmt = -200;
   pv = -500;
   type = 1;
   r = finance('fv', rate, nper, pmt, pv, type);
   put r = ;
run;
```

The value of *r* that is returned is 2581.4033741.

Example 21: Computing Description: FVSCHEDULE The following example computes the future value of the initial principal after applying a series of compound interest rates.

```
data _null_;
   principal = 1;
   r1 = 0.09;
   r2 = 0.11;
   r3 = 0.1;
   r = finance('fvschedule', principal, r1, r2, r3);
   put r = ;
run;
```

The value of *r* that is returned is 1.33089.

Example 22: Computing Description: INTRATE The following example computes the interest rate for a fully invested security.

```
data _null_;
   settlement = mdy(2,15,2008);
   maturity = mdy(5,15,2008);
   investment = 1000000;
   redemption = 1014420;
   basis = 2;
   r = finance('intrate', settlement, maturity,
        investment, redemption, basis);
   put r = ;
run;
```

The value of *r* that is returned is 0.05768

Example 23: Computing Description: IPMT The following example computes the interest payment for an investment for a specified period.

```
data _null_;
   rate = 0.1/12;
   per = 2;
   nper = 3;
   pv = 100;
```

```
   fv = .;
   type = .;
   r = finance('ipmt', rate, per, nper, pv, fv, type);
   put r = ;
run;
```

The value of *r* that is returned is –0.557857564.

Example 24: Computing Description: IRR The following example computes the internal rate of return for a series of cash flows.

```
data _null_;
   v1 = -70000;
   v2 = 12000;
   v3 = 15000;
   v4 = 18000;
   v5 = 21000;
   v6 = 26000;
   r = finance('irr', v1, v2, v3, v4, v5, v6);
   put r = ;
run;
```

The value of *r* that is returned is 0.086630948.

Example 25: Computing Description: MDURATION The following example computes the Macaulay modified duration for a security with an assumed face value of $100.

```
data _null_;
   settlement = mdy(1,1,2008);
   maturity = mdy(1,1,2016);
   couponrate = 0.08;
   yield = 0.09;
   frequency = 2;
   basis = 1;
   r = finance('mduration', settlement, maturity,
     couponrate, yield, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 5.7356698139.

Example 26: Computing Description: MIRR The following example computes the internal rate of return where positive and negative cash flows are financed at different rates.

```
data _null_;
   v1 = -1000;
   v2 = 3000;
   v3 = 4000;
   v4 = 5000;
   financerate = 0.08;
   reinvestrate = 0.10;
   r = finance('mirr', v1, v2, v3, v4, financerate, reinvestrate);
   put r = ;
run;
```

The value of *r* that is returned is 1.3531420172.

Example 27: Computing Description: NOMINAL The following example computes the annual nominal interest rate.

```
data _null_;
   effectrate = 0.08;
   npery = 4;
   r = finance('nominal', effectrate, npery);
   put r = ;
run;
```

The value of *r* that is returned is 0.0777061876.

Example 28: Computing Description: NPER The following example computes the number of periods for an investment.

```
data _null_;
   rate = 0.08;
   pmt = 200;
   pv = 1000;
   fv = 0;
   type = 0;
   r = finance('nper', rate, pmt, pv, fv, type);
   put r = ;
run;
```

The value of *r* that is returned is –4.371981351.

Example 29: Computing Description: NPV The following example computes the net present value of an investment based on a series of periodic cash flows and a discount rate.

```
data _null_;
   rate = 0.08;
   v1 = 200;
   v2 = 1000;
   v3 = 0.;
   r = finance('npv', rate, v1, v2, v3);
   put r = ;
run;
```

The value of *r* that is returned is 1042.5240055.

Example 30: Computing Description: ODDFPRICE The following example computes the price of a security per $100 face value with an odd first period.

```
data _null_;
   settlement = mdy(1,15,93);
   maturity = mdy(1,1,98);
   issue = mdy(1,1,93);
   firstcoupon = mdy(7,1,94);
   rate = 0.07;
   yld = 0.06;
   redemption = 100;
   frequency = 2;
   basis = 0;
   r = finance('oddfprice',
    settlement, maturity, issue, firstcoupon, rate, yld, redemption,
    frequency, basis);
```

```
   put r = ;
run;
```

The value of *r* that is returned is 103.94103984.

Example 31: Computing Description: ODDFYIELD The following example computes the interest of a yield with an odd first period.

```
data _null_;
   settlement = mdy(1,15,93);
   maturity = mdy(1,1,98);
   issue = mdy(1,1,93);
   firstcoupon = mdy(7,1,94);
   rate = 0.07;
   pr = 103.94103984;
   redemption = 100;
   frequency = 2;
   basis = 0;
   r = finance('oddfyield',
    settlement, maturity, issue, firstcoupon, rate, pr, redemption,
    frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 0.06.

Example 32: Computing Description: ODDLPRICE The following example computes the price of a security per $100 face value with an odd last period.

```
data _null_;
   settlement = mdy(2,7,2008);
   maturity = mdy(6,15,2008);
   lastinterest = mdy(10,15,2007);
   rate = 0.0375;
   yield = 0.0405;
   redemption = 100;
   frequency = 2;
   basis = 0;
   r = finance('oddlprice', settlement, maturity, lastinterest,
   rate, yield, redemption, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 99.878286015.

Example 33: Computing Description: ODDLYIELD The following example computes the yield of a security with an odd last period.

```
data _null_;
   settlement = mdy(2,7,2008);
   maturity = mdy(6,15,2008);
   lastinterest = mdy(10,15,2007);
   rate = 0.0375;
   pr = 99.878286015;
   redemption = 100;
   frequency = 2;
   basis = 0;
   r = finance('oddlyield', settlement, maturity, lastinterest,
```

```
   rate, pr, redemption, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 0.0405.

Example 34: Computing Description: PMT The following example computes the periodic payment for an annuity.

```
data _null_;
   rate = 0.08;
   nper = 5;
   pv = 91;
   fv = 3;
   type = 0;
   r = finance('pmt', rate, nper, pv, fv, type);
   put r = ;
run;
```

The value of *r* that is returned is -23.30290673.

Example 35: Computing Description: PPMT The following example computes the payment on the principal for an investment for a specified period.

```
data _null_;
   rate = 0.08;
   per = 10;
   nper = 10;
   pv = 200000;
   fv = 0;
   type = 0;
   r = finance('ppmt', rate, per, nper, pv, fv, type);
   put r = ;
run;
```

The value of *r* that is returned is –27598.05346.

Example 36: Computing Description: PRICE The following example computes the price of a security per $100 face value that pays periodic interest.

```
data _null_;
   settlement = mdy(2,15,2008);
   maturity = mdy(11,15,2017);
   rate = 0.0575;
   yield = 0.065;
   redemption = 100;
   frequency = 2;
   basis = 0;
   r = finance('price', settlement, maturity, rate, yield, redemption,
     frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 94.634361621.

Example 37: Computing Description: PRICEDISC The following example computes the price of a discounted security per $100 face value.

```
data _null_;
   settlement = mdy(2,15,2008);
   maturity = mdy(11,15,2017);
   discount = 0.0525;
   redemption = 100;
   basis = 0;
   r = finance('pricedisc', settlement, maturity, discount, redemption, basis);
   put r = ;
run;
```

The value of *r* that is returned is 48.8125.

Example 38: Computing Description: PRICEMAT The following example computes the price of a security per $100 face value that pays interest at maturity.

```
data _null_;
   settlement = mdy(2,15,2008);
   maturity = mdy(4,13,2008);
   issue = mdy(11,11,2007);
   rate = 0.061;
   yield = 0.061;
   basis = 0;
   r = finance('pricemat', settlement, maturity, issue, rate, yield, basis);
   put r = ;
run;
```

The value of *r* that is returned is 99.98449888.

Example 39: Computing Description: PV The following example computes the present value of an investment.

```
data _null_;
   rate = 0.05;
   nper = 10;
   pmt = 1000;
   fv = 200;
   type = 0;
   r = finance('pv', rate, nper, pmt, fv, type);
   put r = ;
run;
```

The value of *r* that is returned is –7844.51758.

Example 40: Computing Description: RATE The following example computes the interest rate per period of an annuity.

```
data _null_;
   nper = 4;
   pmt = -2481;
   pv = 8000;
   r = finance('rate', nper, pmt, pv);
   put r = ;
run;
```

The value of *r* that is returned is 0.0921476841.

Example 41: Computing Description: RECEIVED The following example computes the amount that is received at maturity for a fully invested security.

```
data _null_;
   settlement = mdy(2,15,2008);
   maturity = mdy(5,15,2008);
   investment = 1000000;
   discount = 0.0575;
   basis = 2;
   r = finance('received', settlement, maturity, investment, discount, basis);
   put r = ;
run;
```

The value of *r* that is returned is 1014584.6544.

Example 42: Computing Description: SLN The following example computes the straight-line depreciation of an asset for one period.

```
data _null_;
   cost = 2000;
   salvage = 200;
   life = 11;
   r = finance('sln', cost, salvage, life);
   put r = ;
run;
```

The value of *r* that is returned is 163.63636364.

Example 43: Computing Description: SYD The following example computes the sum-of-years digits depreciation of an asset for a specified period.

```
data _null_;
   cost = 2000;
   salvage = 200;
   life = 11;
   per = 1;
   r = finance('syd', cost, salvage, life, per);
   put r = ;
run;
```

The value of *r* that is returned is 300.

Example 44: Computing Description: TBILLEQ The following example computes the bond-equivalent yield for a treasury bill.

```
data _null_;
   settlement = mdy(3,31,2008);
   maturity = mdy(6,1,2008);
   discount = 0.0914;
   r = finance('tbilleq', settlement, maturity, discount);
   put r = ;
run;
```

The value of *r* that is returned is 0.0941514936.

Example 45: Computing Description: TBILLPRICE The following example computes the price of a treasury bill per $100 face value.

```
data _null_;
   settlement = mdy(3,31,2008);
   maturity = mdy(6,1,2008);
   discount = 0.09;
```

```
   r = finance('tbillprice', settlement, maturity, discount);
   put r = ;
run;
```

The value of *r* that is returned is 98.45.

Example 46: Computing Description: TBILLYIELD The following example computes the yield for a treasury bill.

```
data _null_;
   settlement = mdy(3,31,2008);
   maturity = mdy(6,1,2008);
   pr = 98;
   r = finance('tbillyield', settlement, maturity, pr);
   put r = ;
run;
```

The value of *r* that is returned is 0.1184990125.

Example 47: Computing Description: VDB The following example computes the depreciation of an asset for a specified or partial period by using a declining balance method.

```
data _null_;
   cost = 2400;
   salvage = 300;
   life = 10;
   startperiod = 0;
   endperiod = 1;
   factor = 1.5;
   r = finance('vdb', cost, salvage, life, startperiod, endperiod, factor);
   put r = ;
run;
```

The value of *r* that is returned is 360.

Example 48: Computing Description: XIRR The following example computes the internal rate of return for a schedule of cash flows that is not necessarily periodic.

```
data _null_;
   v1 = -10000; d1 = mdy(1,1,2008);
   v2 = 2750; d2 = mdy(3,1,2008);
   v3 = 4250; d3 = mdy(10,30,2008);
   v4 = 3250; d4 = mdy(2,15,2009);
   v5 = 2750; d5 = mdy(4,1,2009);
   r = finance('xirr', v1, v2, v3, v4, v5, d1, d2, d3, d4, d5, 0.1);
   put r = ;
run;
```

The value of *r* that is returned is 0.3733625335.

Example 49: Computing Description: XNPV The following example computes the net present value for a schedule of cash flows that is not necessarily periodic.

```
data _null_;
   r = 0.09;
   v1 = -10000; d1 = mdy(1,1,2008);
   v2 = 2750; d2 = mdy(3,1,2008);
   v3 = 4250; d3 = mdy(10,30,2008);
```

```
   v4 = 3250; d4 = mdy(2,15,2009);
   v5 = 2750; d5 = mdy(4,1,2009);
   r = finance('xnpv', r, v1, v2, v3, v4, v5, d1, d2, d3, d4, d5);
   put r = ;
run;
```

The value of *r* that is returned is 2086.647602.

Example 50: Computing Description: YIELD The following example computes the yield on a security that pays periodic interest.

```
data _null_;
   settlement = mdy(2,15,2008);
   maturity = mdy(11,15,2016);
   rate = 0.0575;
   pr = 95.04287;
   redemption = 100;
   frequency = 2;
   basis = 0;
   r = finance('yield', settlement, maturity, rate, pr, redemption, frequency, basis);
   put r = ;
run;
```

The value of *r* that is returned is 0.0650000069.

Example 51: Computing Description: YIELDDISC The following example computes the annual yield for a discounted security (for example, a treasury bill).

```
data _null_;
   settlement = mdy(2,15,2008);
   maturity = mdy(11,15,2016);
   pr = 95.04287;
   redemption = 100;
   basis = 0;
   r = finance('yielddisc', settlement, maturity, pr, redemption, basis);
   put r = ;
run;
```

The value of *r* that is returned is 0.0059607748.

Example 52: Computing Description: YIELDMAT The following example computes the annual yield of a security that pays interest at maturity.

```
data _null_;
   settlement = mdy(3,15,2008);
   maturity = mdy(11,3,2008);
   issue = mdy(11,8,2007);
   rate = 0.0625;
   pr = 100.0123;
   basis = 0;
   r = finance('yieldmat', settlement, maturity, issue, rate, pr, basis);
   put r = ;
run;
```

The value of *r* that is returned is 0.0609543337.

FIND Function

Searches for a specific substring of characters within a character string.

Category: Character

Restriction: "I18N Level 1" on page 314

Tip: Use the KINDEX function in *SAS National Language Support (NLS): Reference Guide* instead to write encoding independent code.

Syntax

FIND(*string,substring*<,*modifiers*><,*startpos*>)

FIND(*string,substring*<,*startpos*><,*modifiers*>)

Arguments

string
specifies a character constant, variable, or expression that will be searched for substrings.

Tip: Enclose a literal string of characters in quotation marks.

substring
is a character constant, variable, or expression that specifies the substring of characters to search for in *string*.

Tip: Enclose a literal string of characters in quotation marks.

modifiers
is a character constant, variable, or expression that specifies one or more modifiers. The following *modifiers* can be in uppercase or lowercase:

i — ignores character case during the search. If this modifier is not specified, FIND only searches for character substrings with the same case as the characters in *substring*.

t — trims trailing blanks from *string* and *substring*.

Note: If you want to remove trailing blanks from only one character argument instead of both (or all) character arguments, use the TRIM function instead of the FIND function with the T modifier. △

Tip: If *modifier* is a constant, enclose it in quotation marks. Specify multiple constants in a single set of quotation marks. *Modifier* can also be expressed as a variable or an expression.

startpos
is a numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction of the search.

Details

The FIND function searches *string* for the first occurrence of the specified *substring*, and returns the position of that substring. If the substring is not found in *string*, FIND returns a value of 0.

If *startpos* is not specified, FIND starts the search at the beginning of the *string* and searches the *string* from left to right. If *startpos* is specified, the absolute value of *startpos* determines the position at which to start the search. The sign of *startpos* determines the direction of the search.

Value of *startpos*	Action
greater than 0	starts the search at position *startpos* and the direction of the search is to the right. If *startpos* is greater than the length of *string*, FIND returns a value of 0.
less than 0	starts the search at position *–startpos* and the direction of the search is to the left. If *–startpos* is greater than the length of *string*, the search starts at the end of *string*.
equal to 0	returns a value of 0.

Comparisons

- The FIND function searches for substrings of characters in a character string, whereas the FINDC function searches for individual characters in a character string.
- The FIND function and the INDEX function both search for substrings of characters in a character string. However, the INDEX function does not have the *modifiers* nor the *startpos* arguments.

Examples

SAS Statements	Results
`whereisshe=find('She sells seashells? Yes, she does.','she ');` `put whereisshe;`	`27`

`variable1='She sells seashells? Yes, she does.';`
`variable2='she ';`
`variable3='i';`
`whereisshe_i=find(variable1,variable2,variable3);`
`put whereisshe_i;`
	`1`

`expression1='She sells seashells? '||'Yes, she does.';`
`expression2=kscan('he or she',3)||'  ';`
`expression3=trim('t   ');`
`whereisshe_t=find(expression1,expression2,expression3);`
`put whereisshe_t;`
	`14`

`xyz='She sells seashells? Yes, she does.';`
`startposvar=22;`
`whereisshe_22=find(xyz,'she',startposvar);`
`put whereisshe_22;`
	`27`

`xyz='She sells seashells? Yes, she does.';`
`startposexp=1-23;`
`whereisShe_ineg22=find(xyz,'She','i',startposexp);`
`put whereisShe_ineg22;`
	`14`

See Also

Functions:

"COUNT Function" on page 616
"FINDC Function" on page 737
"INDEX Function" on page 817

FINDC Function

Searches a string for any character in a list of characters.

Category: Character

Restriction: "I18N Level 1" on page 314

Tip: Use the KINDEXC function in *SAS National Language Support (NLS): Reference Guide* instead to write encoding independent code.

Syntax

FINDC(*string* <, *charlist*>)

FINDC(*string*, *charlist* <, *modifiers*>)

FINDC(*string*, *charlist*, *modifier(s)* <, *startpos*>)

FINDC(*string*, *charlist*, <*startpos*>, <*modifiers*)>

Arguments

string
is a character constant, variable, or expression that specifies the character string to be searched.

Tip: Enclose a literal string of characters in quotation marks.

charlist
is an optional constant, variable, or character expression that initializes a list of characters. FINDC searches for the characters in this list provided that you do not specify the K modifier in the *modifier* argument. If you specify the K modifier, FINDC searches for all characters that are not in this list of characters. You can add more characters to the list by using other modifiers.

modifier
is an optional character constant, variable, or expression in which each character modifies the action of the FINDC function. The following characters, in upper- or lowercase, can be used as modifiers:

blank	is ignored.
a or A	adds alphabetic characters to the list of characters.
b or B	searches from right to left, instead of from left to right, regardless of the sign of the *startpos* argument.

c or C	adds control characters to the list of characters.
d or D	adds digits to the list of characters.
f or F	adds an underscore and English letters (that is, the characters that can begin a SAS variable name using VALIDVARNAME=V7) to the list of characters.
g or G	adds graphic characters to the list of characters.
h or H	adds a horizontal tab to the list of characters.
i or I	ignores character case during the search.
k or K	searches for any character that does not appear in the list of characters. If you do not specify this modifier, then FINDC searches for any character that appears in the list of characters.
l or L	adds lowercase letters to the list of characters.
n or N	adds digits, an underscore, and English letters (that is, the characters that can appear in a SAS variable name using VALIDVARNAME=V7) to the list of characters.
o or O	processes the *charlist* and the *modifier* arguments only once, rather than every time the FINDC function is called. Using the O modifier in the DATA step (excluding WHERE clauses), or in the SQL procedure can make FINDC run faster when you call it in a loop where the *charlist* and the *modifier* arguments do not change.
p or P	adds punctuation marks to the list of characters.
s or S	adds space characters to the list of characters (blank, horizontal tab, vertical tab, carriage return, line feed, and form feed).
t or T	trims trailing blanks from the *string* and *charlist* arguments. *Note:* If you want to remove trailing blanks from just one character argument instead of both (or all) character arguments, use the TRIM function instead of the FINDC function with the T modifier. △
u or U	adds uppercase letters to the list of characters.
w or W	adds printable characters to the list of characters.
x or X	adds hexadecimal characters to the list of characters.

Tip: If *modifier* is a constant, then enclose it in quotation marks. Specify multiple constants in a single set of quotation marks. *Modifier* can also be expressed as a variable or an expression.

startpos
is an optional numeric constant, variable, or expression having an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The FINDC function searches *string* for the first occurrence of the specified characters, and returns the position of the first character found. If no characters are found in *string*, then FINDC returns a value of 0.

The FINDC function allows character arguments to be null. Null arguments are treated as character strings that have a length of zero. Numeric arguments cannot be null.

If *startpos* is not specified, FINDC begins the search at the end of the string if you use the B modifier, or at the beginning of the string if you do not use the B modifier.

If *startpos* is specified, the absolute value of *startpos* specifies the position at which to begin the search. If you use the B modifier, the search always proceeds from right to left. If you do not use the B modifier, the sign of *startpos* specifies the direction in which to search. The following table summarizes the search directions:

Value of *startpos*	Action
greater than 0	search begins at position *startpos* and proceeds to the right. If *startpos* is greater than the length of the string, FINDC returns a value of 0.
less than 0	search begins at position *–startpos* and proceeds to the left. If *startpos* is less than the negative of the length of the string, the search begins at the end of the string.
equal to 0	returns a value of 0.

Comparisons

- The FINDC function searches for individual characters in a character string, whereas the FIND function searches for substrings of characters in a character string.
- The FINDC function and the INDEXC function both search for individual characters in a character string. However, the INDEXC function does not have the *modifier* nor the *startpos* arguments.
- The FINDC function searches for individual characters in a character string, whereas the VERIFY function searches for the first character that is unique to an expression. The VERIFY function does not have the *modifier* nor the *startpos* arguments.

Examples

Example 1: Searching for Characters in a String This example searches a character string and returns the characters that are found.

```
data _null_;
   string = 'Hi, ho!';
   charlist = 'hi';
   j = 0;
   do until (j = 0);
      j = findc(string, charlist, j+1);
      if j = 0 then put +3 "That's all";
      else do;
         c = substr(string, j, 1);
         put +3 j= c=;
      end;
   end;
run;
```

SAS writes the following output to the log:

```
j=2 c=i
j=5 c=h
That's all
```

Example 2: Searching for Characters in a String and Ignoring Case This example searches a character string and returns the characters that are found. The I modifier is used to ignore the case of the characters.

```
data _null_;
   string = 'Hi, ho!';
   charlist = 'ho';
   j = 0;
   do until (j = 0);
      j = findc(string, charlist, j+1, "i");
      if j = 0 then put +3 "That's all";
      else do;
         c = substr(string, j, 1);
         put +3 j= c=;
      end;
   end;
run;
```

SAS writes the following output to the log:

```
j=1 c=H
j=5 c=h
j=6 c=o
That's all
```

Example 3: Searching for Characters and Using the K Modifier This example searches a character string and returns the characters that do not appear in the character list.

```
data _null_;
   string = 'Hi, ho!';
   charlist = 'hi';
   j = 0;
   do until (j = 0);
      j = findc(string, charlist, "k", j+1);
      if j = 0 then put +3 "That's all";
      else do;
         c = substr(string, j, 1);
         put +3 j= c=;
      end;
   end;
run;
```

SAS writes the following output to the log:

```
j=1 c=H
j=3 c=,
j=4 c=
j=6 c=o
j=7 c=!
That's all
```

Example 4: Searching for the Characters h, i, and Blank This example searches for the three characters h, i, and blank. The characters h and i are in lowercase. The uppercase characters H and I are ignored in this search.

```
data _null_;
   whereishi=0;
   do until(whereishi=0);
      whereishi=findc('Hi there, Ian!','hi ',whereishi+1);
      if whereishi=0 then put ''The End'';
      else do;
         whatfound=substr('Hi there, Ian!',whereishi,1);
         put whereishi= whatfound=;
      end;
   end;
run;
```

SAS writes the following output to the log:

```
whereishi=2 whatfound=i
whereishi=3 whatfound=
whereishi=5 whatfound=h
whereishi=10 whatfound=
The End
```

Example 5: Searching for the Characters h and i While Ignoring Case This example searches for the four characters h, i, H, and I. FINDC with the i modifier ignores character case during the search.

```
data _null_;
   whereishi_i=0;
   do until(whereishi_i=0);
      variable1='Hi there, Ian!';
      variable2='hi';
      variable3='i';
      whereishi_i=findc(variable1,variable2,variable3,whereishi_i+1);
      if whereishi_i=0 then put ''The End'';
      else do;
         whatfound=substr(variable1,whereishi_i,1);
         put whereishi_i= whatfound=;
      end;
   end;
run;
```

SAS writes the following output to the log:

```
whereishi_i=1 whatfound=H
whereishi_i=2 whatfound=i
whereishi_i=5 whatfound=h
whereishi_i=11 whatfound=I
The End
```

Example 6: Searching for the Characters h and i with Trailing Blanks Trimmed This example searches for the two characters h and i. FINDC with the t modifier trims trailing blanks from the string argument and the characters argument.

```
data _null_;
   whereishi_t=0;
   do until(whereishi_t=0);
```

```
      expression1='Hi there, '||'Ian!';
      expression2=kscan('bye or hi',3)||'  ';
      expression3=trim('t   ');
      whereishi_t=findc(expression1,expression2,expression3,whereishi_t+1);
      if whereishi_t=0 then put ''The End'';
      else do;
         whatfound=substr(expression1,whereishi_t,1);
         put whereishi_t= whatfound=;
      end;
   end;
run;
```

SAS writes the following lines output to the log:

```
whereishi_t=2 whatfound=i
whereishi_t=5 whatfound=h
The End
```

Example 7: Searching for all Characters, Excluding h, i, H, and I This example searches for all of the characters in the string, excluding the characters h, i, H, and I. FINDC with the v modifier counts only the characters that do not appear in the characters argument. This example also includes the i modifier and therefore ignores character case during the search.

```
data _null_;
   whereishi_iv=0;
   do until(whereishi_iv=0);
      xyz='Hi there, Ian!';
      whereishi_iv=findc(xyz,'hi',whereishi_iv+1,'iv');
      if whereishi_iv=0 then put ''The End'';
      else do;
         whatfound=substr(xyz,whereishi_iv,1);
         put whereishi_iv= whatfound=;
      end;
   end;
run;
```

SAS writes the following output to the log:

```
whereishi_iv=3 whatfound=
whereishi_iv=4 whatfound=t
whereishi_iv=6 whatfound=e
whereishi_iv=7 whatfound=r
whereishi_iv=8 whatfound=e
whereishi_iv=9 whatfound=,
whereishi_iv=10 whatfound=
whereishi_iv=12 whatfound=a
whereishi_iv=13 whatfound=n
whereishi_iv=14 whatfound=!
The End
```

See Also

Functions:

"ANYALNUM Function" on page 379
"ANYALPHA Function" on page 381

"ANYCNTRL Function" on page 383
"ANYDIGIT Function" on page 384
"ANYGRAPH Function" on page 388
"ANYLOWER Function" on page 390
"ANYPRINT Function" on page 394
"ANYPUNCT Function" on page 396
"ANYSPACE Function" on page 397
"ANYUPPER Function" on page 399
"ANYXDIGIT Function" on page 401
"COUNTC Function" on page 618
"INDEXC Function" on page 819
"VERIFY Function" on page 1193
"NOTALNUM Function" on page 948
"NOTALPHA Function" on page 950
"NOTCNTRL Function" on page 952
"NOTDIGIT Function" on page 954
"NOTGRAPH Function" on page 959
"NOTLOWER Function" on page 961
"NOTPRINT Function" on page 965
"NOTPUNCT Function" on page 967
"NOTSPACE Function" on page 969
"NOTUPPER Function" on page 971
"NOTXDIGIT Function" on page 973

FINDW Function

Returns the character position of a word in a string, or returns the number of the word in a string.

Category: Character

Syntax

FINDW(*string*, *word* <, *chars*>)

FINDW(*string*, *word*, *chars*, *modifiers* <, *startpos*>)

FINDW(*string*, *word*, *chars*, *startpos* <, *modifiers*>)

FINDW(*string*, *word*, *startpos* <, *chars* <, *modifiers*>>)

Arguments

string
is a character constant, variable, or expression that specifies the character string to be searched.

word
is a character constant, variable, or expression that specifies the word to be searched.

chars
is an optional character constant, variable, or expression that initializes a list of characters.

The characters in this list are the delimiters that separate words, provided that you do not specify the K modifier in the *modifier* argument. If you specify the K modifier, then all characters that are not in this list are delimiters. You can add more characters to this list by using other modifiers.

startpos
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should begin and the direction in which to search.

modifier
specifies a character constant, variable, or expression in which each non-blank character modifies the action of the FINDW function.

Tip: If you use the *modifier* argument, then it must be positioned after the *chars* argument.

You can use the following characters as modifiers:

blank	is ignored.
a or A	adds alphabetic characters to the list of characters.
b or B	scans from right to left instead of from left to right, regardless of the sign of the *startpos* argument.
c or C	adds control characters to the list of characters.
d or D	adds digits to the list of characters.
e or E	counts the words that are scanned until the specified word is found, instead of determining the character position of the specified word in the string. Fragments of a word are not counted.
f or F	adds an underscore and English letters (that is, the characters that can begin a SAS variable name using VALIDVARNAME=V7) to the list of characters.
g or G	adds graphic characters to the list of characters.
h or H	adds a horizontal tab to the list of characters.
i or I	ignores the case of the characters.
k or K	causes all characters that are not in the list of characters to be treated as delimiters. If K is not specified, then all characters that are in the list of characters are treated as delimiters.
l or L	adds lowercase letters to the list of characters.
m or M	specifies that multiple consecutive delimiters, and delimiters at the beginning or end of the *string* argument, refer to words that have a length of zero.
n or N	adds digits, an underscore, and English letters (that is, the characters that can appear after the first character in a SAS variable name using VALIDVARNAME=V7) to the list of characters.

o or O	processes the *chars* and *modifier* arguments only once, rather than every time the FINDW function is called. Using the O modifier in the DATA step (excluding WHERE clauses), or in the SQL procedure, can make FINDW run faster when you call it in a loop where the *chars* and *modifier* arguments do not change.
p or P	adds punctuation marks to the list of characters.
q or Q	ignores delimiters that are inside of substrings that are enclosed in quotation marks. If the value of the *string* argument contains unmatched quotation marks, then scanning from left to right will produce different words than scanning from right to left.
r or R	removes leading and trailing delimiters from the *word* argument.
s or S	adds space characters (blank, horizontal tab, vertical tab, carriage return, line feed, and form feed) to the list of characters.
t or T	trims trailing blanks from the *string*, *word*, and *chars* arguments.
u or U	adds uppercase letters to the list of characters.
w or W	adds printable characters to the list of characters.
x or X	adds hexadecimal characters to the list of characters.

Details

Definition of "Delimiter" "Delimiter" refers to any of several characters that are used to separate words. You can specify the delimiters by using the *chars* argument, the *modifier* argument, or both. If you specify the Q modifier, then the characters inside of substrings that are enclosed in quotation marks are not treated as delimiters.

Definition of "Word" "Word" refers to a substring that has both of the following characteristics:

- bounded on the left by a delimiter or the beginning of the string
- bounded on the right by a delimiter or the end of the string

Note: A word can contain delimiters. In this case, the FINDW function differs from the SCAN function, in which words are defined as not containing delimiters. △

Searching for a String If the FINDW function fails to find a substring that both matches the specified word and satisfies the definition of a word, then FINDW returns a value of 0.

If the FINDW function finds a substring that both matches the specified word and satisfies the definition of a word, the value that is returned by FINDW depends on whether the E modifier is specified:

- If you specify the E modifier, then FINDW returns the number of complete words that were scanned while searching for the specified word. If *startpos* specifies a position in the middle of a word, then that word is not counted.
- If you do not specify the E modifier, then FINDW returns the character position of the substring that is found.

If you specify the *startpos* argument, then the absolute value of *startpos* specifies the position at which to begin the search. The sign of *startpos* specifies the direction in which to search:

Value of *startpos*	Action
greater than 0	search begins at position *startpos* and proceeds to the right. If *startpos* is greater than the length of the string, then FINDW returns a value of 0.
less than 0	search begins at position –*startpos* and proceeds to the left. If *startpos* is less than the negative of the length of the string, then the search begins at the end of the string.
equal to 0	FINDW returns a value of 0.

If you do not specify the *startpos* argument or the B modifier, then FINDW searches from left to right starting at the beginning of the string. If you specify the B modifier, but do not use the *startpos* argument, then FINDW searches from right to left starting at the end of the string.

Using the FINDW Function in ASCII and EBCDIC Environments If you use the FINDW function with only two arguments, the default delimiters depend on whether your computer uses ASCII or EBCDIC characters.

- If your computer uses ASCII characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ^ |

  In ASCII environments that do not contain the ^ character, the FINDW function uses the ~ character instead.
- If your computer uses EBCDIC characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ¬ | ¢

Using Null Arguments The FINDW function allows character arguments to be null. Null arguments are treated as character strings with a length of zero. Numeric arguments cannot be null.

Examples

Example 1: Searching a Character String for a Word The following example searches a character string for the word “she”, and returns the position of the beginning of the word.

```
data _null_;
   whereisshe=findw('She sells sea shells? Yes, she does.','she');
   put whereisshe=;
run;
```

SAS writes the following output to the log:

```
whereisshe=28
```

Example 2: Searching a Character String and Using the *Chars* and *Startpos* Arguments The following example contains two occurrences of the word "rain." Only the second occurrence is found by FINDW because the search begins in position 25. The *chars* argument specifies a space as the delimiter.

```
data _null_;
   result = findw('At least 2.5 meters of rain falls in a rain forest.',
                  'rain', ' ', 25);
   put result=;
run;
```

SAS writes the following output to the log:

```
result=40
```

Example 3: Searching a Character String and Using the I Modifier and the *Startpos* Argument The following example uses the I modifier and returns the position of the beginning of the word. The I modifier disregards case, and the *startpos* argument identifies the starting position from which to search.

```
data _null_;
   string='Artists from around the country display their art at
           an art festival.';
   result=findw(string, 'Art',' ', 'i', 10);
   put result=;
run;
```

SAS writes the following output to the log:

```
result=47
```

Example 4: Searching a Character String and Using the E Modifier The following example uses the E modifier and returns the number of complete words that are scanned while searching for the word "art."

```
data _null_;
   string='Artists from around the country display their art at
           an art festival.';
   result=findw(string,'art',' ','E');
   put result=;
run;
```

SAS writes the following output to the log:

```
result=8
```

Example 5: Searching a Character String and Using the E Modifier and the *Startpos* Argument The following example uses the E modifier to count words in a character string. The word count begins at position 50 in the string. The result is 3 because "art" is the third word after the 50th character position.

```
data _null_;
   string='Artists from around the country display their art at
           an art festival.';
   result=findw(string, 'art',' ','E',50);
   put result=;
run;
```

SAS writes the following output to the log:

```
result=3
```

Example 6: Searching a Character String and Using Two Modifiers The following example uses the I and the E modifiers to find a word in a string.

```
data _null_;
   string='The Great Himalayan National Park was created in 1984. Because
           of its terrain and altitude, the park supports a diversity
           of wildlife and vegetation.';
   result=findw(string,'park',' ','I E');
   put result=;
run;
```

SAS writes the following output to the log:

```
result=5
```

Example 7: Searching a Character String and Using the R Modifier The following example uses the R modifier to remove leading and trailing delimiters from a word.

```
data _null_;
    string='Artists from around the country display their art at
               an art festival.';
    word='  art  ';
    result=findw(string, word, ' ', 'R');
    put result=;
run;
```

SAS writes the following output to the log:

```
result=47
```

See Also

Functions and CALL Routines:

"CALL SCAN Routine" on page 516
"COUNTW Function" on page 621
"FIND Function" on page 735
"FINDC Function" on page 737
"INDEXW Function" on page 820
"SCAN Function" on page 1111

FINFO Function

Returns the value of a file information item.

Category: External Files

See: FINFO Function in the documentation for your operating environment.

Syntax

FINFO(*file-id*,*info-item*)

Arguments

file-id
is a numeric constant, variable, or expression that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

info-item
is a character constant, variable, or expression that specifies the name of the file information item to be retrieved.

Details

FINFO returns the value of a system-dependent information item for an external file. FINFO returns a blank if the value given for *info-item* is invalid.

Operating Environment Information: The information available on files depends on the operating environment. △

Comparisons

- The FOPTNAME function determines the names of the available file information items.
- The FOPTNUM function determines the number of system-dependent information items that are available.

Examples

This example stores information items about an external file in a SAS data set:

```
data info;
   length infoname infoval $60;
   drop rc fid infonum i close;
   rc=filename('abc','physical-filename');
   fid=fopen('abc');
   infonum=foptnum(fid);
   do i=1 to infonum;
      infoname=foptname(fid,i);
      infoval=finfo(fid,infoname);
      output;
   end;
   close=fclose(fid);
run;
```

See Also

Functions:

"FCLOSE Function" on page 680
"FOPTNUM Function" on page 766
"MOPEN Function" on page 936

FINV Function

Returns a quantile from the *F* distribution.

Category: Quantile

Syntax

FINV (*p*, *ndf*, *ddf* <,*nc*>)

Arguments

p
is a numeric probability.
Range: $0 \leq p$

ndf
is a numeric numerator degrees of freedom parameter.

Range: $ndf > 0$

ddf
is a numeric denominator degrees of freedom parameter.

Range: $ddf > 0$

nc
is an optional numeric noncentrality parameter.

Range: $nc \geq 0$

Details

The FINV function returns the p^{th} quantile from the *F* distribution with numerator degrees of freedom *ndf*, denominator degrees of freedom *ddf*, and noncentrality parameter *nc*. The probability that an observation from the *F* distribution is less than the quantile is *p*. This function accepts noninteger degrees of freedom parameters *ndf* and *ddf*.

If the optional parameter *nc* is not specified or has the value 0, the quantile from the central *F* distribution is returned. The noncentrality parameter *nc* is defined such that if X and Y are normal random variables with means μ and 0, respectively, and variance 1, then X^2/Y^2 has a noncentral *F* distribution with $nc = \mu^2$.

CAUTION:
For large values of *nc,* the algorithm could fail. In that case, a missing value is returned. △

Note: FINV is the inverse of the PROBF function. △

Examples

These statements compute the 95^{th} quantile value of a central *F* distribution with 2 and 10 degrees of freedom and a noncentral *F* distribution with 2 and 10.3 degrees of freedom and a noncentrality parameter equal to 2:

SAS Statements	Results
`q1=finv(.95,2,10);`	`4.1028210151`
`q2=finv(.95,2,10.3,2);`	`7.583766024`

See Also

Functions:

"QUANTILE Function" on page 1064

FIPNAME Function

Converts two-digit FIPS codes to uppercase state names.

Category: State and ZIP Code

Syntax

FIPNAME(*expression*)

Arguments

expression
specifies a numeric constant, variable, or expression that represents a U.S. FIPS code.

Details

If the FIPNAME function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

The FIPNAME function converts a U.S. Federal Information Processing Standards (FIPS) code to the corresponding state or U.S. territory name in uppercase, returning a value of up to 20 characters.

Comparisons

The FIPNAME, FIPNAMEL, and FIPSTATE functions take the same argument but return different values. FIPNAME returns uppercase state names. FIPNAMEL returns mixed case state names. FIPSTATE returns a two-character state postal code (or world-wide GSA geographic code for U.S. territories) in uppercase.

Examples

The examples show the differences when using FIPNAME, FIPNAMEL, and FIPSTATE.

SAS Statements	Results
`x=fipname(37);` `put x;`	`NORTH CAROLINA`
`x=fipnamel(37);` `put x;`	`North Carolina`
`x=fipstate(37);` `put x;`	`NC`

See Also

Functions:

"FIPNAMEL Function" on page 753
"FIPSTATE Function" on page 754
"STFIPS Function" on page 1134
"STNAME Function" on page 1136
"STNAMEL Function" on page 1137

FIPNAMEL Function

Converts two-digit FIPS codes to mixed case state names.

Category: State and ZIP Code

Syntax

FIPNAMEL(*expression*)

Arguments

expression
specifies a numeric constant, variable, or expression that represents a U.S. FIPS code.

Details

If the FIPNAMEL function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

The FIPNAMEL function converts a U.S. Federal Information Processing Standards (FIPS) code to the corresponding state or U.S. territory name in mixed case, returning a value of up to 20 characters.

Comparisons

The FIPNAME, FIPNAMEL, and FIPSTATE functions take the same argument but return different values. FIPNAME returns uppercase state names. FIPNAMEL returns mixed case state names. FIPSTATE returns a two-character state postal code (or world-wide GSA geographic code for U.S. territories) in uppercase.

Examples

The examples show the differences when using FIPNAME, FIPNAMEL, and FIPSTATE.

SAS Statements	Results
`x=fipname(37);` `put x;`	`NORTH CAROLINA`
`x=fipnamel(37);` `put x;`	`North Carolina`
`x=fipstate(37);` `put x;`	`NC`

See Also

Functions:

"FIPNAME Function" on page 752
"FIPSTATE Function" on page 754
"STFIPS Function" on page 1134
"STNAME Function" on page 1136
"STNAMEL Function" on page 1137

FIPSTATE Function

Converts two-digit FIPS codes to two-character state postal codes.

Category: State and ZIP Code

Syntax

FIPSTATE(*expression*)

Arguments

expression
specifies a numeric constant, variable, or expression that represents a U.S. FIPS code.

Details

If the FIPSTATE function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

The FIPSTATE function converts a U.S. Federal Information Processing Standards (FIPS) code to a two-character state postal code (or world-wide GSA geographic code for U.S. territories) in uppercase.

Comparisons

The FIPNAME, FIPNAMEL, and FIPSTATE functions take the same argument but return different values. FIPNAME returns uppercase state names. FIPNAMEL returns mixed case state names. FIPSTATE returns a two-character state postal code (or world-wide GSA geographic code for U.S. territories) in uppercase.

Examples

The examples show the differences when using FIPNAME, FIPNAMEL, and FIPSTATE.

SAS Statements	Results
`x=fipname(37);` `put x;`	`NORTH CAROLINA`
`x=fipnamel(37);` `put x;`	`North Carolina`
`x=fipstate(37);` `put x;`	`NC`

See Also

Functions:

"FIPNAME Function" on page 752
"FIPNAMEL Function" on page 753
"STFIPS Function" on page 1134
"STNAME Function" on page 1136
"STNAMEL Function" on page 1137

FIRST Function

Returns the first character in a character string.

Category: Character

Syntax

FIRST(*string*)

Arguments

string
specifies a character string.

Details

In a DATA step, the default length of the target variable for the FIRST function is 1.

The FIRST function returns a string with a length of 1. If *string* has a length of 0, then the FIRST function returns a single blank.

Comparisons

The FIRST function returns the same result as CHAR(*string*, 1) and SUBPAD(*string*, 1, 1). While the results are the same, the default length of the target variable is different.

Examples

The following example shows the results of using the FIRST function.

```
options pageno=1 ps=64 ls=80 nodate;

data test;
   string1="abc";
   result1=first(string1);

   string2="";
   result2=first(string2);
run;

proc print noobs data=test;
run;
```

Output 4.46 Output from the FIRST Function

```
                              The SAS System                                1

                string1    result1    string2    result2

                  abc          a
```

See Also

Functions:

"CHAR Function" on page 578

FLOOR Function

Returns the largest integer that is less than or equal to the argument, fuzzed to avoid unexpected floating-point results.

Category: Truncation

Syntax

FLOOR (*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

If the argument is within 1E-12 of an integer, the function returns that integer.

Comparisons

Unlike the FLOORZ function, the FLOOR function fuzzes the result. If the argument is within 1E-12 of an integer, the FLOOR function fuzzes the result to be equal to that integer. The FLOORZ function does not fuzz the result. Therefore, with the FLOORZ function you might get unexpected results.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`var1=2.1;`
`a=floor(var1);`
`put a;`
	`2`

`var2=-2.4;`
`b=floor(var2);`
`put b;`
	`-3`

`c=floor(-1.6);` `put c;`	`-2`
`d=floor(1.-1.e-13);` `put d;`	`1`
`e=floor(763);` `put e;`	`763`
`f=floor(-223.456);` `put f;`	`-224`

See Also

Functions:

"FLOORZ Function" on page 758

FLOORZ Function

Returns the largest integer that is less than or equal to the argument, using zero fuzzing.

Category: Truncation

Syntax

FLOORZ (*argument*)

Arguments

argument
is a numeric constant, variable, or expression.

Comparisons

Unlike the FLOOR function, the FLOORZ function uses zero fuzzing. If the argument is within 1E-12 of an integer, the FLOOR function fuzzes the result to be equal to that integer. The FLOORZ function does not fuzz the result. Therefore, with the FLOORZ function you might get unexpected results.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`var1=2.1;`
`a=floorz(var1);`
`put a;`
	`2`

`var2=-2.4;`
`b=floorz(var2);`
`put b;`
	`-3`

`c=floorz(-1.6);` `put c;`	`-2`

`var6=(1.-1.e-13);`
`d=floorz(1-1.e-13);`
`put d;`
	`0`

SAS Statements	Results
`e=floorz(763);` `put e;`	`763`
`f=floorz(-223.456);` `put f;`	`-224`

See Also

Functions:

"FLOOR Function" on page 757

FNONCT Function

Returns the value of the noncentrality parameter of an F distribution.

Category: Mathematical

Syntax

FNONCT(*x*,*ndf*,*ddf*,*prob*)

Arguments

x
is a numeric random variable.
Range: $x \geq 0$

ndf
is a numeric numerator degree of freedom parameter.
Range: $ndf > 0$

ddf
is a numeric denominator degree of freedom parameter.
Range: $ddf > 0$

prob
is a probability.
Range: $0 < prob < 1$

Details

The FNONCT function returns the nonnegative noncentrality parameter from a noncentral *F* distribution whose parameters are *x*, *ndf*, *ddf*, and *nc*. If *prob* is greater than the probability from the central *F* distribution whose parameters are *x*, *ndf*, and *ddf*, a root to this problem does not exist. In this case a missing value is returned. A Newton-type algorithm is used to find a nonnegative root *nc* of the equation

$$P_f\left(x|ndf, ddf, nc\right) - prob = 0$$

where

$$P_f\left(x|ndf, ddf, nc\right) = e^{\frac{-nc}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{nc}{2}\right)^j}{j!} I_{\frac{(ndf)x}{ddf+(ndf)x}} \left(\frac{ddf}{2} + j, \frac{ddf}{2}\right)$$

where $I\left(. \ . \ .\right)$ is the probability from the beta distribution that is given by

$$I_x\left(a, b\right) = \frac{}{\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}} \int_0^x t^{a-1}\left(1-t\right)^{b-1} dt$$

If the algorithm fails to converge to a fixed point, a missing value is returned.

Examples

```
data work;
   x=2;
   df=4;
   ddf=5;
   do nc=1 to 3 by .5;
      prob=probf(x,df,ddf,nc);
      ncc=fnonct(x,df,ddf,prob);
      output;
   end;
run;
proc print;
run;
```

Output 4.47 FNONCT Example Output

OBS	x	df	ddf	nc	prob	ncc
1	2	4	5	1.0	0.69277	1.0
2	2	4	5	1.5	0.65701	1.5
3	2	4	5	2.0	0.62232	2.0
4	2	4	5	2.5	0.58878	2.5
5	2	4	5	3.0	0.55642	3.0

FNOTE Function

Identifies the last record that was read, and returns a value that the FPOINT function can use.

Category: External Files

Syntax

FNOTE(*file-id*)

Argument

file-id
is a numeric constant, variable, or expression that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

Details

You can use FNOTE like a bookmark, marking the position in the file so that your application can later return to that position using FPOINT. The value that is returned by FNOTE is required by the FPOINT function to reposition the file pointer on a specific record.

To free the memory associated with each FNOTE identifier, use DROPNOTE.

Note: You cannot write a new record in place of the current record if the new record has a length that is greater than the current record. △

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file. If the file is opened successfully, indicated by a positive value in the variable FID, then it reads the records, stores in the variable NOTE 3 the position of the third record read, and then later uses FPOINT to point back to NOTE3 to update the file. After updating the record, it closes the file:

```
%let
fref=MYFILE;
%let rc=%sysfunc(filename(fref,
   physical-filename));
%let fid=%sysfunc(fopen(&fref,u));
%if &fid > 0 %then
   %do;
      %let rc=%sysfunc(fread(&fid));
         /* Read second record. */
      %let rc=%sysfunc(fread(&fid));
         /* Read third record. */
      %let rc=%sysfunc(fread(&fid));
         /* Note position of third record. */
      %let note3=%sysfunc(fnote(&fid));
         /* Read fourth record. */
      %let rc=%sysfunc(fread(&fid));
         /* Read fifth record. */
      %let rc=%sysfunc(fread(&fid));
         /* Point to third record. */
      %let rc=%sysfunc(fpoint(&fid,&note3));
         /* Read third record. */
      %let rc=%sysfunc(fread(&fid));
         /* Copy new text to FDB. */
      %let rc=%sysfunc(fput(&fid,New text));
         /* Update third record  */
```

```
         /* with data in FDB. */
       %let rc=%sysfunc(fwrite(&fid));
         /* Close file. */
       %let rc=%sysfunc(fclose(&fid));
    %end;
%let rc=%sysfunc(filename(fref));
```

See Also

Functions:

"DROPNOTE Function" on page 667
"FCLOSE Function" on page 680
"FILENAME Function" on page 690
"FOPEN Function" on page 762
"FPOINT Function" on page 767
"FPUT Function" on page 771
"FREAD Function" on page 772
"FREWIND Function" on page 773
"FWRITE Function" on page 778
"MOPEN Function" on page 936

FOPEN Function

Opens an external file and returns a file identifier value.

Category: External Files

See: FOPEN Function in the documentation for your operating environment.

Syntax

FOPEN(*fileref*<,*open-mode*<,*record-length*<,*record-format*>>>)

Arguments

fileref
is a character constant, variable, or expression that specifies the fileref assigned to the external file.

Tip: If *fileref* is longer than eight characters, then it will be truncated to eight characters.

open-mode
is a character constant, variable, or expression that specifies the type of access to the file:

A	APPEND mode allows writing new records after the current end of the file.

I	INPUT mode allows reading only (default).
O	OUTPUT mode defaults to the OPEN mode specified in the operating environment option in the FILENAME statement or function. If no operating environment option is specified, it allows writing new records at the beginning of the file.
S	Sequential input mode is used for pipes and other sequential devices such as hardware ports.
U	UPDATE mode allows both reading and writing.

Default: I

record-length
is a numeric constant, variable, or expression that specifies the logical record length of the file. To use the existing record length for the file, specify a length of 0, or do not provide a value here.

record-format
is a character constant, variable, or expression that specifies the record format of the file. To use the existing record format, do not specify a value here. Valid values are:

B	data are to be interpreted as binary data.
D	use default record format.
E	use editable record format.
F	file contains fixed length records.
P	file contains printer carriage control in operating environment-dependent record format. *Note:* For z/OS data sets with FBA or VBA record format, specify 'P' for the *record-format* argument.
V	file contains variable length records.

Note: If an argument is invalid, FOPEN returns 0, and you can obtain the text of the corresponding error message from the SYSMSG function. Invalid arguments do not produce a message in the SAS log and do not set the _ERROR_ automatic variable. △

Details

CAUTION:
Use OUTPUT mode with care. Opening an existing file for output overwrites the current contents of the file without warning. △

The FOPEN function opens an external file for reading or updating and returns a file identifier value that is used to identify the open file to other functions. You must associate a fileref with the external file before calling the FOPEN function. FOPEN returns a 0 if the file could not be opened. You can assign filerefs by using the FILENAME statement or the FILENAME external file access function. Under some operating environments, you can also assign filerefs by using system commands.

If you call the FOPEN function from a macro, then the result of the call is valid only when it is passed to functions in a macro. If you call the FOPEN function from the DATA step, then the result is valid only when it is passed to functions in the same DATA step.

Operating Environment Information: It is good practice to use the FCLOSE function at the end of a DATA step if you used FOPEN to open the file, even though using FCLOSE might not be required in your operating environment. For more information about FOPEN, see the SAS documentation for your operating environment. △

Examples

Example 1: Opening a File Using Defaults This example assigns the fileref MYFILE to an external file and attempts to open the file for input using all defaults. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf));
```

Example 2: Opening a File without Using Defaults This example attempts to open a file for input without using defaults. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let fid=%sysfunc(fopen(file2,o,132,e));
```

Example 3: Handling Errors This example shows how to check for errors and write an error message from the SYSMSG function.

```
data _null_;
   f=fopen('bad','?');
   if not f then do;
      m=sysmsg();
      put m;
      abort;
      end;
      ... more SAS statements ...
   run;
```

See Also

Functions:

"DOPEN Function" on page 661
"FCLOSE Function" on page 680
"FILENAME Function" on page 690
"FILEREF Function" on page 692
"MOPEN Function" on page 936
"SYSMSG Function" on page 1154

Statement:

"FILENAME Statement" on page 1520

FOPTNAME Function

Returns the name of an item of information about a file.

Category: External Files

See: FOPTNAME Function in the documentation for your operating environment.

Syntax

FOPTNAME(*file-id*,*nval*)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

nval
is a numeric constant, variable, or expression that specifies the number of the information item.

Details

FOPTNAME returns a blank if an error occurred.

Operating Environment Information: The number, value, and type of information items that are available depend on the operating environment. △

Examples

Example 1: Retrieving File Information Items and Writing Them to the Log This example retrieves the system-dependent file information items that are available and writes them to the log:

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf));
%let infonum=%sysfunc(foptnum(&fid));
%do j=1 %to &infonum;
   %let name=%sysfunc(foptname(&fid,&j));
   %let value=%sysfunc(finfo(&fid,&name));
   %put File attribute &name equals &value;
%end;
%let rc=%sysfunc(fclose(&fid));
%let rc=%sysfunc(filename(filrf));
```

Example 2: Creating a Data Set with Names and Values of File Attributes This example creates a data set that contains the name and value of the available file attributes:

```
data fileatt;
   length name $ 20 value $ 40;
   drop rc fid j infonum;
   rc=filename("myfile","physical-filename");
   fid=fopen("myfile");
   infonum=foptnum(fid);
   do j=1 to infonum;
      name=foptname(fid,j);
      value=finfo(fid,name);
      put 'File attribute ' name
       'has a value of ' value;
      output;
```

```
   end;
   rc=filename("myfile");
run;
```

See Also

Functions:

"DINFO Function" on page 657
"DOPTNAME Function" on page 662
"DOPTNUM Function" on page 664
"FCLOSE Function" on page 680
"FILENAME Function" on page 690
"FINFO Function" on page 749
"FOPEN Function" on page 762
"FOPTNUM Function" on page 766
"MOPEN Function" on page 936

FOPTNUM Function

Returns the number of information items that are available for an external file.

Category: External Files

See: FOPTNUM Function in the documentation for your operating environment.

Syntax

FOPTNUM(*file-id*)

Argument

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

Details

Operating Environment Information: The number, value, and type of information items that are available depend on the operating environment. △

Comparisons

- Use FOPTNAME to determine the names of the items that are available for a particular operating environment.
- Use FINFO to retrieve the value of a particular information item.

Examples

This example opens the external file with the fileref MYFILE and determines the number of system-dependent file information items available:

```
%let fid=%sysfunc(fopen(myfile));
%let infonum=%sysfunc(foptnum(&fid));
```

See Also

Functions:

"DINFO Function" on page 657
"DOPTNAME Function" on page 662
"DOPTNUM Function" on page 664
"FINFO Function" on page 749
"FOPEN Function" on page 762
"FOPTNAME Function" on page 764
"MOPEN Function" on page 936

See the "Examples" on page 765 in the FOPTNAME Function.

FPOINT Function

Positions the read pointer on the next record to be read.

Category: External Files

Syntax

FPOINT(*file-id*,*note-id*)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

note-id
specifies the identifier that was assigned by the FNOTE function.

Details

FPOINT returns 0 if the operation was successful, or ≠0 if it was not successful. FPOINT determines only the record to read next. It has no impact on which record is written next. When you open the file for update, FWRITE writes to the most recently read record.

Note: You cannot write a new record in place of the current record if the new record has a length that is greater than the current record. △

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file. If the file is opened successfully, it reads the records and uses NOTE3 to store the position of the third record read. Later, it points back to NOTE3 to update the file, and closes the file afterward:

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf,u));
%if &fid > 0 %then
   %do;
         /* Read first record. */
      %let rc=%sysfunc(fread(&fid));
         /* Read second record. */
      %let rc=%sysfunc(fread(&fid));
         /* Read third record. */
      %let rc=%sysfunc(fread(&fid));
         /* Note position of third record. */
      %let note3=%sysfunc(fnote(&fid));
         /* Read fourth record. */
      %let rc=%sysfunc(fread(&fid));
         /* Read fifth record. */
      %let rc=%sysfunc(fread(&fid));
         /* Point to third record. */
      %let rc=%sysfunc(fpoint(&fid,&note3));
         /* Read third record. */
      %let rc=%sysfunc(fread(&fid));
         /* Copy new text to FDB. */
      %let rc=%sysfunc(fput(&fid,New text));
         /* Update third record  */
         /* with data in FDB. */
      %let rc=%sysfunc(fwrite(&fid));
         /* Close file. */
      %let rc=%sysfunc(fclose(&fid));
   %end;
%let rc=%sysfunc(filename(filrf));
```

See Also

Functions:

“DROPNOTE Function” on page 667
“FCLOSE Function” on page 680
“FILENAME Function” on page 690
“FNOTE Function” on page 760
“FOPEN Function” on page 762
“FPUT Function” on page 771
“FREAD Function” on page 772
“FREWIND Function” on page 773
“FWRITE Function” on page 778
“MOPEN Function” on page 936

FPOS Function

Sets the position of the column pointer in the File Data Buffer (FDB).

Category: External Files

Syntax

FPOS(*file-id*,*nval*)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

nval
is a numeric constant, variable, or expression that specifies the column at which to set the pointer.

Details

FPOS returns 0 if the operation was successful, ≠0 if it was not successful. If you open a file in output mode and the specified position is past the end of the current record, the size of the record is increased appropriately. However, in a fixed block or VBA file, if you specify a column position beyond the end of the record, the record size does not change and the text string is not written to the file.

If you open a file in update mode and the specified position is not past the end of the current record, then SAS writes the record to the file. If the specified position is past the end of the current record, then SAS returns an error message and does not write the new record:

```
ERROR: Cannot increase record length in update mode.
```

Note: If you use the update mode with the FOPEN function, then you must execute FREAD before you execute FWRITE functions. △

Examples

This example assigns the fileref MYFILE to an external file and opens the file in update mode. If the file is opened successfully, indicated by a positive value in the variable FID, SAS reads a record and places data into the file's buffer at column 12. If the resulting record length is less than or equal to the original record length, then SAS writes the record and closes the file. If the resulting record length is greater than the original record length, then SAS writes an error message to the log.

```
%macro ptest;
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,test.txt));
%let fid=%sysfunc(fopen(&filrf,o));
```

```
%let rc=%sysfunc(fread(&fid));
%put &fid;
%if (&fid > 0) %then
   %do;
      %let dataline=This is some data.;
         /* Position at column 12 in the FDB. */
      %let rc=%sysfunc(fpos(&fid,12));
      %put &rc one;
         /* Put the data in the FDB. */
      %let rc=%sysfunc(fput(&fid,&dataline));
      %put &rc two;
      %if (&rc ne 0) %then
         %do;
            %put %sysfunc(sysmsg());
         %end;
      %else %do;
         /* Write the record. */
      %let rc=%sysfunc(fwrite(&fid));
      %if (&rc ne 0) %then
         %do;
            %put write fails &rc;
         %end;
      %end;
         /* Close the file. */
      %let rc=%sysfunc(fclose(&fid));
   %end;
%let rc=%sysfunc(filename(filrf));
%mend;
%ptest;
```

Output 4.48 Output from the FPOS Function

```
1
0 one
0 two
```

See Also

Functions:

“FCLOSE Function” on page 680
“FCOL Function” on page 681
“FILENAME Function” on page 690
“FOPEN Function” on page 762
“FPUT Function” on page 771
“FWRITE Function” on page 778
“MOPEN Function” on page 936

FPUT Function

Moves data to the File Data Buffer (FDB) of an external file, starting at the FDB's current column position.

Category: External Files

Syntax

FPUT(*file-id*,*cval*)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

cval
is a character constant, variable, or expression that specifies the file data.

Details

FPUT returns 0 if the operation was successful, ≠0 if it was not successful. The number of bytes moved to the FDB is determined by the length of the variable. The value of the column pointer is then increased to one position past the end of the new text.

Note: You cannot write a new record in place of the current record if the new record has a length that is greater than the current record. △

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file in APPEND mode. If the file is opened successfully, indicated by a positive value in the variable FID, it moves data to the FDB using FPUT, appends a record using FWRITE, and then closes the file. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%macro ptest;
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,test.txt));
%let fid=%sysfunc(fopen(&filrf,a));
%if &fid > 0 %then
   %do;
      %let rc=%sysfunc(fread(&fid));
      %let mystring=This is some data.;
      %let rc=%sysfunc(fput(&fid,&mystring));
      %let rc=%sysfunc(fwrite(&fid));
      %let rc=%sysfunc(fclose(&fid));
   %end;
```

```
%else
   %put %sysfunc(sysmsg());
%let rc=%sysfunc(filename(filrf));
%put return code = &rc;
%mend;
%ptest;
```

SAS writes the following output to the log:

```
return code = 0
```

See Also

Functions:

"FCLOSE Function" on page 680
"FILENAME Function" on page 690
"FNOTE Function" on page 760
"FOPEN Function" on page 762
"FPOINT Function" on page 767
"FPOS Function" on page 769
"FWRITE Function" on page 778
"MOPEN Function" on page 936
"SYSMSG Function" on page 1154

FREAD Function

Reads a record from an external file into the File Data Buffer (FDB).

Category: External Files

Syntax

FREAD(*file-id*)

Argument

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

Details

FREAD returns 0 if the operation was successful, ≠0 if it was not successful. The position of the file pointer is updated automatically after the read operation so that successive FREAD functions read successive file records.

To position the file pointer explicitly, use FNOTE, FPOINT, and FREWIND.

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file. If the file opens successfully, it lists all of the file's records in the log:

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf));
%if &fid > 0 %then
   %do %while(%sysfunc(fread(&fid)) = 0);
      %let rc=%sysfunc(fget(&fid,c,200));
           %put &c;
      %end;
   %let rc=%sysfunc(fclose(&fid));
%let rc=%sysfunc(filename(filrf));
```

See Also

Functions:

"FCLOSE Function" on page 680
"FGET Function" on page 687
"FILENAME Function" on page 690
"FNOTE Function" on page 760
"FOPEN Function" on page 762
"FREWIND Function" on page 773
"FREWIND Function" on page 773
"MOPEN Function" on page 936

FREWIND Function

Positions the file pointer to the start of the file.

Category: External Files

Syntax

FREWIND(*file-id*)

Argument

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

Details

FREWIND returns 0 if the operation was successful, ≠0 if it was not successful.
FREWIND has no effect on a file opened with sequential access.

Examples

This example assigns the fileref MYFILE to an external file. Then it opens the file and reads the records until the end of the file is reached. The FREWIND function then repositions the pointer to the beginning of the file. The first record is read again and stored in the File Data Buffer (FDB). The first token is retrieved and stored in the macro variable VAL:

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
   physical-filename));
%let fid=%sysfunc(fopen(&filrf));
%let rc=0;
%do %while (&rc ne -1);
   /* Read a record. */
   %let rc=%sysfunc(fread(&fid));
%end;
   /* Reposition pointer to beginning of file. */
%if &rc = -1 %then
  %do;
    %let rc=%sysfunc(frewind(&fid));
       /* Read first record. */
    %let rc=%sysfunc(fread(&fid));
       /* Read first token  */
       /* into macro variable VAL. */
    %let rc=%sysfunc(fget(&fid,val));
    %put val=&val;
  %end;
%else
  %put Error on fread=%sysfunc(sysmsg());
%let rc=%sysfunc(fclose(&fid));
%let rc=%sysfunc(filename(filrf));
```

See Also

Functions:

"FCLOSE Function" on page 680
"FGET Function" on page 687
"FILENAME Function" on page 690
"FOPEN Function" on page 762
"FREAD Function" on page 772
"MOPEN Function" on page 936
"SYSMSG Function" on page 1154

FRLEN Function

Returns the size of the last record that was read, or, if the file is opened for output, returns the current record size.

Category: External Files

Syntax

FRLEN(*file-id*)

Argument

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

Examples

This example opens the file that is identified by the fileref MYFILE. It determines the minimum and maximum length of records in the external file and writes the results to the log:

```
%let fid=%sysfunc(fopen(myfile));
%let min=0;
%let max=0;
%if (%sysfunc(fread(&fid)) = 0) %then
   %do;
      %let min=%sysfunc(frlen(&fid));
      %let max=&min;
      %do %while(%sysfunc(fread(&fid)) = 0);
         %let reclen=%sysfunc(frlen(&fid));
         %if (&reclen > &max) %then
            %let max=&reclen;
         %if (&reclen < &min) %then
            %let min=&reclen;
      %end;
%end;
%let rc=%sysfunc(fclose(&fid));
%put max=&max min=&min;
```

See Also

Functions:

"FCLOSE Function" on page 680
"FOPEN Function" on page 762
"FREAD Function" on page 772
"MOPEN Function" on page 936

FSEP Function

Sets the token delimiters for the FGET function.

Category: External Files

Syntax

FSEP(*file-id*,*characters*<,'x' |'X'>)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

character
is a character constant, variable, or expression that specifies one or more delimiters that separate items in the File Data Buffer (FDB). Each character listed is a delimiter. That is, if *character* is #@, either # or @ can separate items. Multiple consecutive delimiters, such as @#@, are treated as a single delimiter.

Default: blank

'x' | 'X'
specifies that the character delimiter is a hexadecimal value.

Restriction: 'x' and 'X' are the only valid values for this argument. All other values will cause an error to occur.

Restriction: If you pass 'x' or 'X' as the third argument, a valid hexadecimal string must be passed as the second argument, *character*. Otherwise, the function will fail. A valid hexadecimal string is an even number of 0–9 and A–F characters.

Tip: If you use a macro statement, then quotation marks enclosing x or X are not required.

Details

FSEP returns 0 if the operation was successful, ≠0 if it was not successful.

Examples

An external file has data in this form:

```
John J. Doe,Male,25,Weight Lifter
Pat O'Neal,Female,22,Gymnast
```

Note that each field is separated by a comma.

This example reads the file that is identified by the fileref MYFILE, using the comma as a separator, and writes the values for NAME, GENDER, AGE, and WORK to the SAS log. Note that in a macro statement you do not enclose character strings in quotation marks, but a literal comma in a function argument must be enclosed in a macro quoting function such as %STR.

```
%let fid=%sysfunc(fopen(myfile));
%let rc=%sysfunc(fsep(&fid,%str(,)));
%do %while(%sysfunc(fread(&fid)) = 0);
   %let rc=%sysfunc(fget(&fid,name));
   %let rc=%sysfunc(fget(&fid,gender));
```

```
   %let rc=%sysfunc(fget(&fid,age));
   %let rc=%sysfunc(fget(&fid,work));
   %put name=%bquote(&name) gender=&gender
      age=&age work=&work;
%end;
%let rc=%sysfunc(fclose(&fid));
```

See Also

Functions:

"FCLOSE Function" on page 680
"FGET Function" on page 687
"FOPEN Function" on page 762
"FREAD Function" on page 772
"MOPEN Function" on page 936

FUZZ Function

Returns the nearest integer if the argument is within 1E–12 of that integer.

Category: Truncation

Syntax

FUZZ(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The FUZZ function returns the nearest integer value if the argument is within 1E–12 of the integer (that is, if the absolute difference between the integer and argument is less than 1E–12). Otherwise, the argument is returned.

Examples

SAS Statements	Results

`var1=5.9999999999999;`
`x=fuzz(var1);`
`put x 16.14`
	`6.000000000000000`

`x=fuzz(5.99999999);` `put x 16.14;`	`5.999999990000000`

FWRITE Function

Writes a record to an external file.

Category: External Files

Syntax

FWRITE(*file-id*<,*cc*>)

Arguments

file-id
is a numeric variable that specifies the identifier that was assigned when the file was opened, generally by the FOPEN function.

cc
is a character constant, variable, or expression that specifies a carriage-control character:

blank	starts the record on a new line.
0	skips one blank line before a new line.
-	skips two blank lines before a new line.
1	starts the line on a new page.
+	overstrikes the line on a previous line.
P	interprets the line as a computer prompt.
=	interprets the line as carriage control information.
all else	starts the line record on a new line.

Details

FWRITE returns 0 if the operation was successful, ≠0 if it was not successful. FWRITE moves text from the File Data Buffer (FDB) to the external file. In order to use the carriage control characters, you must open the file with a record format of **`P`** (print format) in FOPEN.

Note: When you use the update mode, you must execute FREAD before you execute FWRITE. You cannot write a new record in place of the current record if the new record has a length that is greater than the current record. △

Examples

This example assigns the fileref MYFILE to an external file and attempts to open the file. If the file is opened successfully, it writes the numbers 1 to 50 to the external file, skipping two blank lines. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let filrf=myfile;
%let rc=%sysfunc(filename(filrf,
```

```
   physical-filename));
%let fid=%sysfunc(fopen(&filrf,o,0,P));

%do i=1 %to 50;
   %let rc=%sysfunc(fput(&fid,
      %sysfunc(putn(&i,2.))));

   %if (%sysfunc(fwrite(&fid,-)) ne 0) %then
     %put %sysfunc(sysmsg());
%end;

%let rc=%sysfunc(fclose(&fid));
```

See Also

Functions:

"FAPPEND Function" on page 679
"FCLOSE Function" on page 680
"FGET Function" on page 687
"FILENAME Function" on page 690
"FOPEN Function" on page 762
"FPUT Function" on page 771
"SYSMSG Function" on page 1154

GAMINV Function

Returns a quantile from the gamma distribution.

Category: Quantile

Syntax

GAMINV(*p*,*a*)

Arguments

p
is a numeric probability.
Range: $0 \leq p < 1$

a
is a numeric shape parameter.
Range: $a > 0$

Details

The GAMINV function returns the p^{th} quantile from the gamma distribution, with shape parameter a. The probability that an observation from a gamma distribution is less than or equal to the returned quantile is p.

Note: GAMINV is the inverse of the PROBGAM function. △

Examples

SAS Statements	Results
`q1=gaminv(0.5,9);`	`8.6689511844`
`q2=gaminv(0.1,2.1);`	`0.5841932369`

See Also

Functions:

"QUANTILE Function" on page 1064

GAMMA Function

Returns the value of the gamma function.

Category: Mathematical

Syntax

GAMMA(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Restriction: Nonpositive integers are invalid.

Details

The GAMMA function returns the integral given by

$$\mathrm{GAMMA}(x) = \int_0^\infty t^{x-1} e^{-t}\, dt.$$

For positive integers, GAMMA(x) is (x – 1)!. This function is commonly denoted by $\Gamma(x)$.

Examples

SAS Statements	Results
`x=gamma(6);`	`120`

GARKHCLPRC Function

Calculates call prices for European options on stocks, based on the Garman-Kohlhagen model.

Category: Financial

Syntax

GARKHCLPRC(*E*, *t*, *S*, R_d, R_f, *sigma*)

Arguments

E
is a non-missing, positive value that specifies the exercise price.

Requirement: Specify *E* and *S* in the same units.

t
is a non-missing value that specifies the time to maturity.

S
is a non-missing, positive value that specifies the spot currency price.

Requirement: Specify *S* and *E* in the same units.

R_d
is a non-missing, positive fraction that specifies the risk-free domestic interest rate for period *t*.

Requirement: Specify a value for R_d for the same time period as the unit of *t*.

R_f
is a non-missing, positive fraction that specifies the risk-free foreign interest rate for period *t*.

Requirement: Specify a value for R_f for the same time period as the unit of *t*.

sigma
is a non-missing, positive fraction that specifies the volatility of the currency rate.

Requirement: Specify a value for *sigma* for the same time period as the unit of *t*.

Details

The GARKHCLPRC function calculates the call prices for European options on stocks, based on the Garman-Kohlhagen model. The function is based on the following relationship:

$$\texttt{CALL} = SN\left(d_1\right)\left(e^{-R_{f^t}}\right) - EN\left(d_2\right)\left(e^{-R_{d^t}}\right)$$

where

S	specifies the spot currency price.
N	specifies the cumulative normal density function.
E	specifies the exercise price of the option.
t	specifies the time to expiration.
R_{d}	specifies the risk-free domestic interest rate for period t.
R_{f}	specifies the risk-free foreign interest rate for period t.

$$d_1 = \frac{\left(\ln\left(\frac{S}{E}\right) + \left(R_d - R_f + \frac{\sigma^2}{2}\right)t\right)}{\sigma\sqrt{t}}$$
$$d_2 = d_1 - \sigma\sqrt{t}$$

where

σ	specifies the volatility of the underlying asset.
σ^2	specifies the variance of the rate of return.

For the special case of t=0, the following equation is true:

$$\texttt{CALL} = \max\left(\left(S - E\right), 0\right)$$

For information about the basics of pricing, see “Using Pricing Functions” on page 311.

Comparisons

The GARKHCLPRC function calculates the call prices for European options on stocks, based on the Garman-Kohlhagen model. The GARKHPTPRC function calculates the put prices for European options on stocks, based on the Garman-Kohlhagen model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=garkhclprc(1000, .5, 950, 4, 4, 2);` `put a;`	`65.335687119`
`b=garkhclprc(850, 1.2, 125, 5, 3, 1);` `put b;`	`1.9002767538`

SAS Statements	Results
`c=garkhclprc(7500, .9, 950, 3, 2, 2);` `put c;`	`69.328647279`
`d=garkhclprc(5000, -.5, 237, 3, 3, 2);` `put d;`	`0`

See Also

Function:

"GARKHPTPRC Function" on page 783

GARKHPTPRC Function

Calculates put prices for European options on stocks, based on the Garman-Kohlhagen model.

Category: Financial

Syntax

GARKHPTPRC(*E*, *t*, *S*, R_d, R_f, *sigma*)

Arguments

E
is a non-missing, positive value that specifies the exercise price.

Requirement: Specify *E* and *S* in the same units.

t
is a non-missing value that specifies the time to maturity.

S
is a non-missing, positive value that specifies the spot currency price.

Requirement: Specify *S* and *E* in the same units.

R_d
is a non-missing, positive fraction that specifies the risk-free domestic interest rate for period *t*.

Requirement: Specify a value for R_d for the same time period as the unit of *t*.

R_f
is a non-missing, positive fraction that specifies the risk-free foreign interest rate for period *t*.

Requirement: Specify a value for R_t for the same time period as the unit of *t*.

sigma
is a non-missing, positive fraction that specifies the volatility of the currency rate.

Requirement: Specify a value for *sigma* for the same time period as the unit of *t*.

Details

The GARKHPTPRC function calculates the put prices for European options on stocks, based on the Garman-Kohlhagen model. The function is based on the following relationship:

$$\texttt{PUT} = \texttt{CALL} - S\left(e^{-R_{f^t}}\right) + E\left(e^{-R_{d^t}}\right)$$

where

S	specifies the spot currency price.
E	specifies the exercise price of the option.
t	specifies the time to expiration.
R_d	specifies the risk-free domestic interest rate for period t.
R_f	specifies the risk-free foreign interest rate for period t.

$$d_1 = \frac{\left(\ln\left(\frac{S}{E}\right) + \left(R_d - R_f + \frac{\sigma^2}{2}\right)t\right)}{\sigma\sqrt{t}}$$

$$d_2 = d_1 - \sigma\sqrt{t}$$

where

σ	specifies the volatility of the underlying asset.
σ^2	specifies the variance of the rate of return.

For the special case of t=0, the following equation is true:

$$\texttt{PUT} = \max\left(\left(E - S\right), 0\right)$$

For information about the basics of pricing, see “Using Pricing Functions” on page 311.

Comparisons

The GARKHPTPRC function calculates the put prices for European options on stocks, based on the Garman-Kohlhagen model. The GARKHCLPRC function calculates the call prices for European options on stocks, based on the Garman-Kohlhagen model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=garkhptprc(1000, .5. 950, 4, 4, 2): put a;`	`72.102451281`
`b=garkhptprc(850, 1.2. 125, 5, 3, 1): put b;`	`0.5917507981`

SAS Statements	Results
`c=garkhptprc(7500, .9. 950, 3, 2, 2):` `put c;`	`416.33604902`
`d=garkhptprc(5000, -.5, 237, 3, 3, 2):` `put d;`	`0`

See Also

Function:

"GARKHCLPRC Function" on page 781

GCD Function

Returns the greatest common divisor for one or more integers.

Category: Mathematical

Syntax

GCD(*x1, x2, x3, …, xn*)

Arguments

x
specifies a numeric constant, variable, or expression that has an integer value.

Details

The GCD (greatest common divisor) function returns the greatest common divisor of one or more integers. For example, the greatest common divisor for 30 and 42 is 6. The greatest common divisor is also called the highest common factor.

If any of the arguments are missing, then the returned value is a missing value.

Examples

The following example returns the greatest common divisor of the integers 10 and 15.

```
data _null_;
   x=gcd(10, 15);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=5
```

See Also

Functions:

"LCM Function" on page 879

GEODIST Function

Returns the geodetic distance between two latitude and longitude coordinates.

Category: Distance

Syntax

GEODIST(*latitude-1*, *longitude-1*, *latitude-2*, *longitude-2* <,*options*>)

Arguments

latitude
is a numeric constant, variable, or expression that specifies the coordinate of a given position north or south of the equator. Coordinates that are located north of the equator have positive values; coordinates that are located south of the equator have negative values.

Restriction: If the value is expressed in degrees, it must be between 90 and –90. If the value is expressed in radians, it must be between pi/2 and –pi/2.

longitude
is a numeric constant, variable, or expression that specifies the coordinate of a given position east or west of the prime meridian, which runs through Greenwich, England. Coordinates that are located east of the prime meridian have positive values; coordinates that are located west of the prime meridian have negative values.

Restriction: If the value is expressed in degrees, it must be between 180 and –180. If the value is expressed in radians, it must be between pi and –pi.

option
specifies a character constant, variable, or expression that contains any of the following characters:

M	specifies distance in miles.
K	specifies distance in kilometers. K is the default value for distance.
D	specifies that input values are expressed in degrees. D is the default for input values.
R	specifies that input values are expressed in radians.

Details

The GEODIST function computes the geodetic distance between any two arbitrary latitude and longitude coordinates. Input values can be expressed in degrees or in radians.

Examples

Example 1: Calculating the Geodetic Distance in Kilometers The following example shows the geodetic distance in kilometers between Mobile, AL (latitude 30.68 N, longitude 88.25 W), and Asheville, NC (latitude 35.43 N, longitude 82.55 W). The program uses the default K option.

```
data _null_;
   distance=geodist(30.68, -88.25, 35.43, -82.55);
   put 'Distance= ' distance 'kilometers';
run;
```

SAS writes the following output to the log:

```
Distance= 748.6529147 kilometers
```

Example 2: Calculating the Geodetic Distance in Miles The following example uses the M option to compute the geodetic distance in miles between Mobile, AL (latitude 30.68 N, longitude 88.25 W), and Asheville, NC (latitude 35.43 N, longitude 82.55 W).

```
data _null_;
   distance=geodist(30.68, -88.25, 35.43, -82.55, 'M');
   put 'Distance = ' distance 'miles';
run;
```

SAS writes the following output to the log:

```
Distance = 465.29081088 miles
```

Example 3: Calculating the Geodetic Distance with Input Measured in Degrees The following example uses latitude and longitude values that are expressed in degrees to compute the geodetic distance between two locations. Both the D and the M options are specified in the program.

```
data _null_;
   input lat1 long1 lat2 long2;
   Distance = geodist(lat1,long1,lat2,long2,'DM');
   put 'Distance = ' Distance 'miles';
   datalines;
35.2 -78.1 37.6 -79.8
;
run;
```

SAS writes the following output to the log:

```
Distance = 190.72474282 miles
```

Example 4: Calculating the Geodetic Distance with Input Measured in Radians The following example uses latitude and longitude values that are expressed in radians to compute the geodetic distance between two locations. The program converts degrees to radians before executing the GEODIST function. Both the R and the M options are specified in this program.

```
data _null_;
   input lat1 long1 lat2 long2;
   pi = constant('pi');
   lat1 = (pi*lat1)/180;
   long1 = (pi*long1)/180;
   lat2 = (pi*lat2)/180;
   long2 = (pi*long2)/180;
```

```
   Distance = geodist(lat1,long1,lat2,long2,'RM');
   put 'Distance= ' Distance 'miles';
   datalines;
35.2 -78.1 37.6 -79.8
;
run;
```

SAS writes the following output to the log:

```
Distance= 190.72474282 miles
```

References

Vincenty, T. 1975. "Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations." *Survey Review* 22:88–93.

GEOMEAN Function

Returns the geometric mean.

Category: Descriptive Statistics

Syntax

GEOMEAN(*argument*<,*argument*,...>)

Arguments

argument
is a non-negative numeric constant, variable, or expression.

Tip: The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The MEAN function returns the arithmetic mean (average), and the HARMEAN function returns the harmonic mean, whereas the GEOMEAN function returns the geometric mean of the non-missing values. Unlike GEOMEANZ, GEOMEAN fuzzes the values of the arguments that are approximately zero.

Details

If any argument is negative, then the result is a missing value. A message appears in the log that the negative argument is invalid, and _ERROR_ is set to 1. If any argument is zero, then the geometric mean is zero. If all the arguments are missing values, then the result is a missing value. Otherwise, the result is the geometric mean of the non-missing values.

Let n be the number of arguments with non-missing values, and let $x_1, x_2, \ldots, x_n$ be the values of those arguments. The geometric mean is the n^{th} root of the product of the values:

$$\sqrt[n]{(x_1 * x_2 * \ldots * x_n)}$$

Equivalently, the geometric mean is

$$\exp\left(\frac{(\log(x_1) + \log(x_2) + \ldots + \log(x_n))}{n}\right)$$

Floating-point arithmetic often produces tiny numerical errors. Some computations that result in zero when exact arithmetic is used might result in a tiny non-zero value when floating-point arithmetic is used. Therefore, GEOMEAN fuzzes the values of arguments that are approximately zero. When the value of one argument is extremely small relative to the largest argument, then the former argument is treated as zero. If you do not want SAS to fuzz the extremely small values, then use the GEOMEANZ function.

Examples

SAS Statements	Results
`x1=geomean(1,2,2,4);`	`2`
`x2=geomean(.,2,4,8);`	`4`
`x3=geomean(of x1-x2);`	`2.8284271247`

See Also

Function:

"GEOMEANZ Function" on page 790
"HARMEAN Function" on page 799
"HARMEANZ Function" on page 801
"MEAN Function" on page 925

GEOMEANZ Function

Returns the geometric mean, using zero fuzzing.

Category: Descriptive Statistics

Syntax

GEOMEANZ(*argument*<,*argument*,...>)

Arguments

argument
is a non-negative numeric constant, variable, or expression.

Tip: The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The MEAN function returns the arithmetic mean (average), and the HARMEAN function returns the harmonic mean, whereas the GEOMEANZ function returns the geometric mean of the non-missing values. Unlike GEOMEAN, GEOMEANZ does not fuzz the values of the arguments that are approximately zero.

Details

If any argument is negative, then the result is a missing value. A message appears in the log that the negative argument is invalid, and _ERROR_ is set to 1. If any argument is zero, then the geometric mean is zero. If all the arguments are missing values, then the result is a missing value. Otherwise, the result is the geometric mean of the non-missing values.

Let n be the number of arguments with non-missing values, and let $x_1, x_2, \ldots, x_n$ be the values of those arguments. The geometric mean is the n^{th} root of the product of the values:

$$\sqrt[n]{(x_1 * x_2 * \ldots * x_n)}$$

Equivalently, the geometric mean is

$$\exp\left(\frac{(\log(x_1) + \log(x_2) + \ldots + \log(x_n))}{n}\right)$$

Examples

SAS Statements	Results
`x1=geomeanz(1,2,2,4);`	`2`
`x2=geomeanz(.,2,4,8);`	`4`
`x3=geomeanz(of x1-x2);`	`2.8284271247`

See Also

Function:

"GEOMEAN Function" on page 788
"HARMEAN Function" on page 799
"HARMEANZ Function" on page 801
"MEAN Function" on page 925

GETOPTION Function

Returns the value of a SAS system or graphics option.

Category: Special

Syntax

GETOPTION(*option-name*<,*reporting-options*<,...>>)

Arguments

option-name
is a character constant, variable, or expression that specifies the name of the system option.

Tip: Do not put an equal sign after the name. For example, write PAGESIZE= as PAGESIZE.

Tip: SAS options that are passwords, such as EMAILPW and METAPASS, return the value **xxxxxxxx**, and not the actual password.

reporting-options
is a character constant, variable, or expression that specifies the reporting options. You can separate the options with blanks, or you can specify each reporting option as a separate argument to the GETOPTION function. The following is a list of reporting options:

IN	reports graphic units of measure in inches.
CM	reports graphic units of measure in centimeters.
EXPAND | KEYEXPAND	for options that contain environment variables or keywords, returns the value of the environment variable or keyword in the option value.
HOWSET	returns a character string that specifies how an option value was set. **Restriction:** HOWSET is valid only for SAS system options. SAS issues an error message when the HOWSET option is specified and *option-name* is a graphics option.
HOWSCOPE	returns a character string that specifies the scope of an option. **Restriction:** HOWSCOPE is valid only for SAS system options. SAS issues an error message when the HOWSCOPE option is specified and *option-name* is a graphics option.
KEYWORD	returns option values in a KEYWORD= format that would be suitable for direct use in the SAS OPTIONS or GOPTIONS global statements. *Note:* For a system option with a null value, the GETOPTION function returns a value of ”(single quotation marks with a blank space between them), for example **`EMAILID=' '`**. △

Examples

Example 1: Using GETOPTION to Save and Restore the YEARCUTOFF Option This example saves the initial value of the YEARCUTOFF option and then resets the value to 1920. The DATA step that follows verifies the option setting and performs date processing. When the DATA step ends, the YEARCUTOFF option is set to its original value.

```
%let cutoff=%sysfunc(getoption
                     (yearcutoff,keyword));
options yearcutoff=1920;
data ages;
  if getoption('yearcutoff') = '1920' then
     do;
        ...more statements...
     end;
  else put 'Set Option YEARCUTOFF to 1920';
run;

options &cutoff;
```

Example 2: Using GETOPTION to Obtain Different Reporting Options This example defines a macro to illustrate the use of the GETOPTION function to obtain the value of system and graphics options by using different reporting options.

```
%macro showopts;
  %put MAPS= %sysfunc(
     getoption(MAPS));
  %put MAPSEXPANDED= %sysfunc(
     getoption(MAPS, EXPAND));
  %put PAGESIZE= %sysfunc(
     getoption(PAGESIZE));
  %put PAGESIZESETBY= %sysfunc(
     getoption(PAGESIZE, HOWSET));
  %put PAGESIZESCOPE= %sysfunc(
     getoption(PAGESIZE, HOWSCOPE));
  %put PS= %sysfunc(
     getoption(PS));
  %put LS= %sysfunc(
     getoption(LS));
  %put PS(keyword form)= %sysfunc(
     getoption(PS,keyword));
  %put LS(keyword form)= %sysfunc(
     getoption(LS,keyword));
  %put FORMCHAR= %sysfunc(
     getoption(FORMCHAR));
  %put HSIZE= %sysfunc(
     getoption(HSIZE));
  %put VSIZE= %sysfunc(
     getoption(VSIZE));
  %put HSIZE(in/keyword form)= %sysfunc(
     getoption(HSIZE,in,keyword));
  %put HSIZE(cm/keyword form)= %sysfunc(
     getoption(HSIZE,cm,keyword));
  %put VSIZE(in/keyword form)= %sysfunc(
     getoption(VSIZE,in,keyword));
  %put HSIZE(cm/keyword form)= %sysfunc(
     getoption(VSIZE,cm,keyword));
%mend;
goptions VSIZE=8.5 in HSIZE=11 in;
options PAGESIZE=67;

%showopts
```

The following is SAS output from the example.

```
MAPS= !SASROOT\MAPS
MAPSEXPANDED= ( 'C:\PROGRAM FILES\SAS\SAS 9.1\MAPS' )
PAGESIZE= 67
PAGESIZESETBY= OPTIONS STATEMENT
PAGESIZESCOPE= DMS PROCESS
PS= 23
LS= 76
PS(keyword form)= PS=23
LS(keyword form)= LS=76
FORMCHAR= |----|+|---+=|-/\<>*
```

```
HSIZE= 11.0000 in.
VSIZE= 8.5000 in.
HSIZE(in/keyword form)= HSIZE=11.0000 in.
HSIZE(cm/keyword form)= HSIZE=27.9400 cm.
VSIZE(in/keyword form)= VSIZE=8.5000 in.
HSIZE(cm/keyword form)= VSIZE=21.5900 cm.
```

Note: The default settings for the PAGESIZE= and the LINESIZE= options depend on the mode you use to run SAS. △

GETVARC Function

Returns the value of a SAS data set character variable.

Category: SAS File I/O

Syntax

GETVARC(*data-set-id*,*var-num*)

Arguments

data-set-id
is a numeric constant, variable, or expression that specifies the data set identifier that the OPEN function returns.

var-num
is a numeric constant, variable, or expression that specifies the number of the variable in the Data Set Data Vector (DDV).

Tip: You can obtain this value by using the VARNUM function.

Tip: This value is listed next to the variable when you use the CONTENTS procedure.

Details

Use VARNUM to obtain the number of a variable in a SAS data set. VARNUM can be nested or it can be assigned to a variable that can then be passed as the second argument, as shown in the following examples. GETVARC reads the value of a character variable from the current observation in the Data Set Data Vector (DDV) into a macro or DATA step variable.

Examples

- This example opens the SASUSER.HOUSES data set and gets the entire tenth observation. The data set identifier value for the open data set is stored in the macro variable MYDATAID. This example nests VARNUM to return the position of the variable in the DDV, and reads in the value of the character variable STYLE.

```
%let mydataid=%sysfunc(open
                      (sasuser.houses,i));
%let rc=%sysfunc(fetchobs(&mydataid,10));
%let style=%sysfunc(getvarc(&mydataid,
                   %sysfunc(varnum
                   (&mydataid,STYLE))));
%let rc=%sysfunc(close(&mydataid));
```

- This example assigns VARNUM to a variable that can then be passed as the second argument. This example fetches data from observation 10.

```
%let namenum=%sysfunc(varnum(&mydataid,NAME));
%let rc=%sysfunc(fetchobs(&mydataid,10));
%let user=%sysfunc(getvarc
                  (&mydataid,&namenum));
```

See Also

Functions:

"FETCH Function" on page 684
"FETCHOBS Function" on page 685
"GETVARN Function" on page 795
"VARNUM Function" on page 1188

GETVARN Function

Returns the value of a SAS data set numeric variable.

Category: SAS File I/O

Syntax

GETVARN(*data-set-id*,*var-num*)

Arguments

data-set-id
is a numeric constant, variable, or expression that specifies the data set identifier that the OPEN function returns.

var-num
is a numeric constant, variable, or expression that specifies the number of the variable in the Data Set Data Vector (DDV).

Tip: You can obtain this value by using the VARNUM function.

Tip: This value is listed next to the variable when you use the CONTENTS procedure.

Details

Use VARNUM to obtain the number of a variable in a SAS data set. You can nest VARNUM or you can assign it to a variable that can then be passed as the second argument, as shown in the "Examples" section. GETVARN reads the value of a numeric variable from the current observation in the Data Set Data Vector (DDV) into a macro variable or DATA step variable.

Examples

□ This example obtains the entire tenth observation from a SAS data set. The data set must have been previously opened using OPEN. The data set identifier value for the open data set is stored in the variable MYDATAID. This example nests VARNUM, and reads in the value of the numeric variable PRICE from the tenth observation of an open SAS data set.

```
%let rc=%sysfunc(fetchobs(&mydataid,10));
%let price=%sysfunc(getvarn(&mydataid,
                  %sysfunc(varnum
                       (&mydataid,price))));
```

□ This example assigns VARNUM to a variable that can then be passed as the second argument. This example fetches data from observation 10.

```
%let pricenum=%sysfunc(varnum
                     (&mydataid,price));
%let rc=%sysfunc(fetchobs(&mydataid,10));
%let price=%sysfunc(getvarn
                  (&mydataid,&pricenum));
```

See Also

Functions:

“FETCH Function” on page 684

“FETCHOBS Function” on page 685

“GETVARC Function” on page 794

“VARNUM Function” on page 1188

GRAYCODE Function

Generates all subsets of *n* items in a minimal change order.

Category: Combinatorial

Restriction: The GRAYCODE function cannot be executed when you use the %SYSFUNC macro.

Syntax

GRAYCODE(*k*, *numeric-variable-1*, ..., *numeric-variable-n*)

GRAYCODE(*k*, *character-variable* <, *n* <, *in-out*>>)

Arguments

k
specifies a numeric variable. Initialize *k* to either of the following values before executing the GRAYCODE function:

- a negative number to cause GRAYCODE to initialize the subset to be empty
- the number of items in the initial set indicated by *numeric-variable-1* through *numeric-variable-n*, or *character-variable*, which must be an integer value between 0 and *n* inclusive

The value of *k* is updated when GRAYCODE is executed. The value that is returned is the number of items in the subset.

numeric-variable
specifies numeric variables that have values of 0 or 1 which are updated when GRAYCODE is executed. A value of 1 for *numeric-variable-j* indicates that the j^{th} item is in the subset. A value of 0 for *numeric-variable-j* indicates that the j^{th} item is not in the subset.

If you assign a negative value to *k* before you execute GRAYCODE, then you do not need to initialize *numeric-variable-1* through *numeric-variable-n* before executing GRAYCODE unless you want to suppress the note about uninitialized variables.

If you assign a value between 0 and *n* inclusive to *k* before you execute GRAYCODE, then you must initialize *numeric-variable-1* through *numeric-variable-n* to *k* values of 1 and *n*-*k* values of 0.

character-variable
specifies a character variable that has a length of at least *n* characters. The first *n* characters indicate which items are in the subset. By default, an "I" in the j^{th} position indicates that the j^{th} item is in the subset, and an "O" in the j^{th} position indicates that the j^{th} item is out of the subset. You can change the two characters by specifying the *in-out* argument.

If you assign a negative value to *k* before you execute GRAYCODE, then you do not need to initialize *character-variable* before executing GRAYCODE unless you want to suppress the note about an uninitialized variable.

If you assign a value between 0 and *n* inclusive to *k* before you execute GRAYCODE, then you must initialize *character-variable* to *k* characters that indicate an item is in the subset, and *n*-*k* characters that indicate an item is out of the subset.

n
specifies a numeric constant, variable, or expression. By default, *n* is the length of *character-variable*.

in-out
specifies a character constant, variable, or expression. The default value is "IO." The first character is used to indicate that an item is in the subset. The second character is used to indicate that an item is out of the subset.

Details

When you execute GRAYCODE with a negative value of *k*, the subset is initialized to be empty. The GRAYCODE function returns zero.

When you execute GRAYCODE with an integer value of *k* between 0 and *n* inclusive, one item is either added to the subset or removed from the subset, and the value of *k* is updated to equal the number of items in the subset. If the j^{th} item is added to the subset or removed from the subset, the GRAYCODE function returns *j*.

To generate all subsets of *n* items, you can initialize *k* to a negative value and execute GRAYCODE in a loop that iterates 2**n times. If you want to start with a non-empty subset, then initialize *k* to be the number of items in the subset, initialize the other arguments to specify the desired initial subset, and execute GRAYCODE in a loop that iterates 2**n-1 times. The sequence of subsets that are generated by GRAYCODE is cyclical, so you can begin with any subset you want.

Examples

Example 1: Using *n*=4 Numeric Variables and Negative Initial *k* The following program uses numeric variables to generate subsets in a minimal change order.

```
data _null_;
   array x[4];
   n=dim(x);
   k=-1;
   nsubs=2**n;
   do i=1 to nsubs;
      rc=graycode(k, of x[*]);
      put i 5. +3 k= ' x=' x[*] +3 rc=;
   end;
run;
```

SAS writes the following output to the log:

```
    1   k=0  x=0 0 0 0    rc=0
    2   k=1  x=1 0 0 0    rc=1
    3   k=2  x=1 1 0 0    rc=2
    4   k=1  x=0 1 0 0    rc=1
    5   k=2  x=0 1 1 0    rc=3
    6   k=3  x=1 1 1 0    rc=1
    7   k=2  x=1 0 1 0    rc=2
    8   k=1  x=0 0 1 0    rc=1
    9   k=2  x=0 0 1 1    rc=4
   10   k=3  x=1 0 1 1    rc=1
   11   k=4  x=1 1 1 1    rc=2
   12   k=3  x=0 1 1 1    rc=1
   13   k=2  x=0 1 0 1    rc=3
   14   k=3  x=1 1 0 1    rc=1
```

```
15   k=2  x=1 0 0 1    rc=2
16   k=1  x=0 0 0 1    rc=1
```

Example 2: Using a Character Variable and Positive Initial *k* The following example uses a character variable to generate subsets in a minimal change order.

```
data _null_;
   x='++++';
   n=length(x);
   k=countc(x, '+');
   put '    1' +3 k= +2 x=;
   nsubs=2**n;
   do i=2 to nsubs;
      rc=graycode(k, x, n, '+-');
      put i 5. +3 k= +2 x= +3 rc=;
   end;
run;
```

SAS writes the following output to the log:

```
 1   k=4   x=++++
 2   k=3   x=-+++    rc=1
 3   k=2   x=-+-+    rc=3
 4   k=3   x=++-+    rc=1
 5   k=2   x=+--+    rc=2
 6   k=1   x=---+    rc=1
 7   k=0   x=----    rc=4
 8   k=1   x=+---    rc=1
 9   k=2   x=++--    rc=2
10   k=1   x=-+--    rc=1
11   k=2   x=-++-    rc=3
12   k=3   x=+++-    rc=1
13   k=2   x=+-+-    rc=2
14   k=1   x=--+-    rc=1
15   k=2   x=--++    rc=4
16   k=3   x=+-++    rc=1
```

See Also

Functions and CALL Routines:

"CALL GRAYCODE Routine" on page 450

HARMEAN Function

Returns the harmonic mean.

Category: Descriptive Statistics

Syntax

HARMEAN(*argument*<,*argument*,...>)

Arguments

argument
is a non-negative numeric constant, variable, or expression.

Tip: The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The MEAN function returns the arithmetic mean (average), and the GEOMEAN function returns the geometric mean, whereas the HARMEAN function returns the harmonic mean of the non-missing values. Unlike HARMEANZ, HARMEAN fuzzes the values of the arguments that are approximately zero.

Details

If any argument is negative, then the result is a missing value. A message appears in the log that the negative argument is invalid, and _ERROR_ is set to 1. If all the arguments are missing values, then the result is a missing value. Otherwise, the result is the harmonic mean of the non-missing values.

If any argument is zero, then the harmonic mean is zero. Otherwise, the harmonic mean is the reciprocal of the arithmetic mean of the reciprocals of the values.

Let n be the number of arguments with non-missing values, and let $x_1, x_2, \ldots, x_n$ be the values of those arguments. The harmonic mean is

$$\frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

Floating-point arithmetic often produces tiny numerical errors. Some computations that result in zero when exact arithmetic is used might result in a tiny non-zero value when floating-point arithmetic is used. Therefore, HARMEAN fuzzes the values of arguments that are approximately zero. When the value of one argument is extremely small relative to the largest argument, then the former argument is treated as zero. If you do not want SAS to fuzz the extremely small values, then use the HARMEANZ function.

Examples

SAS Statements	Results
`x1=harmean(1,2,4,4);`	**2**
`x2=harmean(.,4,12,24);`	**8**
`x3=harmean(of x1-x2);`	**3.2**

See Also

Function:

"GEOMEAN Function" on page 788
"GEOMEANZ Function" on page 790
"HARMEANZ Function" on page 801
"MEAN Function" on page 925

HARMEANZ Function

Returns the harmonic mean, using zero fuzzing.

Category: Descriptive Statistics

Syntax

HARMEANZ(*argument*<,*argument*,...>)

Arguments

argument
is a non-negative numeric constant, variable, or expression.

Tip: The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The MEAN function returns the arithmetic mean (average), and the GEOMEAN function returns the geometric mean, whereas the HARMEANZ function returns the harmonic mean of the non-missing values. Unlike HARMEAN, HARMEANZ does not fuzz the values of the arguments that are approximately zero.

Details

If any argument is negative, then the result is a missing value. A message appears in the log that the negative argument is invalid, and _ERROR_ is set to 1. If all the arguments are missing values, then the result is a missing value. Otherwise, the result is the harmonic mean of the non-missing values.

If any argument is zero, then the harmonic mean is zero. Otherwise, the harmonic mean is the reciprocal of the arithmetic mean of the reciprocals of the values.

Let n be the number of arguments with non-missing values, and let $x_1, x_2, \ldots, x_n$ be the values of those arguments. The harmonic mean is

$$\frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

Examples

SAS Statements	Results
`x1=harmeanz(1,2,4,4);`	`2`
`x2=harmeanz(.,4,12,24);`	`8`
`x3=harmeanz(of x1-x2);`	`3.2`

See Also

Function:

"GEOMEAN Function" on page 788
"GEOMEANZ Function" on page 790
"HARMEAN Function" on page 799
"MEAN Function" on page 925

HBOUND Function

Returns the upper bound of an array.

Category: Array

Syntax

HBOUND<*n*>(*array-name*)

HBOUND(*array-name*,*bound-n*)

Arguments

n
is an integer constant that specifies the dimension for which you want to know the upper bound. If no *n* value is specified, the HBOUND function returns the upper bound of the first dimension of the array.

array-name
is the name of an array that was defined previously in the same DATA step.

bound-n
is a numeric constant, variable, or expression that specifies the dimension for which you want to know the upper bound. Use *bound-n* only if *n* is not specified.

Details

The HBOUND function returns the upper bound of a one-dimensional array or the upper bound of a specified dimension of a multidimensional array. Use HBOUND in

array processing to avoid changing the upper bound of an iterative DO group each time you change the bounds of the array. HBOUND and LBOUND can be used together to return the values of the upper and lower bounds of an array dimension.

Comparisons

- HBOUND returns the literal value of the upper bound of an array dimension.
- DIM always returns a total count of the number of elements in an array dimension.

  Note: This distinction is important when the lower bound of an array dimension has a value other than 1 and the upper bound has a value other than the total number of elements in the array dimension. △

Examples

Example 1: One-dimensional Array In this example, HBOUND returns the upper bound of the dimension, a value of 5. Therefore, SAS repeats the statements in the DO loop five times.

```
array big{5} weight sex height state city;
do i=1 to hbound(big5);
   more SAS statements;
end;
```

Example 2: Multidimensional Array This example shows two ways of specifying the HBOUND function for multidimensional arrays. Both methods return the same value for HBOUND, as shown in the table that follows the SAS code example.

```
array mult{2:6,4:13,2} mult1-mult100;
```

Syntax	Alternative Syntax	Value
HBOUND(MULT)	HBOUND(MULT,1)	6
HBOUND2(MULT)	HBOUND(MULT,2)	13
HBOUND3(MULT)	HBOUND(MULT,3)	2

See Also

Functions:

"DIM Function" on page 655
"LBOUND Function" on page 878

Statements:

"ARRAY Statement" on page 1440
"Array Reference Statement" on page 1445

"Array Processing" in *SAS Language Reference: Concepts*

HMS Function

Returns a SAS time value from hour, minute, and second values.

Category: Date and Time

Syntax

HMS(*hour,minute,second*)

Arguments

hour
is numeric.

minute
is numeric.

second
is numeric.

Details

The HMS function returns a positive numeric value that represents a SAS time value.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`hrid=hms(12,45,10);`	
`put hrid`	`45910`
`/ hrid time.;`	`12:45:10`

See Also

Functions:

"DHMS Function" on page 651
"HOUR Function" on page 807
"MINUTE Function" on page 928
"SECOND Function" on page 1122

HOLIDAY Function

Returns a SAS date value of a specified holiday for a specified year.

Category: Date and Time

Syntax

HOLIDAY('*holiday*', *year*)

Arguments

'holiday'
is a character constant, variable, or expression that specifies one of the values listed in the following table.

Values for *holiday* can be in uppercase or lowercase.

Table 4.5 Holidays Recognized By SAS

Holiday Value	Description	Date Celebrated
BOXING	Boxing Day	December 26
CANADA	Canadian Independence Day	July 1
CANADAOBSERVED	Canadian Independence Day observed	July 1, or July 2 if July 1 is a Sunday
CHRISTMAS	Christmas	December 25
COLUMBUS	Columbus Day	2nd Monday in October
EASTER	Easter Sunday	date varies
FATHERS	Father's Day	3rd Sunday in June
HALLOWEEN	Halloween	October 31
LABOR	Labor Day	1st Monday in September
MLK	Martin Luther King, Jr. 's birthday	3rd Monday in January beginning in 1986
MEMORIAL	Memorial Day	last Monday in May (since 1971)
MOTHERS	Mother's Day	2nd Sunday in May
NEWYEAR	New Year's Day	January 1
THANKSGIVING	U.S. Thanksgiving Day	4th Thursday in November
THANKSGIVINGCANADA	Canadian Thanksgiving Day	2nd Monday in October
USINDEPENDENCE	U.S. Independence Day	July 4
USPRESIDENTS	Abraham Lincoln's and George Washington's birthdays observed	3rd Monday in February (since 1971)
VALENTINES	Valentine's Day	February 14
VETERANS	Veterans Day	November 11
VETERANSUSG	Veterans Day - U.S. government-observed	U.S. government-observed date for Monday–Friday schedule

Holiday Value	Description	Date Celebrated
VETERANSUSPS	Veterans Day - U.S. post office observed	U.S. government-observed date for Monday–Saturday schedule (U.S. Post Office)
VICTORIA	Victoria Day	Monday on or preceding May 24

year
is a numeric constant, variable, or expression that specifies a four-digit year. If you use a two-digit year, then you must specify the YEARCUTOFF= system option.

Details

The HOLIDAY function computes the date on which a specific holiday occurs in a specified year. Only certain common U.S. and Canadian holidays are defined for use with this function. (See Table 4.5 on page 805 for a list of valid holidays.)

The HOLIDAY function returns a SAS date value. To convert the SAS date value to a calendar date, use any valid SAS date format, such as the DATE9. format.

Comparisons

In some cases, the HOLIDAY function and the NWKDOM function return the same result. For example, the statement **`HOLIDAY('THANKSGIVING', 2007);`** returns the same value as **`NWKDOM(4, 5, 11, 2007);`**.

In other cases, the HOLIDAY function and the MDY function return the same result. For example, the statement **`HOLIDAY('CHRISTMAS', 2007);`** returns the same value as **`MDY(12, 25, 2007);`**.

Examples

The following examples give these results:

SAS Statements	Results

`thanks = holiday('thanksgiving', 2007);`
`format thanks date9.;`
`put thanks;`
	`22NOV2007`

`boxing = holiday('boxing', 2007);`
`format boxing date9.;`
`put boxing;`
	`26DEC2007`

`easter = holiday('easter', 2007);`
`format easter date9.;`
`put easter;`
	`08APR2007`

`canada = holiday('canada', 2007);`
`format canada date9.;`
`put canada;`
	`01JUL2007`

`fathers = holiday('fathers', 2007);`
`format fathers date9.;`
`put fathers;`
	`17JUN2007`

SAS Statements	Results

`valentines = holiday('valentines', 2007);`
`format valentines date9.;`
`put valentines;`
	`14FEB2007`

`victoria = holiday('victoria', 2007);`
`format victoria date9.;`
`put victoria;`
	`21MAY2007`

See Also

Functions:

"NWKDOM Function" on page 978

"MDY Function" on page 924

HOUR Function

Returns the hour from a SAS time or datetime value.

Category: Date and Time

Syntax

HOUR(<*time* | *datetime*>)

Arguments

time
is a numeric constant, variable, or expression that specifies a SAS time value.

datetime
is a numeric constant, variable, or expression that specifies a SAS datetime value.

Details

The HOUR function returns a numeric value that represents the hour from a SAS time or datetime value. Numeric values can range from 0 through 23. HOUR always returns a positive number.

Examples

The following SAS statements produce these results:

SAS Statements	Results

`now='1:30't;`
`h=hour(now);`
`put h;`
	`1`

See Also

Functions:

"MINUTE Function" on page 928
"SECOND Function" on page 1122

HTMLDECODE Function

Decodes a string that contains HTML numeric character references or HTML character entity references, and returns the decoded string.

Category: Web Tools

Restriction: "I18N Level 2" on page 314

Syntax

HTMLDECODE(*expression*)

Arguments

expression
specifies a character constant, variable, or expression.

Details

The HTMLDECODE function recognizes the following character entity references:

Character entity reference	decoded character
&amp;	&
&lt;	<
&gt;	>
&quot;	"
&apos;	'

Unrecognized entities (&<name>;) are left unmodified in the output string.

The HTMLDECODE function recognizes numeric entity references that are of the form

&#*nnn*; where *nnn* specifies a decimal number that contains one or more digits.

&#X*nnn*; where *nnn* specifies a hexadecimal number that contains one or more digits.

Operating Environment Information: Numeric character references that cannot be represented in the current SAS session encoding will not be decoded. The reference will be copied unchanged to the output string. △

Examples

SAS Statements	Results
`x1=htmldecode('not a <tag>');`	`not a <tag>`
`x2=htmldecode('&');`	`'&'`
`x3=htmldecode ('&#65;&#66;&#67;');`	`'ABC'`

See Also

Function:

"HTMLENCODE Function" on page 809

HTMLENCODE Function

Encodes characters using HTML character entity references, and returns the encoded string.

Category: Web Tools

Restriction: "I18N Level 2" on page 314

Syntax

HTMLENCODE(*expression*, <*options*>)

Arguments

expression
specifies a character constant, variable, or expression. By default, any greater-than (>), less-than (<), and ampersand (&) characters are encoded as **>**, **<**, and **&**, respectively. In SAS 9 only, this behavior can be modified with the *options* argument.

Note: The encoded string can be longer than the output string. You should take the additional length into consideration when you define your output variable. If the encoded string exceeds the maximum length that is defined, the output string might be truncated. △

options
is a character constant, variable, or expression that specifies the type of characters to encode. If you use more than one option, separate the options by spaces. The following options are available:

Option	Character	Character Entity Reference	Description
amp	&	&amp;	The HTMLENCODE function encodes these characters by default. If you need to encode these characters only, then you do not need to specify the options argument. However, if you specify any value for the options argument, then the defaults are overridden, and you must explicitly specify the options for all of the characters you want to encode.
gt	>	&gt;	
lt	<	&lt;	
apos	’	&apos;	Use this option to encode the apostrophe (’) character in text that is used in an HTML or XML tag attribute.
quot	"	&quot;	Use this option to encode the double quotation mark (") character in text that is used in an HTML or XML tag attribute.
7bit	any character that is not represented in 7-bit ASCII encoding	&#x*nnn*; (Unicode)	*nnn* is a one or more digit hexadecimal number. Encode these characters to create HTML or XML that is easily transferred through communication paths that might support only 7-bit ASCII encodings (for example, ftp or e-mail).

Examples

SAS Statements	Results
`htmlencode("John’s test <tag>")`	`John’s test <tag>`
`htmlencode("John’s test <tag>",’apos’)`	`John&apos;s test <tag>`
`htmlencode(’John "Jon" Smith <tag>’,’quot’)`	`John &quot;Jon&quot; Smith <tag>`

SAS Statements	Results
`htmlencode("'A&B&C'",'amp lt gt apos')`	`&apos;A&B&C&apos;`
`htmlencode('80'x, '7bit')` ('80'x is the euro symbol in Western European locales.)	`&#x20AC;` (20AC is the Unicode code point for the euro symbol.)

See Also

Function:

"HTMLDECODE Function" on page 808

IBESSEL Function

Returns the value of the modified Bessel function.

Category: Mathematical

Syntax

IBESSEL(*nu*,*x*,*kode*)

Arguments

nu
specifies a numeric constant, variable, or expression.

Range: $nu \geq 0$

x
specifies a numeric constant, variable, or expression.

Range: $x \geq 0$

kode
is a numeric constant, variable, or expression that specifies a nonnegative integer.

Details

The IBESSEL function returns the value of the modified Bessel function of order *nu* evaluated at *x* (Abramowitz, Stegun 1964; Amos, Daniel, Weston 1977). When *kode* equals 0, the Bessel function is returned. Otherwise, the value of the following function is returned:

$$e^{-x} I_{nm}(x)$$

Examples

SAS Statements	Results
`x=ibessel(2,2,0);`	`0.6889484477`
`x=ibessel(2,2,1);`	`0.0932390333`

IFC Function

Returns a character value based on whether an expression is true, false, or missing.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

IFC(*logical-expression*, *value-returned-when-true*, *value-returned-when-false* <,*value-returned-when-missing*>)

Arguments

logical-expression
specifies a numeric constant, variable, or expression.

value-returned-when-true
specifies a character constant, variable, or expression that is returned when the value of *logical-expression* is true.

value-returned-when-false
specifies a character constant, variable, or expression that is returned when the value of *logical-expression* is false.

value-returned-when-missing
specifies a character constant, variable, or expression that is returned when the value of *logical-expression* is missing.

Details

Length of Returned Variable In a DATA step, if the IFC function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The IFC function uses conditional logic that enables you to select among several values based on the value of a logical expression.

IFC evaluates the first argument, *logical-expression*. If *logical-expression* is true (that is, not zero and not missing), then IFC returns the value in the second argument. If *logical-expression* is a missing value, and you have a fourth argument, then IFC returns

the value in the fourth argument. Otherwise, if *logical-expression* is false, IFC returns the value in the third argument.

The IFC function is useful in DATA step expressions, and even more useful in WHERE clauses and other expressions where it is not convenient or possible to use an IF/THEN/ELSE construct.

Comparisons

The IFC function is similar to the IFN function except that IFC returns a character value while IFN returns a numeric value.

Examples

In the following example, IFC evaluates the expression **grade>80** to implement the logic that determines the performance of several members on a team. The results are written to the SAS log.

```
data _null_;
   input name $ grade;
   performance = ifc(grade>80, 'Pass             ', 'Needs Improvement');
   put name= performance=;
   datalines;
John 74
Kareem 89
Kati 100
Maria 92
;

run;
```

Output 4.49 Partial SAS Log: IFC Function

```
name=John performance=Needs Improvement
name=Kareem performance=Pass
name=Kati performance=Pass
name=Maria performance=Pass
```

This example uses an IF/THEN/ELSE construct to generate the same output that is generated by the IFC function. The results are written to the SAS log.

```
data _null_;
   input name $ grade;
   if grade>80 then performance='Pass             ';
      else performance = 'Needs Improvement';
   put name= performance=;
   datalines;
John 74
Sam 89
Kati 100
Maria 92
;

run;
```

Output 4.50 Partial SAS Log: IF/THEN/ELSE Construct

```
name=John performance=Needs Improvement
name=Sam performance=Pass
name=Kati performance=Pass
name=Maria performance=Pass
```

See Also

Functions:

"IFN Function" on page 814

IFN Function

Returns a numeric value based on whether an expression is true, false, or missing.

Category: Numeric

Restriction: "I18N Level 2" on page 314

Syntax

IFN(*logical-expression*, *value-returned-when-true*, *value-returned-when-false* <,*value-returned-when-missing*>)

Arguments

logical-expression
specifies a numeric constant, variable, or expression.

value-returned-when-true
specifies a numeric constant, variable, or expression that is returned when the value of *logical-expression* is true.

value-returned-when-false
specifies a numeric constant, variable, or expression that is returned when the value of *logical-expression* is false.

value-returned-when-missing
specifies a numeric constant, variable or expression that is returned when the value of *logical-expression* is missing.

Details

The IFN function uses conditional logic that enables you to select among several values based on the value of a logical expression.

IFN evaluates the first argument, then *logical-expression*. If *logical-expression* is true (that is, not zero and not missing), then IFN returns the value in the second argument. If *logical-expression* is a missing value, and you have a fourth argument, then IFN

returns the value in the fourth argument. Otherwise, if *logical-expression* is false, IFN returns the value in the third argument.

The IFN function, an IF/THEN/ELSE construct, or a WHERE statement can produce the same results. (See "Examples" on page 815.) However, the IFN function is useful in DATA step expressions when it is not convenient or possible to use an IF/THEN/ELSE construct or a WHERE statement.

Comparisons

The IFN function is similar to the IFC function, except that IFN returns a numeric value whereas IFC returns a character value.

Examples

Example 1: Calculating Sales Commission The following examples show how to calculate sales commission using the IFN function, an IF/THEN/ELSE construct, and a WHERE statement. In each of the examples, the commission that is calculated is the same.

Calculating Commission Using the IFN Function In the following example, IFN evaluates the expression **`TotalSales > 10000`**. If total sales exceeds $10,000, then the sales commission is 5% of the total sales. If total sales is less than $10,000, then the sales commission is 2% of the total sales.

```
data _null_;
   input TotalSales;
   commission=ifn(TotalSales > 10000, TotalSales*.05, TotalSales*.02);
   put commission=;
   datalines;
25000
10000
500
10300
;

run;
```

SAS writes the following output to the log:

```
commission=1250
commission=200
commission=10
commission=515
```

Calculating Commission Using an IF/THEN/ELSE Construct In the following example, an IF/THEN/ELSE construct evaluates the expression **`TotalSales > 10000`**. If total sales exceeds $10,000, then the sales commission is 5% of the total sales. If total sales is less than $10,000, then the sales commission is 2% of the total sales.

```
data _null_;
   input TotalSales;
   if TotalSales > 10000 then commission = .05 * TotalSales;
      else commission = .02 * TotalSales;
   put commission=;
   datalines;
25000
```

```
10000
500
10300
;

run;
```

SAS writes the following output to the log:

```
commission=1250
commission=200
commission=10
commission=515
```

Calculating Commission Using a WHERE Statement In the following example, a WHERE statement evaluates the expression **TotalSales > 10000**. If total sales exceeds $10,000, then the sales commission is 5% of the total sales. If total sales is less than $10,000, then the sales commission is 2% of the total sales. The output shows only those salespeople whose total sales exceed $10,000.

```
options pageno=1 nodate ls=80 ps=64;

data sales;
   input SalesPerson $ TotalSales;
   datalines;
Michaels 25000
Janowski 10000
Chen 500
Gupta 10300
;

data commission;
   set sales;
   where TotalSales > 10000;
   commission = TotalSales * .05;
run;

proc print data=commission;
   title 'Commission for Total Sales > 1000';
run;
```

Output 4.51 Output from a WHERE Statement

```
                    Commission for Total Sales > 1000                    1

                          Sales       Total
                    Obs   Person      Sales    commission

                     1    Michaels    25000       1250
                     2    Gupta       10300        515
```

See Also

Functions:

"IFC Function" on page 812

INDEX Function

Searches a character expression for a string of characters, and returns the position of the string's first character for the first occurrence of the string.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KINDEX in *SAS National Language Support (NLS): Reference Guide*. See "DBCS Compatibility" on page 817.

Syntax

INDEX(*source,excerpt*)

Arguments

source
specifies a character constant, variable, or expression to search.

excerpt
is a character constant, variable, or expression that specifies the string of characters to search for in *source*.

Tip: Enclose a literal string of characters in quotation marks.

Tip: Both leading and trailing spaces are considered part of the *excerpt* argument. To remove trailing spaces, include the TRIM function with the *excerpt* variable inside the INDEX function.

Details

The Basics The INDEX function searches *source*, from left to right, for the first occurrence of the string specified in *excerpt*, and returns the position in *source* of the string's first character. If the string is not found in *source*, INDEX returns a value of 0. If there are multiple occurrences of the string, INDEX returns only the position of the first occurrence.

DBCS Compatibility

The DBCS equivalent function is KINDEX, which is documented in *SAS National Language Support (NLS): Reference Guide*. However, there is a minor difference in the way trailing blanks are handled. In KINDEX, multiple blanks in the second argument match a single blank in the first argument. The following example shows the differences between the two functions:

```
index('ABC,DE F(X=Y)',','       ')            => 0
kindex('ABC,DE F(X=Y)',','       ')          => 7
```

Examples

Example 1: Finding the Position of a Variable in the Source String The following example finds the first position of the *excerpt* argument in *source*.

```
data _null_;
   a = 'ABC.DEF(X=Y)';
   b = 'X=Y';
   x = index(a,b);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=9
```

Example 2: Removing Trailing Spaces When You Use the INDEX Function with the TRIM Function The following example shows the results when you use the INDEX function with and without the TRIM function. If you use INDEX without the TRIM function, leading and trailing spaces are considered part of the *excerpt* argument. If you use INDEX with the TRIM function, TRIM removes trailing spaces from the *excerpt* argument as you can see in this example. Note that the TRIM function is used inside the INDEX function.

```
options nodate nostimer ls=78 ps=60;

data _null_;
   length a b $14;
   a='ABC.DEF (X=Y)';
   b='X=Y';
   q=index(a,b);
   w=index(a,trim(b));
   put q= w=;
run;
```

SAS writes the following output to the log:

```
q=0 w=10
```

See Also

Functions:

"FIND Function" on page 735
"INDEXC Function" on page 819
"INDEXW Function" on page 820

INDEXC Function

Searches a character expression for any of the specified characters, and returns the position of that character.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KINDEXC in *SAS National Language Support (NLS): Reference Guide*.

Syntax

INDEXC(*source,excerpt-1*<,... *excerpt-n*>)

Arguments

source
specifies a character constant, variable, or expression to search.

excerpt
specifies the character constant, variable, or expression to search for in *source*.

Tip: If you specify more than one excerpt, separate them with a comma.

Details

The INDEXC function searches *source*, from left to right, for the first occurrence of any character present in the excerpts and returns the position in *source* of that character. If none of the characters in *excerpt-1* through *excerpt-n* in *source* are found, INDEXC returns a value of 0.

Comparisons

The INDEXC function searches for the first occurrence of any individual character that is present within the character string, whereas the INDEX function searches for the first occurrence of the character string as a substring. The FINDC function provides more options.

Examples

SAS Statements	Results

`a='ABC.DEP (X2=Y1)';`
`x=indexc(a,'0123',';()=.');`
`put x;`
	`4`

`b='have a good day';`
`x=indexc(b,'pleasant','very');`
`put x;`
	`2`

See Also

Functions:

"FINDC Function" on page 737
"INDEX Function" on page 817
"INDEXW Function" on page 820

INDEXW Function

Searches a character expression for a string that is specified as a word, and returns the position of the first character in the word.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

INDEXW(*source, excerpt*<,*delimiters*>)

Arguments

source
specifies a character constant, variable, or expression to search.

excerpt
specifies a character constant, variable, or expression to search for in *source*. SAS removes leading and trailing delimiters from *excerpt*.

delimiter
specifies a character constant, variable, or expression containing the characters that you want INDEXW to use as delimiters in the character string. The default delimiter is the blank character.

Details

The INDEXW function searches *source*, from left to right, for the first occurrence of *excerpt* and returns the position in *source* of the substring's first character. If the substring is not found in *source*, then INDEXW returns a value of 0. If there are multiple occurrences of the string, then INDEXW returns only the position of the first occurrence.

The substring pattern must begin and end on a word boundary. For INDEXW, word boundaries are delimiters, the beginning of *source*, and the end of *source*. If you use an alternate delimiter, then INDEXW does not recognize the end of the text as the end data.

INDEXW has the following behavior when the second argument contains blank spaces or has a length of 0:

- If both *source* and *excerpt* contain only blank spaces or have a length of 0, then INDEXW returns a value of 1.

- If *excerpt* contains only blank spaces or has a length of 0, and *source* contains character or numeric data, then INDEXW returns a value of 0.

Comparisons

The INDEXW function searches for strings that are words, whereas the INDEX function searches for patterns as separate words or as parts of other words. INDEXC searches for any characters that are present in the excerpts. The FINDW function provides more options.

Examples

Example 1: Table of SAS Examples The following SAS statements give these results.

SAS Statements	Results

`s='asdf adog dog';`
`p='dog  ';`
`x=indexw(s,p);`
`put x;`
	`11`

`s='abcdef x=y';`
`p='def';`
`x=indexw(s,p);`
`put x;`
	`0`

`x="abc,def@ xyz";`
`abc=indexw(x, " abc ", "@");`
`put abc;`
	`0`

`x="abc,def@ xyz";`
`comma=indexw(x, ",", "@");`
`put comma;`
	`0`

`x='abc,def% xyz';`
`def=indexw(x, 'def', '%,');`
`put def;`
	`5`

`x="abc,def@ xyz";`
`at=indexw(x, "@", "@");`
`put at;`
	`0`

`x="abc,def@ xyz";`
`xyz=indexw(x, " xyz", "@");`
`put xyz;`
	`9`

`c=indexw(trimn(' '), '        ');`	`1`
`g=indexw('  x  y  ', trimn(' '));`	`0`

Example 2: Using a Semicolon (;) As the Delimiter

The following example shows how to use the semicolon delimiter in a SAS program that also calls the CATX function. A semicolon delimeter must be in place after each call to CATX, and the second argument in the INDEXW function must be trimmed or searches will not be successful.

```
data temp;
   infile datalines;
   input name $12.;
   datalines;
abcdef
```

```
abcdef
;
run;

data temp2;
   set temp;
   format name_list $1024.;
   retain name_list ' ';
   exists=indexw(name_list, trim(name), ';');
   if exists=0 then do
      name_list=catx(';', name_list, name)||';' ;
   name_count +1;
      put '-------------------------------';
      put exists= ;
      put name_list= ;
      put name_count= ;
   end;
 run;
```

Output 4.52 Output from Using a Semicolon As the Delimiter

```
-------------------------------
exists=0
name_list=abcdef;
name_count=1
```

In this example, the first time that CATX is called *name_list* is blank and the value of *name* is 'abcdef'. CATX returns 'abcdef' with no semicolon appended. However, when INDEXW is called the second time, the value of *name_list* is 'abcdef' followed by 1018 (1024–6) blanks, and the value of *name* is 'abcdef' followed by six blanks. Because the third argument in INDEXW is a semicolon (;), the blanks are significant and do not denote a word boundary. Therefore, the second argument cannot be found in the first argument.

If the example has no blanks, the behavior of INDEXW is easier to understand. In the following example, we expect the value of *x* to be 0 because the complete word ABCDE was not found in the first argument:

```
x = indexw('ABCDEF;XYZ', 'ABCDE', ';');
```

The only values for the second argument that would return a nonzero result are ABCDEF and XYZ.

Example 3: Using a Space As the Delimiter

The following example uses a space as a delimiter:

```
data temp;
   infile datalines;
   input name $12.;
   datalines;
abcdef
abcdef
;
run;

data temp2;
```

```
   set temp;
   format name_list $1024.;
   retain name_list ' ';
   exists=indexw(name_list, name, ' ');
   if exists=0 then do
      name_list=catx(' ', name_list, name) ;
   name_count +1;
      put '-------------------------------';
      put exists= ;
      put name_list= ;
      put name_count= ;
   end;
run;
```

Output 4.53 Output from Using a Space as the Delimiter

```
-------------------------------
exists=0
name_list=abcdef
name_count=1
```

See Also

Functions:

"FINDW Function" on page 743
"INDEX Function" on page 817
"INDEXC Function" on page 819

INPUT Function

Returns the value that is produced when SAS converts an expression using the specified informat.

Category: Special

Syntax

INPUT(*source*, <? | ??>*informat*.)

Arguments

source
specifies a character constant, variable, or expression to which you want to apply a specific informat.

? or ??
specifies the optional question mark (?) and double question mark (??) modifiers that suppress the printing of both the error messages and the input lines when invalid data values are read. The ? modifier suppresses the invalid data message. The ?? modifier also suppresses the invalid data message and, in addition, prevents the automatic variable _ERROR_ from being set to 1 when invalid data are read.

informat.
is the SAS informat that you want to apply to the source. This argument must be the name of an informat followed by a period, and cannot be a character constant, variable, or expression.

Details

If the INPUT function returns a character value to a variable that has not yet been assigned a length, by default the variable length is determined by the width of the informat.

The INPUT function enables you to convert the value of *source* by using a specified informat. The informat determines whether the result is numeric or character. Use INPUT to convert character values to numeric values or other character values.

Comparisons

The INPUT function returns the value produced when a SAS expression is converted using a specified informat. You must use an assignment statement to store that value in a variable. The INPUT statement uses an informat to read a data value. Storing that value in a variable is optional.

The INPUT function requires the informat to be specified as a name followed by a period and optional decimal specification. The INPUTC and INPUTN functions allow the informat to be specified as a character constant, variable, or expression.

Examples

Example 1: Converting Character Values to Numeric Values This example uses the INPUT function to convert a character value to a numeric value and store it in another variable. The COMMA9. informat reads the value of the SALE variable, stripping the commas. The resulting value, 2115353, is stored in FMTSALE.

```
data testin;
   input sale $9.;
   fmtsale=input(sale,comma9.);
   datalines;
2,115,353
;
```

Example 2: Using PUT and INPUT Functions In this example, PUT returns a numeric value as a character string. The value 122591 is assigned to the CHARDATE variable. INPUT returns the value of the character string as a SAS date value using a SAS date informat. The value 11681 is stored in the SASDATE variable.

```
numdate=122591;
chardate=put(numdate,z6.);
sasdate=input(chardate,mmddyy6.);
```

Example 3: Suppressing Error Messages In this example, the question mark (?) modifier tells SAS not to print the invalid data error message if it finds data errors. The automatic variable _ERROR_ is set to 1 and input data lines are written to the SAS log.

```
y=input(x,? 3.1);
```

Because the double question mark (??) modifier suppresses printing of error messages and input lines and prevents the automatic variable _ERROR_ from being set to 1 when invalid data are read, the following two examples produce the same result:

- **y=input(x,?? 2.);**
- **y=input(x,? 2.); _error_=0;**

See Also

Functions:

"INPUTC Function" on page 826
"INPUTN Function" on page 827
"PUT Function" on page 1056
"PUTC Function" on page 1058
"PUTN Function" on page 1060

Statements:

"INPUT Statement" on page 1617

INPUTC Function

Enables you to specify a character informat at run time.

Category: Special

Syntax

INPUTC(*source*, *informat*<,*w*>)

Arguments

source
specifies a character constant, variable, or expression to which you want to apply the informat.

informat
is a character constant, variable, or expression that contains the character informat you want to apply to *source*.

w
is a numeric constant, variable, or expression that specifies a width to apply to the informat.

Interaction: If you specify a width here, it overrides any width specification in the informat.

Details

If the INPUTC function returns a value to a variable that has not yet been assigned a length, by default the variable length is determined by the length of the first argument.

Comparisons

The INPUTN function enables you to specify a numeric informat at run time. Using the INPUT function is faster because you specify the informat at compile time.

Examples

Example 1: Specifying Character Informats The PROC FORMAT step in this example creates a format, TYPEFMT., that formats the variable values 1, 2, and 3 with the name of one of the three informats that this step also creates. The informats store responses of "positive," "negative," and "neutral" as different words, depending on the type of question. After PROC FORMAT creates the format and informats, the DATA step creates a SAS data set from raw data consisting of a number identifying the type of question and a response. After reading a record, the DATA step uses the value of TYPE to create a variable, RESPINF, that contains the value of the appropriate informat for the current type of question. The DATA step also creates another variable, WORD, whose value is the appropriate word for a response. The INPUTC function assigns the value of WORD based on the type of question and the appropriate informat.

```
proc format;
   value typefmt 1='$groupx'
```

```
                  2='$groupy'
                  3='$groupz';
   invalue $groupx 'positive'='agree'
                   'negative'='disagree'
                   'neutral'='notsure';
   invalue $groupy 'positive'='accept'
                   'negative'='reject'
                   'neutral'='possible';

   invalue $groupz 'positive'='pass'
                   'negative'='fail'
                   'neutral'='retest';
run;

data answers;
   input type response $;
   respinformat = put(type, typefmt.);
   word = inputc(response, respinformat);
   datalines;
1 positive
1 negative
1 neutral
2 positive
2 negative
2 neutral
3 positive
3 negative
3 neutral
;
```

The value of WORD for the first observation is **agree**. The value of WORD for the last observation is **retest**.

See Also

Functions:

"INPUT Function" on page 823
"INPUTN Function" on page 827
"PUT Function" on page 1056
"PUTC Function" on page 1058
"PUTN Function" on page 1060

INPUTN Function

Enables you to specify a numeric informat at run time.

Category: Special

Syntax

INPUTN(*source*, *informat*<,*w*<,*d*>>)

Arguments

source
specifies a character constant, variable, or expression to which you want to apply the informat.

informat
is a character constant, variable or expression that contains the numeric informat you want to apply to *source*.

w
is a numeric constant, variable, or expression that specifies a width to apply to the informat.

Interaction: If you specify a width here, it overrides any width specification in the informat.

d
is a numeric constant, variable, or expression that specifies the number of decimal places to use.

Interaction: If you specify a number here, it overrides any decimal-place specification in the informat.

Comparisons

The INPUTC function enables you to specify a character informat at run time. Using the INPUT function is faster because you specify the informat at compile time.

Examples

Example 1: Specifying Numeric Informats The PROC FORMAT step in this example creates a format, READDATE., that formats the variable values 1 and 2 with the name of a SAS date informat. The DATA step creates a SAS data set from raw data originally from two different sources (indicated by the value of the variable SOURCE). Each source specified dates differently. After reading a record, the DATA step uses the value of SOURCE to create a variable, DATEINF, that contains the value of the appropriate informat for reading the date. The DATA step also creates a new variable, NEWDATE, whose value is a SAS date. The INPUTN function assigns the value of NEWDATE based on the source of the observation and the appropriate informat.

```
proc format;
   value readdate 1='date7.'
                  2='mmddyy8.';
run;

options yearcutoff=1920;

data fixdates (drop=start dateinformat);
   length jobdesc $12;
   input source id lname $ jobdesc $ start $;
   dateinformat=put(source, readdate.);
   newdate = inputn(start, dateinformat);
   datalines;
1 1604 Ziminski writer 09aug90
1 2010 Clavell editor 26jan95
2 1833 Rivera writer 10/25/92
```

```
2 2222 Barnes proofreader 3/26/98
;
```

See Also

Functions:

"INPUT Function" on page 823
"INPUTC Function" on page 826
"PUT Function" on page 1056
"PUTC Function" on page 1058
"PUTN Function" on page 1060

INT Function

Returns the integer value, fuzzed to avoid unexpected floating-point results.

Category: Truncation

Syntax

INT(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The INT function returns the integer portion of the argument (truncates the decimal portion). If the argument's value is within 1E-12 of an integer, the function results in that integer. If the value of *argument* is positive, the INT function has the same result as the FLOOR function. If the value of *argument* is negative, the INT function has the same result as the CEIL function.

Comparisons

Unlike the INTZ function, the INT function fuzzes the result. If the argument is within 1E-12 of an integer, the INT function fuzzes the result to be equal to that integer. The INTZ function does not fuzz the result. Therefore, with the INTZ function you might get unexpected results.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`var1=2.1;`
`x=int(var1);`
`put x;`
	`2`

`var2=-2.4;`
`y=int(var2);`
`put y;`
	`-2`

`a=int(1+1.e-11);` `put a;`	`1`
`b=int(-1.6);` `put b;`	`-1`

See Also

Functions:

"CEIL Function" on page 573
"FLOOR Function" on page 757
"INTZ Function" on page 861

INTCINDEX Function

Returns the cycle index when a date, time, or datetime interval and value are specified.

Category: Date and Time

Syntax

INTCINDEX(*interval*<<*multiple*.<*shift-index*>>>, *date-time-value*)

Arguments

interval
specifies a character constant, a variable, or an expression that contains an interval name such as WEEK, MONTH, or QTR. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Tip: If *interval* is a character constant, then enclose the value in quotation marks.

Requirement: Valid values for *interval* depend on whether *date-time-value* is a date, time, or datetime value.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<multiple.shift-index>
The three parts of the interval name are as follows:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval, then only multiperiod intervals can be shifted with the optional shift index. For example, because MONTH intervals shift by MONTH periods by default, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

date-time-value
specifies a date, time, or datetime value that represents a time period of a specified interval.

Details

The INTCINDEX function returns the index of the seasonal cycle when you specify an interval and a SAS date, time, or datetime value. For example, if the interval is MONTH, each observation in the data corresponds to a particular month. Monthly data is considered to be periodic for a one-year period. A year contains 12 months, so the number of intervals (months) in a seasonal cycle (year) is 12. WEEK is the seasonal cycle for an interval that is equal to DAY. Therefore, `intcindex('day','01SEP78'd);` returns a value of 35 because September 1, 1978, is the sixth day of the 35th week of the year. For more information about working with date and time intervals, see "Date and Time Intervals" on page 328.

The INTCINDEX function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For a list of these intervals, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Comparisons

The INTCINDEX function returns the cycle index, whereas the INTINDEX function returns the seasonal index.

In the example **`cycle_index = intcindex('day','04APR2005'd);`**, the INTCINDEX function returns the week of the year. In the example **`index = intindex('day','04APR2005'd);`**, the INTINDEX function returns the day of the week.

In the example **`cycle_index = intcindex('minute','01Sep78:00:00:00'dt);`**, the INTCINDEX function returns the hour of the day. In the example **`index = intindex('minute','01Sep78:00:00:00'dt);`**, the INTINDEX function returns the minute of the hour.

In the example **`intseas(intcycle('interval'));`**, the INTSEAS function returns the maximum number that could be returned by **`intcindex('interval',date);`**.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`cycle_index1 = intcindex('day', '01SEP05'd);` `put cycle_index1;`	`35`
`cycle_index2 = intcindex('dtqtr', '23MAY2005:05:03:01'dt);` `put cycle_index2;`	`1`
`cycle_index3 = intcindex('tenday', '13DEC2005' d);` `put cycle_index3;`	`1`
`cycle_index4 = intcindex('minute', '23:13:02't);` `put cycle_index4;`	`24`

`var1 = 'semimonth';`
`cycle_index5 = intcindex(var1, '05MAY2005:10:54:03'dt);`
`put cycle_index5;`
	`1`

See Also

Functions:

"INTINDEX Function" on page 845
"INTCYCLE Function" on page 836
"INTSEAS Function" on page 855

INTCK Function

Returns the count of the number of interval boundaries between two dates, two times, or two datetime values.

Category: Date and Time

Syntax

INTCK(*interval*<*multiple*><*.shift-index*>, *start-from*, *increment*, <*'alignment'*>)

INTCK(*custom-interval*, *start-from*, *increment*, <*'alignment'*>)

Arguments

interval
specifies a character constant, a variable, or an expression that contains an interval name. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Requirement: The type of interval (date, datetime, or time) must match the type of value in *from*.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<*multiple.shift-index*>
The three parts of the interval name are

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329for more information.

custom-interval
specifies a user-defined interval that is defined by a SAS data set. Each observation contains two variables, *begin* and *end*.

See: "Details" on page 834 for more information about custom intervals.

Requirement: You must use the INTERVALDS system option if you use the *custom-interval* variable.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, MONTH type intervals shift by MONTH subperiods by default. Thus, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329for more information.

start-from
specifies a SAS expression that represents the starting SAS date, time, or datetime value.

increment
specifies a SAS expression that represents the ending SAS date, time, or datetime value.

'alignment'
controls the position of SAS dates within the interval. You must enclose *alignment* in quotation marks. *Alignment* can be one of these values:

CONTINUOUS
specifies that continuous time is measured (the interval is shifted based on the starting date).

Alias: C or CONT

DISCRETE
specifies that discrete time is measured.

Alias: D or DISC

Details

Time Series Analysis: The Basics Times series analysis uses time intervals to analyze events. All values within the interval are interpreted as being equivalent. This means that the dates of January 1, 2005 and January 15, 2005 are equivalent when you specify a monthly interval. Both of these dates represent the interval that begins on January 1, 2005 and ends on January 31, 2005. You can use the date for the beginning of the interval (January 1, 2005) or the date for the end of the interval (January 31, 2005) to identify the interval. These dates represent all of the dates within the monthly interval.

In the example `intck('qtr','14JAN2005'd,'02SEP2005'd);`, the *start-from* argument ('14JAN2005'd) is equivalent to the first quarter of 2005. The *increment* argument ('02SEP2005'd) is equivalent to the third quarter of 2005. The interval count, that is, the number of times the beginning of an interval is reached in moving from the *start-from* argument to the *increment* argument is 2.

WEEK intervals are determined by the number of Sundays that occur between the *start-from* argument and the *increment* argument, and not by how many seven-day periods fall between the *start-from* argument and the *increment* argument.

Both the *multiple* and the *shift-index* arguments are optional and default to 1. For example, YEAR, YEAR1, YEAR.1, and YEAR1.1 are all equivalent ways of specifying ordinary calendar years.

For more information about working with date and time intervals, see "Date and Time Intervals" on page 328.

Custom Intervals

A custom interval is defined by a SAS data set. The data set must contain two variables, *begin* and *end*. Each observation represents one interval with the *begin* variable containing the start of the interval, and the *end* variable containing the end of the interval. The intervals must be listed in ascending order. You cannot have gaps between intervals, and intervals cannot overlap.

The SAS system option INTERVALDS is used to define custom intervals and associate interval data sets with new interval names. The following example shows how to specify the INTERVALDS system option:

```
options intervalds=(interval=libref.dataset-name);
```

where

interval
specifies the name of an interval. The value of *interval* is the data set that is named in *libref.dataset-name*.

libref.dataset-name
specifies the libref and data set name of the file that contains user-supplied holidays.

Retail Calendar Intervals

The retail industry often accounts for its data by dividing the yearly calendar into four 13-week periods, based on one of the following formats: 4-4-5, 4-5-4, or 5-4-4. The first, second, and third numbers specify the number of weeks in the first, second, and third month of each period, respectively. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`qtr=intck('qtr','10jan95'd,'01jul95'd);` `put qtr;`	`2`

`year=intck('year','31dec94'd,`
`'01jan95'd);`
`put year;`
	`1`

`year=intck('year','01jan94'd,`
`'31dec94'd);`
`put year;`
	`0`

`semi=intck('semiyear','01jan95'd,`
`'01jan98'd);`
`put semi;`
	`6`

`weekvar=intck('week2.2','01jan97'd,`
`'31mar97'd);`
`put weekvar;`
	`7`

`wdvar=intck('weekday7w','01jan97'd,`
`'01feb97'd);`
`put wdvar;`
	`26`

SAS Statements	Results

`y='year';`
`date1='1sep1991'd;`
`date2='1sep2001'd;`
`newyears=intck(y,date1,date2);`
`put newyears;`
	`10`

`y=trim('year     ');`
`date1='1sep1991'd + 300;`
`date2='1sep2001'd - 300;`
`newyears=intck(y,date1,date2);`
`put newyears;`
	`8`

In the second example, INTCK returns a value of 1 even though only one day has elapsed. This result is because the interval from December 31, 1994, to January 1, 1995, contains the starting point for the YEAR interval. However, in the third example, a value of 0 is returned even though 364 days have elapsed. This result is because the period between January 1, 1994, and December 31, 1994, does not contain the starting point for the interval.

In the fourth example, SAS returns a value of 6 because January 1, 1995, through January 1, 1998, contains six semiyearly intervals. (Note that if the ending date were December 31, 1997, SAS would count five intervals.) In the fifth example, SAS returns a value of 6 because there are six two-week intervals beginning on a first Monday during the period of January 1, 1997, through March 31, 1997. In the sixth example, SAS returns the value 26. That indicates that beginning with January 1, 1997, and counting only Saturdays as weekend days through February 1, 1997, the period contains 26 weekdays.

In the seventh example, the use of variables for the arguments is illustrated. The use of expressions for the arguments is illustrated in the last example.

See Also

Functions:

"INTNX Function" on page 848

System Options:

"INTERVALDS= System Option" on page 1930

INTCYCLE Function

Returns the date, time, or datetime interval at the next higher seasonal cycle when a date, time, or datetime interval is specified.

Category: Date and Time

Syntax

INTCYCLE(*interval* <<*multiple*.<*shift-index*>>>)

Arguments

interval
specifies a character constant, a variable, or an expression that contains an interval name such as WEEK, MONTH, or QTR. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<multiple.shift-index>
The three parts of the interval name are as follows:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, because MONTH type intervals shift by MONTH subperiods by default, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

Details

The INTCYCLE function returns the interval of the seasonal cycle, depending on a date, time, or datetime interval. For example, `INTCYCLE('MONTH');` returns the value YEAR because the months from January through December constitute a yearly cycle. `INTCYCLE('DAY');` returns the value WEEK because the days from Sunday through Saturday constitute a weekly cycle.

See "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for information about multipliers and shift indexes. See "Commonly Used Time Intervals" on page 329 for information about how intervals are calculated.

For more information about working with date and time intervals, see "Date and Time Intervals" on page 328.

The INTCYCLE function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

The following examples produce these results:

SAS Statements	Results
`cycle_year = intcycle('year');` `put cycle_year;`	`YEAR`
`cycle_quarter = intcycle('qtr');` `put cycle_quarter;`	`YEAR`
`cycle_month = intcycle('month');` `put cycle_month;`	`YEAR`
`cycle_day = intcycle('day');` `put cycle_day;`	`WEEK`

`var1 = 'second';`
`cycle_second = intcycle(var1);`
`put cycle_second;`
	`DTMINUTE`

See Also

Functions:

"INTSEAS Function" on page 855
"INTINDEX Function" on page 845
"INTCINDEX Function" on page 830

INTFIT Function

Returns a time interval that is aligned between two dates.

Category: Date and Time

Syntax

INTFIT(*argument-1*, *argument-2*, '*type*')

Arguments

argument
specifies a SAS expression that represents a SAS date or datetime value, or an observation.

Tip: Observation numbers are more likely to be used as arguments if date or datetime values are not available.

'type'
specifies whether the arguments are SAS date values, datetime values, or observations.

The following values for *type* are valid:

d	specifies that *argument-1* and *argument-2* are date values.
dt	specifies that *argument-1* and *argument-2* are datetime values.
obs	specifies that *argument-1* and *argument-2* are observations.

Details

The INTFIT function returns the most likely time interval based on two dates, datetime values, or observations that have been aligned within an interval. INTFIT assumes that the alignment value is SAME, which specifies that the date is aligned to the same calendar date with the corresponding interval increment. For more information about the *alignment* argument, see "INTNX Function" on page 848.

If the arguments that are used with INTFIT are observations, you can determine the cycle of an occurrence by using observation numbers. In the following example, the first two arguments of INTFIT are observation numbers, and the *type* argument is **obs**. If Jason used the gym the first time and the 25th time that a researcher recorded data, you could determine the interval by using the following statement: **interval=intfit(1,25,'obs');**. In this case, the value of interval is OBS24.2.

For information about time series, see the *SAS/ETS User's Guide*.

The INTFIT function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

Example 1: Finding Intervals That Are Aligned between Two Dates The following example shows the intervals that are aligned between two dates. The *type* argument in this example identifies the input as date values.

```
options pageno=1 nodate ls=80 ps=64;

data a;
   length interval $20;
   date1='01jan06'd;
   do i=1 to 25;
      date2=intnx('day', date1, i);
      interval=intfit(date1, date2, 'd');
      output;
   end;
   format date1 date2 date.;
run;

proc print data=a;
run;
```

Output 4.54 Interval Output from the INTFIT Function

The SAS System 1

Obs	interval	date1	i	date2
1	DAY	01JAN06	1	02JAN06
2	DAY2	01JAN06	2	03JAN06
3	DAY3.3	01JAN06	3	04JAN06
4	DAY4.3	01JAN06	4	05JAN06
5	DAY5.3	01JAN06	5	06JAN06
6	DAY6.3	01JAN06	6	07JAN06
7	WEEK	01JAN06	7	08JAN06
8	DAY8.3	01JAN06	8	09JAN06
9	DAY9.9	01JAN06	9	10JAN06
10	TENDAY	01JAN06	10	11JAN06
11	DAY11.6	01JAN06	11	12JAN06
12	DAY12.3	01JAN06	12	13JAN06
13	DAY13.7	01JAN06	13	14JAN06
14	WEEK2.8	01JAN06	14	15JAN06
15	SEMIMON	01JAN06	15	16JAN06
16	DAY16.3	01JAN06	16	17JAN06
17	DAY17.7	01JAN06	17	18JAN06
18	DAY18.9	01JAN06	18	19JAN06
19	DAY19.7	01JAN06	19	20JAN06
20	TENDAY2	01JAN06	20	21JAN06
21	WEEK3.8	01JAN06	21	22JAN06
22	DAY22.17	01JAN06	22	23JAN06
23	DAY23.13	01JAN06	23	24JAN06
24	DAY24.3	01JAN06	24	25JAN06
25	DAY25.3	01JAN06	25	26JAN06

The output shows that if the increment value is one day, then the result of the INTFIT function is DAY. If the increment value is two days, then the result of the INTFIT function is DAY2. If the increment value is three days, then the result is DAY3.3, with a shift index of 3. (If the two input dates are a Friday and a Monday, then the result is WEEKDAY.) If the increment value is seven days, then the result is WEEK.

Example 2: Finding Intervals That Are Aligned between Two Dates When the Dates Are Identified As Observations The following example shows the intervals that are aligned between two dates. The *type* argument in this example identifies the input as observations.

```
options pageno=1 nodate ls=80 ps=64;

data a;
   length interval $20;
   date1='01jan06'd;
   do i=1 to 25;
      date2=intnx('day', date1, i);
      interval=intfit(date1, date2, 'obs');
      output;
   end;
   format date1 date2 date.;
run;

proc print data=a;
run;
```

Output 4.55 Interval Output from the INTFIT Function When Dates Are Identified as Observations

The SAS System 1

Obs	interval	date1	i	date2
1	OBS	01JAN06	1	02JAN06
2	OBS2	01JAN06	2	03JAN06
3	OBS3.3	01JAN06	3	04JAN06
4	OBS4.3	01JAN06	4	05JAN06
5	OBS5.3	01JAN06	5	06JAN06
6	OBS6.3	01JAN06	6	07JAN06
7	OBS7.3	01JAN06	7	08JAN06
8	OBS8.3	01JAN06	8	09JAN06
9	OBS9.9	01JAN06	9	10JAN06
10	OBS10.3	01JAN06	10	11JAN06
11	OBS11.6	01JAN06	11	12JAN06
12	OBS12.3	01JAN06	12	13JAN06
13	OBS13.7	01JAN06	13	14JAN06
14	OBS14.3	01JAN06	14	15JAN06
15	OBS15.3	01JAN06	15	16JAN06
16	OBS16.3	01JAN06	16	17JAN06
17	OBS17.7	01JAN06	17	18JAN06
18	OBS18.9	01JAN06	18	19JAN06
19	OBS19.7	01JAN06	19	20JAN06
20	OBS20.3	01JAN06	20	21JAN06
21	OBS21.3	01JAN06	21	22JAN06
22	OBS22.17	01JAN06	22	23JAN06
23	OBS23.13	01JAN06	23	24JAN06
24	OBS24.3	01JAN06	24	25JAN06
25	OBS25.3	01JAN06	25	26JAN06

See Also

Functions:

"INTNX Function" on page 848
"INTCK Function" on page 833

INTFMT Function

Returns a recommended SAS format when a date, time, or datetime interval is specified.

Category: Date and Time

Syntax

INTFMT(*interval*<<*multiple*.<.*shift-index*>>>, *'size'*)

Arguments

interval
specifies a character constant, a variable, or an expression that contains an interval name such as WEEK, MONTH, or QTR. *Interval* can appear in uppercase or

lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<multiple.shift-index>

The three parts of the interval name are as follows:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, because MONTH type intervals shift by MONTH subperiods by default, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

'size'
specifies either LONG or SHORT. When a format includes a year value, LONG or L specifies a format that uses a four-digit year. SHORT or S specifies a format that uses a two-digit year.

Details

The INTFMT function returns a recommended format depending on a date, time, or datetime interval for displaying the time ID values that are associated with a time series of a given interval. The valid values of SIZE (LONG, L, SHORT, or S) specify whether to use a two-digit or a four-digit year when the format refers to a SAS date value. For more information about working with date and time intervals, see "Date and Time Intervals" on page 328.

The INTFMT function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For a list of these intervals, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`fmt1 = intfmt('qtr', 's');` `put fmt1;`	`YYQC4.`
`fmt2 = intfmt('qtr', 'l');` `put fmt2;`	`YYQC6.`
`fmt3 = intfmt('month', 'l');` `put fmt3;`	`MONYY7.`
`fmt4 = intfmt('week', 'short');` `put fmt4;`	`WEEKDATX15.`
`fmt5 = intfmt('week3.2', 'l');` `put fmt5;`	`WEEKDATX17.`
`fmt6 = intfmt('day', 'long');` `put fmt6;`	`DATE9.`

`var1 = 'month2';`
`fmt7 = intfmt(var1, 'long');`
`put fmt7;`
	`MONYY7.`

INTGET Function

Returns a time interval based on three date or datetime values.

Category: Date and Time

Syntax

INTGET(*date-1*, *date-2*, *date-3*)

Arguments

date
specifies a SAS date or datetime value.

Details

INTGET Function Intervals The INTGET function returns a time interval based on three date or datetime values. The function first determines all possible intervals between the first two dates, and then determines all possible intervals between the second and third dates. If the intervals are the same, INTGET returns that interval. If the intervals for the first and second dates differ, and the intervals for the second and third dates differ, INTGET compares the intervals. If one interval is a multiple of the other, then INTGET returns the smaller of the two intervals. Otherwise, INTGET

returns a missing value. INTGET works best with dates generated by the INTNX function whose alignment value is BEGIN.

In the following example, INTGET returns the interval DAY2:

```
interval=intget('01mar00'd, '03mar00'd, '09mar00'd);
```

The interval between the first and second dates is DAY2, because the number of days between March 1, 2000, and March 3, 2000, is two. The interval between the second and third dates is DAY6, because the number of days between March 3, 2000, and March 9, 2000, is six. DAY6 is a multiple of DAY2. INTGET returns the smaller of the two intervals.

In the following example, INTGET returns the interval MONTH4:

```
interval=intget('01jan00'd, '01may00'd, '01may01'd);
```

The interval between the first two dates is MONTH4, because the number of months between January 1, 2000, and May 1, 2000, is four. The interval between the second and third dates is YEAR. INTGET determines that YEAR is a multiple of MONTH4 (there are three MONTH4 intervals in YEAR), and returns the smaller of the two intervals.

In the following example, INTGET returns a missing value:

```
interval=intget('01Jan2006'd, '01Apr2006'd, '01Dec2006'd);
```

The interval between the first two dates is MONTH3, and the interval between the second and third dates is MONTH8. INTGET determines that MONTH8 is not a multiple of MONTH3, and returns a missing value.

The intervals that are returned are valid SAS intervals, including multiples of the intervals and shift intervals. Valid SAS intervals are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Note: If INTGET cannot determine a matching interval, then the function returns a missing value. No message is written to the SAS log. △

Retail Calendar Intervals

The INTGET function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

The following examples produce these results:

SAS Statements	Results
`interval=intget('01jan00'd,'01jan01'd,'01may01'd);` `put interval;`	**MONTH4**
`interval=intget('29feb80'd,'28feb82'd,'29feb84'd);` `put interval;`	**YEAR2.2**
`interval=intget('01feb80'd,'16feb80'd,'01mar80'd);` `put interval;`	**SEMIMONTH**
`interval=intget('2jan09'd,'2feb10'd,'2mar11'd);` `put interval;`	**MONTH13.4**

SAS Statements	Results
`interval=intget('10feb80'd,'19feb80'd,'28feb80'd);` `put interval;`	`DAY9.2`

`interval=intget('01apr2006:00:01:02'dt,`
`'01apr2006:00:02:02'dt,`
`'01apr2006:00:03:02'dt);`
`put interval;`
	`MINUTE`

See Also

Functions:

"INTFIT Function" on page 838
"INTNX Function" on page 848

INTINDEX Function

Returns the seasonal index when a date, time, or datetime interval and value are specified.

Category: Date and Time

Syntax

INTINDEX(*interval*<<*multiple*.<*shift-index*>>>, *date-value*)

Arguments

interval
specifies a character constant, a variable, or an expression that contains an interval name such as WEEK, MONTH, or QTR. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Tip: If *interval* is a character constant, then enclose the value in quotation marks.

Requirement: Valid values for *interval* depend on whether *date-value* is a date, time, or datetime value. For more information, see "Commonly Used Time Intervals" on page 329.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<*multiple.shift-index*>
The three parts of the interval name are as follows:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, because MONTH type intervals shift by MONTH subperiods by default, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

date-value
specifies a date, time, or datetime value that represents a time period of the given interval.

Details

INTINDEX Function Intervals The INTINDEX function returns the seasonal index when you supply an interval and an appropriate date, time, or datetime value. The seasonal index is a number that represents the position of the date, time, or datetime value in the seasonal cycle of the specified interval. For example, **`intindex('month', '01DEC2000'd);`** returns a value of 12 because there are 12 months in a yearly cycle and December is the 12th month of the year. In the following examples, INTINDEX returns the same value because both statements have values that occur in the first quarter of the year 2000: **`intindex('qtr', '01JAN2000'd);`** and **`intindex('qtr', '31MAR2000'd);`**. The statement **`intindex('day', '01DEC2000'd);`** returns a value of 6 because daily data is weekly periodic and December 1, 2000, is a Friday, the sixth day of the week.

How *Interval* and *Date-Time-Value* Are Related To correctly identify the seasonal index, the interval should agree with the date, time, or datetime value. For example, **`intindex('month', '01DEC2000'd);`** returns a value of 12 because there are 12 months in a yearly interval and December is the 12th month of the year. The MONTH interval requires a SAS date value. In the following example, **`intindex('day', '01DEC2000'd);`** returns a value of 6 because there are seven days in a weekly interval and December 1, 2000, is a Friday, the sixth day of the week. The DAY interval requires a SAS date value.

The example **`intindex('qtr', '01JAN2000:00:00:00'dt);`** results in an error because the QTR interval expects the date to be a SAS date value rather than a

datetime value. The example **intindex('dtmonth', '01DEC2000:00:00:00'dt);** returns a value of 12. The DTMONTH interval requires a datetime value.

For more information about working with date and time intervals, see "Date and Time Intervals" on page 328.

Retail Calendar Intervals

The INTINDEX function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Comparisons

The INTINDEX function returns the seasonal index whereas the INTCINDEX function returns the cycle index.

In the example **index = intindex('day', '04APR2005'd);**, the INTINDEX function returns the day of the week. In the example **cycle_index = intcindex('day', '04APR2005'd);**, the INTCINDEX function returns the week of the year.

In the example **index = intindex('minute','01Sep78:00:00:00'dt);**, the INTINDEX function returns the minute of the hour. In the example **cycle_index = intcindex('minute','01Sep78:00:00:00'dt);**, the INTCINDEX function returns the hour of the day.

In the example **intseas('interval');**, INTSEAS returns the maximum number that could be returned by **intindex('interval', date);**.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`interval1 = intindex('qtr', '14AUG2005'd);` `put interval1;`	`3`
`interval2 = intindex('dtqtr','23DEC2005:15:09:19'dt);` `put interval2;`	`4`
`interval3 = intindex('hour', '09:05:15't);` `put interval3;`	`10`
`interval4 = intindex('month', '26FEB2005'd);` `put interval4;`	`2`
`interval5 = intindex('dtmonth', '28MAY2005:05:15:00'dt);` `put interval5;`	`5`
`interval6 = intindex('week', '09SEP2005'd);` `put interval6;`	`36`
`interval7 = intindex('tenday', '16APR2005'd);` `put interval7;`	`11`

See Also

Function:

"INTCINDEX Function" on page 830

INTNX Function

Increments a date, time, or datetime value by a given time interval, and returns a date, time, or datetime value.

Category: Date and Time

Syntax

INTNX(*interval*<*multiple*><.*shift-index*>, *start-from*, *increment*<, '*alignment*'>)

INTNX(*custom-interval*, *start-from*, *increment* <, '*alignment*'>)

Arguments

interval
specifies a character constant, variable, or expression that contains a time interval such as WEEK, SEMIYEAR, QTR, or HOUR. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Requirement: The type of interval (date, datetime, or time) must match the type of value in *start-from* and *increment*.

See: "Commonly Used Time Intervals" on page 329 for a list of commonly used time intervals.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<*multiple.shift-index*>
Here are the three parts of the interval name:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, MONTH type intervals shift by MONTH subperiods by default; thus, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index because there are two MONTH intervals in each MONTH2 interval. The interval name MONTH2.2, for example, specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

start-from
specifies a SAS expression that represents a SAS date, time, or datetime value that identifies a starting point.

increment
specifies a negative, positive, or zero integer that represents the number of date, time, or datetime intervals. *Increment* is the number of intervals to shift the value of *start-from*.

'alignment'
controls the position of SAS dates within the interval. You must enclose *alignment* in quotation marks. *Alignment* can be one of these values:

BEGINNING
specifies that the returned date or datetime value is aligned to the beginning of the interval.

Alias: B

MIDDLE
specifies that the returned date or datetime value is aligned to the midpoint of the interval, which is the average of the beginning and ending alignment values.

Alias: M

END
specifies that the returned date or datetime value is aligned to the end of the interval.

Alias: E

SAME
specifies that the date that is returned has the same alignment as the input date.

Alias: S

Alias: SAMEDAY

See: "SAME Alignment" on page 850 for more information.

Default: BEGINNING

See: "Aligning SAS Date Output within Its Intervals" on page 850 for more information.

Details

The Basics

The INTNX function increments a date, time, or datetime value by intervals such as DAY, WEEK, QTR, and MINUTE, or a custom interval that you define. The increment is based on a starting date, time, or datetime value, and on the number of time intervals that you specify.

The INTNX function returns the SAS date value for the beginning date, time, or datetime value of the interval that you specify in the *start–from* argument. (To convert

the SAS date value to a calendar date, use any valid SAS date format, such as the DATE9. format.) The following example shows how to determine the date of the start of the week that is six weeks from the week of October 17, 2003.

```
x=intnx('week', '17oct03'd, 6);
put x date9.;
```

INTNX returns the value 23NOV2003.

For more information about working with date and time intervals, see "Date and Time Intervals" on page 328.

Aligning SAS Date Output within Its Intervals

SAS date values are typically aligned with the beginning of the time interval that is specified with the *interval* argument.

You can use the optional *alignment* argument to specify the alignment of the date that is returned. The values BEGINNING, MIDDLE, or END align the date to the beginning, middle, or end of the interval, respectively.

SAME Alignment If you use the SAME value of the *alignment* argument, then INTNX returns the same calendar date after computing the interval increment that you specified. The same calendar date is aligned based on the interval's shift period, not the interval. To view the valid shift periods, see the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Most of the values of the shift period are equal to their corresponding intervals. The exceptions are the intervals WEEK, WEEKDAY, QTR, SEMIYEAR, YEAR, and their DT counterparts. WEEK and WEEKDAY intervals have a shift period of DAYS; and QTR, SEMIYEAR, and YEAR intervals have a shift period of MONTH. When you use SAME alignment with YEAR, for example, the result is same-day alignment based on MONTH, the interval's shift period. The result is not aligned to the same day of the YEAR interval. If you specify a multiple interval, then the default shift interval is based on the interval, and not on the multiple interval.

When you use SAME alignment for QTR, SEMIYEAR, and YEAR intervals, the computed date is the same number of months from the beginning of the interval as the input date. The day of the month matches as closely as possible. Because not all months have the same number of days, it is not always possible to match the day of the month.

For more information about shift periods, see the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Alignment Intervals Use the SAME value of the *alignment* argument if you want to base the alignment of the computed date on the alignment of the input date:

```
intnx('week', '15mar2000'd, 1, 'same');         returns 22MAR2000
intnx('dtweek', '15mar2000:8:45'dt, 1, 'same'); returns 22MAR00:08:45:00
intnx('year', '15mar2000'd, 5, 'same');        returns 15MAR2005
```

Adjusting Dates The INTNX function automatically adjusts for the date if the date in the interval that is incremented does not exist. Here is an example:

```
intnx('month', '15mar2000'd, 5, 'same');  returns 15AUG2000
intnx('year', '29feb2000'd, 2, 'same');   returns 28FEB2002
intnx('month', '31aug2001'd, 1, 'same');  returns 30SEP2001
intnx('year', '01mar1999'd, 1, 'same');   returns 01MAR2000
                                          (the first day of the
                                           third month of the year)
```

```
intnx('year', '01mar1999'd, 1, 'same');   returns 29FEB2000 (the 60th day
                                          of the year)
```

In the example **intnx('year', '29feb2000'd, 2);**, the INTNX function returns the value 01JAN2002, which is the beginning of the year two years from the starting date (2000).

In the example **intnx('year', '29feb2000'd, 2, 'same');**, the INTNX function returns the value 28FEB2002. In this case, the starting date begins in the year 2000, the year is two years later (2002), the month is the same (February), and the date is the 28th, because that is the closest date to the 29th in February 2002.

Retail Calendar Intervals

The retail industry often accounts for its data by dividing the yearly calendar into four 13-week periods, based on one of the following formats: 4-4-5, 4-5-4, or 5-4-4. The first, second, and third numbers specify the number of weeks in the first, second, and third month of each period, respectively. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

Example 1: Examples of Using Intervals with the INTNX Function The following SAS statements produce these results.

SAS Statements	Results
`yr=intnx('year','05feb94'd,3);` `put yr / yr date7.;`	**13515** **01JAN97**
`x=intnx('month','05jan95'd,0);` `put x / x date7.;`	**12784** **01JAN95**
`next=intnx('semiyear','01jan97'd,1);` `put next / next date7.;`	**13696** **01JUL97**
`past=intnx('month2','01aug96'd,-1);` `put past / past date7.;`	**13270** **01MAY96**
`sm=intnx('semimonth2.2','01apr97'd,4);` `put sm / sm date7.;`	**13711** **16JUL97**

`x='month';`
`date='1jun1990'd;`
`nextmon=intnx(x,date,1);`
`put nextmon / nextmon date7.;`
	
	11139
	01JUL90

`x1='month     ';`
`x2=trim(x1);`
`date='1jun1990'd - 100;`
`nextmonth=intnx(x2,date,1);`
`put nextmonth / nextmonth date7.;`
	
	11017
	01MAR90

The following examples show the results of advancing a date by using the optional *alignment* argument.

SAS Statements	Results
`date1=intnx('month','01jan95'd,5,'beginning');`	12935
`put date1 / date1 date7.;`	01JUN95
`date2=intnx('month','01jan95'd,5,'middle');`	12949
`put date2 / date2 date7.;`	15JUN95
`date3=intnx('month','01jan95'd,5,'end');`	12964
`put date3 / date3 date7.;`	30JUN95
`date4=intnx('month','01jan95'd,5,'sameday');`	12935
`put date4 / date4 date7.;`	01JUN95
`date5=intnx('month','15mar2000'd,5,'same');`	14837
`put date5 / date5 date9.;`	15AUG2000

`interval='month';`
`date='1sep2001'd;`
`align='m';`
`date4=intnx(interval,date,2,align);`
	15294

`put date4 / date4 date7.;`	15NOV01

`x1='month     ';`
`x2=trim(x1);`
`date='1sep2001'd + 90;`
`date5=intnx(x2,date,2,'m');`
	15356

`put date5 / date5 date7.;`	16JAN02

Example 2: Example of Using Custom Intervals

The following example uses the *custom-interval* form of the INTNX function to increment a date, time, or datetime value by a given time interval.

```
options intervalds=(weekdaycust=dstest);

data dstest;
   format begin end date9.;
   begin='01jan2008'd; end='01jan2008'd; output;
   begin='02jan2008'd; end='02jan2008'd; output;
   begin='03jan2008'd; end='03jan2008'd; output;
   begin='04jan2008'd; end='06jan2008'd; output;
   begin='07jan2008'd; end='07jan2008'd; output;
   begin='08jan2008'd; end='08jan2008'd; output;
   begin='09jan2008'd; end='09jan2008'd; output;
   begin='10jan2008'd; end='10jan2008'd; output;
   begin='11jan2008'd; end='13jan2008'd; output;
   begin='14jan2008'd; end='14jan2008'd; output;
   begin='15jan2008'd; end='15jan2008'd; output;
run;

data _null_;
   format start date9. endcustom date9.;
   start='01jan2008'd;
   do i=0 to 9;
      endcustom=intnx('weekdaycust', start, i);
```

```
      put endcustom;
   end;
run;
```

SAS writes the following output to the log:

```
01JAN2008
02JAN2008
03JAN2008
04JAN2008
07JAN2008
08JAN2008
09JAN2008
10JAN2008
11JAN2008
14JAN2008
```

See Also

Functions:

"INTCK Function" on page 833

"INTSHIFT Function" on page 857

System Options:

"INTERVALDS= System Option" on page 1930

INTRR Function

Returns the internal rate of return as a fraction.

Category: Financial

Syntax

INTRR(*freq*,*c0*, *c1*,..., *cn*)

Arguments

freq
is numeric, the number of payments over a specified base period of time that is associated with the desired internal rate of return.

Range: *freq* > 0

Tip: The case *freq* = 0 is a flag to allow continuous compounding.

c0,c1, ... ,cn
are numeric, the optional cash payments.

Details

The INTRR function returns the internal rate of return over a specified base period of time for the set of cash payments *c0*, *c1*,..., *cn*. The time intervals between any two consecutive payments are assumed to be equal. The argument *freq* > 0 describes the number of payments that occur over the specified base period of time. The number of notes issued from each instance is limited.

The internal rate of return is the interest rate such that the sequence of payments has a 0 net present value. (See the NETPV function.) It is given by

$$r = \begin{cases} \frac{1}{x^{freq}} - 1 & freq > 0 \\ -log_e(x) & freq = 0 \end{cases}$$

where x is the real root of the polynomial.

$$\sum_{i=0}^{n} c_i x^i = 0$$

In the case of multiple roots, one real root is returned and a warning is issued concerning the non-uniqueness of the returned internal rate of return. Depending on the value of payments, a root for the equation does not always exist. In that case, a missing value is returned.

Missing values in the payments are treated as 0 values. When *freq* > 0, the computed rate of return is the effective rate over the specified base period. To compute a quarterly internal rate of return (the base period is three months) with monthly payments, set *freq* to 3.

If *freq* is 0, continuous compounding is assumed and the base period is the time interval between two consecutive payments. The computed internal rate of return is the nominal rate of return over the base period. To compute with continuous compounding and monthly payments, set *freq* to 0. The computed internal rate of return will be a monthly rate.

Comparisons

The IRR function is identical to INTRR, except for in the IRR function, the internal rate of return is a percentage.

Examples

For an initial outlay of $400 and expected payments of $100, $200, and $300 over the following three years, the annual internal rate of return can be expressed as

```
rate=intrr(1,-400,100,200,300);
```

The value returned is 0.19438.

See Also

Functions:

"IRR Function" on page 865

INTSEAS Function

Returns the length of the seasonal cycle when a date, time, or datetime interval is specified.

Category: Date and Time

Syntax

INTSEAS(*interval*<<*multiple*.<*shift-index*>>>)

Arguments

interval
specifies a character constant, a variable, or an expression that contains an interval name such as WEEK, MONTH, or QTR. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<*multiple.shift-index*>
The three parts of the interval name are as follows:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, because MONTH type intervals shift by MONTH subperiods by default, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

Details

The Basics The INTSEAS function returns the number of intervals in a seasonal cycle. For example, when the interval for a time series is described as monthly, then many procedures use the option INTERVAL=MONTH. Each observation in the data then corresponds to a particular month. Monthly data is considered to be periodic for a one-year period. A year contains 12 months, so the number of intervals (months) in a seasonal cycle (year) is 12.

Quarterly data is also considered to be periodic for a one-year period. A year contains four quarters, so the number of intervals in a seasonal cycle is four.

The periodicity is not always one year. For example, INTERVAL=DAY is considered to have a period of one week. Because there are seven days in a week, the number of intervals in the seasonal cycle is seven.

For more information about working with date and time intervals, see "Date and Time Intervals" on page 328.

Retail Calendar Intervals

The retail industry often accounts for its data by dividing the yearly calendar into four 13-week periods, based on one of the following formats: 4-4-5, 4-5-4, or 5-4-4. The first, second, and third numbers specify the number of weeks in the first, second, and third month of each period, respectively. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`cycle_years = intseas('year');` `put cycle_years;`	`1`
`cycle_smiyears = intseas('semiyear');` `put cycle_smiyears;`	`2`
`cycle_quarters = intseas('quarter');` `put cycle_quarters;`	`4`
`cycle_months = intseas('month');` `put cycle_months;`	`12`
`cycle_smimonths = intseas('semimonth');` `put cycle_smimonths;`	`24`
`cycle_tendays = intseas('tenday');` `put cycle_tendays;`	`36`
`cycle_weeks = intseas('week');` `put cycle_weeks;`	`52`
`cycle_wkdays = intseas('weekday');` `put cycle_wkdays;`	`5`
`cycle_hours = intseas('hour');` `put cycle_hours;`	`24`
`cycle_minutes = intseas('minute');` `put cycle_minutes;`	`60`

SAS Statements	Results
`cycle_month2 = intseas('month2.2');` `put cycle_month2;`	`6`
`cycle_week2 = intseas('week2');` `put cycle_week2;`	`26`

`var1 = 'month4.3';`
`cycle_var1 = intseas(var1);`
`put cycle_var1;`
	`3`

`cycle_day1 = intseas('day1');` `put cycle_day1;`	`7`

See Also

Function:

"INTCYCLE Function" on page 836

INTSHIFT Function

Returns the shift interval that corresponds to the base interval.

Category: Date and Time

Syntax

INTSHIFT(*interval* <<*multiple*.<*shift-index*>>>)

Arguments

interval
specifies a character constant, a variable, or an expression that contains a time interval such as WEEK, SEMIYEAR, QTR, or HOUR. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<multiple.shift-index>

The three parts of the interval name are as follows:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, the interval YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 would be an error because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, because MONTH type intervals shift by MONTH subperiods by default, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

Details

The INTSHIFT function returns the shift interval that corresponds to the base interval. For custom intervals, the value that is returned is the base custom interval name. INTSHIFT ignores multiples of the interval and interval shifts.

The INTSHIFT function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

The following examples produce these results:

SAS Statements	Results
`shift1 = intshift('year');` `put shift1;`	`MONTH`
`shift2 = intshift('dtyear');` `put shift2;`	`DTMONTH`
`shift3 = intshift('minute');` `put shift3;`	`DTMINUTE`

`interval = 'weekdays';`
`shift4 = intshift(interval);`
`put shift4;`
	`WEEKDAY`

`shift5 = intshift('weekday5.4');` `put shift5;`	`WEEKDAY`
`shift6 = intshift('qtr');` `put shift6;`	`MONTH`
`shift7 = intshift('dttenday');` `put shift7;`	`DTTENDAY`

INTTEST Function

Returns 1 if a time interval is valid, and returns 0 if a time interval is invalid.

Category: Date and Time

Syntax

INTTEST(*interval*<<*multiple*.<*shift-index*>>>)

Arguments

interval
specifies a character constant, variable, or expression that contains an interval name, such as WEEK, MONTH, or QTR. *Interval* can appear in uppercase or lowercase. The possible values of *interval* are listed in the "Intervals Used with Date and Time Functions" table in *SAS Language Reference: Concepts*.

Multipliers and shift indexes can be used with the basic interval names to construct more complex interval specifications. The general form of an interval name is as follows:

interval<multiple.shift-index>
Here are the three parts of the interval name:

interval
specifies the name of the basic interval type. For example, YEAR specifies yearly intervals.

multiple
specifies an optional multiplier that sets the interval equal to a multiple of the period of the basic interval type. For example, YEAR2 consists of two-year, or biennial, periods.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

shift-index
specifies an optional shift index that shifts the interval to start at a specified subperiod starting point. For example, YEAR.3 specifies yearly periods that are shifted to start on the first of March of each calendar year and to end in February of the following year.

Restriction: The shift index cannot be greater than the number of subperiods in the whole interval. For example, you could use YEAR2.24, but YEAR2.25 is invalid because there is no 25th month in a two-year interval.

Restriction: If the default shift period is the same as the interval type, then only multiperiod intervals can be shifted with the optional shift index. For example, because MONTH type intervals shift by MONTH subperiods by default, monthly intervals cannot be shifted with the shift index. However, bimonthly intervals can be shifted with the shift index, because there are two MONTH intervals in each MONTH2 interval. For example, the interval name MONTH2.2 specifies bimonthly periods starting on the first day of even-numbered months.

See: "Incrementing Dates and Times by Using Multipliers and by Shifting Intervals" on page 329 for more information.

Details

The INTTEST function checks for a valid interval name. This function is useful when checking for valid values of *multiple* and *shift-index*. For more information about multipliers and shift indexes, see "Multiunit Intervals" in *SAS Language Reference: Concepts*.

The INTTEST function can also be used with calendar intervals from the retail industry. These intervals are ISO 8601 compliant. For more information, see "Retail Calendar Intervals: ISO 8601 Compliant" in *SAS Language Reference: Concepts*.

Examples

In the following examples, SAS returns a value of 1 if the *interval* argument is valid, and 0 if the interval argument is invalid.

SAS Statements	Results
`test1 = inttest('month');` `put test1;`	`1`
`test2 = inttest('week6.13');` `put test2;`	`1`
`test3 = inttest('tenday');` `put test3;`	`1`
`test4 = inttest('twoweeks');` `put test4;`	`0`

`var1 = 'hour2.2';`
`test5 = inttest(var1);`
`put test5;`
	`1`

INTZ Function

Returns the integer portion of the argument, using zero fuzzing.

Category: Truncation

Syntax

INTZ (*argument*)

Arguments

argument
is a numeric constant, variable, or expression.

Details

The following rules apply:

- If the value of the argument is an exact integer, INTZ returns that integer.
- If the argument is positive and not an integer, INTZ returns the largest integer that is less than the argument.
- If the argument is negative and not an integer, INTZ returns the smallest integer that is greater than the argument.

Comparisons

Unlike the INT function, the INTZ function uses zero fuzzing. If the argument is within 1E-12 of an integer, the INT function fuzzes the result to be equal to that integer. The INTZ function does not fuzz the result. Therefore, with the INTZ function you might get unexpected results.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`var1=2.1;`
`a=intz(var1);`
`put a;`
	`2`

`var2=-2.4;`
`b=intz(var2);`
`put b;`
	`-2`

`var3=1+1.e-11;`
`c=intz(var3);`
`put c;`
	`1`

`f=intz(-1.6);` `put f;`	`-1`

See Also

Functions:

"CEIL Function" on page 573
"CEILZ Function" on page 575
"FLOOR Function" on page 757
"FLOORZ Function" on page 758

"INT Function" on page 829
"ROUND Function" on page 1099
"ROUNDZ Function" on page 1108

IORCMSG Function

Returns a formatted error message for _IORC_.

Category: SAS File I/O

Syntax

IORCMSG()

Details

If the IORCMSG function returns a value to a variable that has not yet been assigned a length, then by default the variable is assigned a length of 200.

The IORCMSG function returns the formatted error message that is associated with the current value of the automatic variable _IORC_. The _IORC_ variable is created when you use the MODIFY statement, or when you use the SET statement with the KEY= option. The value of the _IORC_ variable is internal and is meant to be read in conjunction with the SYSRC autocall macro. If you try to set _IORC_ to a specific value, you might get unexpected results.

Examples

In the following program, observations are either rewritten or added to the updated master file that contains bank accounts and current bank balance. The program queries the _IORC_ variable and returns a formatted error message if the _IORC_ value is unexpected.

```
libname bank 'SAS-library';

data bank.master(index=(AccountNum));
   infile 'external-file-1';
   format balance dollar8.;
   input @ 1 AccountNum $ 1--3 @ 5 balance 5--9;
run;

data bank.trans(index=(AccountNum));
   infile 'external-file-2';
   format deposit dollar8.;
   input @ 1 AccountNum $ 1--3 @ 5 deposit 5--9;
run;

data bank.master;
   set bank.trans;
```

```
   modify bank.master key=AccountNum;
   if (_IORC_ EQ %sysrc(_SOK)) then
      do;
         balance=balance+deposit;
         replace;
      end;
else
   if (_IORC_ = %sysrc(_DSENOM)) then
      do;
         balance=deposit;
         output;
         _error_=0;
      end;
else
   do;
      errmsg=IORCMSG();
      put 'Unknown error condition:'
      errmsg;
   end;
run;
```

IQR Function

Returns the interquartile range.

Category: Descriptive Statistics

Syntax

IQR(*value-1* <, *value-2...*>)

Arguments

value
specifies a numeric constant, variable, or expression for which the interquartile range is to be computed.

Details

If all arguments have missing values, the result is a missing value. Otherwise, the result is the interquartile range of the non-missing values. The formula for the interquartile range is the same as the one that is used in the UNIVARIATE procedure. For more information, see *Base SAS Procedures Guide*.

Examples

SAS Statements	Results
`iqr=iqr(2,4,1,3,999999);` `put iqr;`	`2`

See Also

Functions:

"MAD Function" on page 916

"PCTL Function" on page 985

IRR Function

Returns the internal rate of return as a percentage.

Category: Financial

Syntax

IRR(*freq,c0,c1,...,cn*)

Arguments

freq
is numeric, the number of payments over a specified base period of time that is associated with the desired internal rate of return.

Range: *freq* > 0.

Tip: The case *freq* = 0 is a flag to allow continuous compounding.

c0,c1,...,cn
are numeric, the optional cash payments.

Details

The IRR function returns the internal rate of return over a specified base period of time for the set of cash payments *c0*, *c1*,...,*cn*. The time intervals between any two consecutive payments are assumed to be equal. The argument *freq* > 0 describes the number of payments that occur over the specified base period of time. The number of notes issued from each instance is limited.

Comparisons

The IRR function is identical to INTRR, except for in the IRR function, the internal rate of return is a percentage.

See Also

Functions:

"INTRR Function" on page 853

JBESSEL Function

Returns the value of the Bessel function.

Category: Mathematical

Syntax

JBESSEL(*nu*,*x*)

Arguments

nu

specifies a numeric constant, variable, or expression.

Range: $nu \geq 0$

x

specifies a numeric constant, variable, or expression.

Range: $x \geq 0$

Details

The JBESSEL function returns the value of the Bessel function of order *nu* evaluated at *x* (For more information, see Abramowitz and Stegun 1964; Amos, Daniel, and Weston 1977).

Examples

SAS Statements	Results
`x=jbessel(2,2);`	`0.3528340286`

JULDATE Function

Returns the Julian date from a SAS date value.

Category: Date and Time

Syntax

JULDATE(*date*)

Arguments

date
specifies a SAS date value.

Details

A SAS date value is a number that represents the number of days from January 1, 1960 to a specific date. The JULDATE function converts a SAS date value to a Julian date. If *date* falls within the 100-year span defined by the system option YEARCUTOFF=, the result has three, four, or five digits. In a five digit result, the first two digits represent the year, and the next three digits represent the day of the year (1 to 365, or 1 to 366 for leap years). Because leading zeros are dropped from the result, the year portion of a Julian date might be omitted (for years ending in 00), or it might have only one digit (for years ending 01–09). Otherwise, the result has seven digits: the first four digits represent the year, and the next three digits represent the day of the year. For example, if YEARCUTOFF=1920, JULDATE would return 97001 for January 1, 1997, and return 1878365 for December 31, 1878.

Comparisons

The function JULDATE7 is similar to JULDATE except that JULDATE7 always returns a four digit year. Thus JULDATE7 is year 2000 compliant because it eliminates the need to consider the implications of a two digit year.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`julian=juldate('31dec99'd);`	`99365`
`julian=juldate('01jan2099'd);`	`2099001`

See Also

Function:

"DATEJUL Function" on page 635
"JULDATE7 Function" on page 868

System Option:

"YEARCUTOFF= System Option" on page 2058

JULDATE7 Function

Returns a seven-digit Julian date from a SAS date value.

Category: Date and Time

Syntax

JULDATE7(*date*)

Arguments

date
specifies a SAS date value.

Details

The JULDATE7 function returns a seven digit Julian date from a SAS date value. The first four digits represent the year, and the next three digits represent the day of the year.

Comparisons

The function JULDATE7 is similar to JULDATE except that JULDATE7 always returns a four digit year. Thus JULDATE7 is year 2000 compliant because it eliminates the need to consider the implications of a two digit year.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`julian=juldate7('31dec96'd);`	`1996366`
`julian=juldate7('01jan2099'd);`	`2099001`

See Also

Function:
"JULDATE Function" on page 866

KURTOSIS Function

Returns the kurtosis.

Category: Descriptive Statistics

Syntax

KURTOSIS(*argument-1,argument-2,argument-3,argument-4*<,…,*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

At least four non-missing arguments are required. Otherwise, the function returns a missing value. If all non-missing arguments have equal values, the kurtosis is mathematically undefined. The KURTOSIS function returns a missing value and sets _ERROR_ equal to 1.

The argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=kurtosis(5,9,3,6);`	`0.928`
`x2=kurtosis(5,8,9,6,.);`	`-3.3`
`x3=kurtosis(8,9,6,1);`	`1.5`
`x4=kurtosis(8,1,6,1);`	`-4.483379501`
`x5=kurtosis(of x1-x4);`	`-5.065692754`

LAG Function

Returns values from a queue.

Category: Special

Syntax

LAG<*n*>(*argument*)

Arguments

n
specifies the number of lagged values.

argument
specifies a numeric or character constant, variable, or expression.

Details

The Basics If the LAG function returns a value to a character variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

The LAG functions, LAG1, LAG2, ..., LAG*n* return values from a queue. LAG1 can also be written as LAG. A LAG*n* function stores a value in a queue and returns a value stored previously in that queue. Each occurrence of a LAG*n* function in a program generates its own queue of values.

The queue for each occurrence of LAG*n* is initialized with *n* missing values, where *n* is the length of the queue (for example, a LAG2 queue is initialized with two missing values). When an occurrence of LAG*n* is executed, the value at the top of its queue is removed and returned, the remaining values are shifted upwards, and the new value of the argument is placed at the bottom of the queue. Hence, missing values are returned for the first *n* executions of each occurrence of LAG*n*, after which the lagged values of the argument begin to appear.

Note: Storing values at the bottom of the queue and returning values from the top of the queue occurs only when the function is executed. An occurrence of the LAG*n* function that is executed conditionally will store and return values only from the observations for which the condition is satisfied. △

If the argument of LAG*n* is an array name, a separate queue is maintained for each variable in the array.

Memory Limit for the LAG Function When the LAG function is compiled, SAS allocates memory in a queue to hold the values of the variable that is listed in the LAG function. For example, if the variable in function LAG100(x) is numeric with a length of 8 bytes, then the memory that is needed is 8 times 100, or 800 bytes. Therefore, the memory limit for the LAG function is based on the memory that SAS allocates, which varies with different operating environments.

Examples

Example 1: Generating Two Lagged Values The following program generates two lagged values for each observation.

```
options pagesize= linesize= pageno=1 nodate;

data one;
   input x @@;
   y=lag1(x);
   z=lag2(x);
   datalines;
1 2 3 4 5 6
;

proc print data=one;
   title 'LAG Output';
run;
```

Output 4.56 Output from Generating Two Lagged Values

LAG Output 1

Obs	x	y	z
1	1	.	.
2	2	1	.
3	3	2	1
4	4	3	2
5	5	4	3
6	6	5	4

LAG1 returns one missing value and the values of X (lagged once). LAG2 returns two missing values and the values of X(lagged twice).

Example 2: Generating Multiple Lagged Values in BY-Groups The following example shows how to generate up to three lagged values within each BY group.

```
/*************************************************************************/
/* This program generates up to three lagged values.  By increasing the  */
/* size of the array and the number of assignment statements that use    */
/* the LAGn functions, you can generate as many lagged values as needed. */
/*************************************************************************/
options pageno=1 ls=80 ps=64 nodate;

/* Create starting data. */

data old;
  input start end;
datalines;
1 1
1 2
1 3
1 4
1 5
1 6
1 7
2 1
2 2
3 1
3 2
3 3
3 4
3 5
;

data new(drop=i count);
  set old;
  by start;

  /* Create and assign values to three new variables.  Use ENDLAG1-    */
  /* ENDLAG3 to store lagged values of END, from the most recent to the */
  /* third preceding value.                                            */
  array x(*) endlag1-endlag3;
  endlag1=lag1(end);
  endlag2=lag2(end);
  endlag3=lag3(end);

  /* Reset COUNT at the start of each new BY-Group */
```

```
  if first.start then count=1;

  /* On each iteration, set to missing array elements   */
  /* that have not yet received a lagged value for the  */
  /* current BY-Group.  Increase count by 1.            */
  do i=count to dim(x);
    x(i)=.;
  end;
  count + 1;
run;

proc print;
run;
```

Output 4.57 Output from Generating Three Lagged Values

The SAS System 1

Obs	start	end	endlag1	endlag2	endlag3
1	1	1	.	.	.
2	1	2	1	.	.
3	1	3	2	1	.
4	1	4	3	2	1
5	1	5	4	3	2
6	1	6	5	4	3
7	1	7	6	5	4
8	2	1	.	.	.
9	2	2	1	.	.
10	3	1	.	.	.
11	3	2	1	.	.
12	3	3	2	1	.
13	3	4	3	2	1
14	3	5	4	3	2

Example 3: Computing the Moving Average of a Variable The following is an example that computes the moving average of a variable in a data set.

```
/* Title: Compute the moving average of a variable
   Goal: Compute the moving average of a variable through the entire data set,
         of the last n observations and of the last n observations within a
         BY-group.
   Input:
*/
options pageno=1 ls=80 ps=64 fullstimer nodate;

data x;
do x=1 to 10;
  output;
  end;
run;

/* Compute the moving average of the entire data set. */
```

```
data avg;
retain s 0;
set x;
s=s+x;
a=s/_n_;
run;
proc print;
run;
/* Compute the moving average of the last 5 observations. */
%let n = 5;
data avg (drop=s);
retain s;
set x;
s = sum (s, x, -lag&n(x)) ;
a = s / min(_n_, &n);
run;
proc print;
run;
/* Compute the moving average within a BY-group of last n observations.
  For the first n-1 observations within the BY-group, the moving average
  is set to missing. */
data ds1;
do patient='A','B','C';
 do month=1 to 7;
  num=int(ranuni(0)*10);
  output;
 end;
end;
run;
proc sort;
by patient;
%let n = 4;
data ds2;
set ds1;
by patient;
retain num_sum 0;
if first.patient then do;
  count=0;
  num_sum=0;
end;
count+1;
last&n=lag&n(num);
if count gt &n then num_sum=sum(num_sum,num,-last&n);
else num_sum=sum(num_sum,num);
if count ge &n then mov_aver=num_sum/&n;
else mov_aver=.;
run;

proc print;
run;
```

Output 4.58 Output from Computing the Moving Average of a Variable

The SAS System 1

Obs	s	x	a
1	1	1	1.0
2	3	2	1.5
3	6	3	2.0
4	10	4	2.5
5	15	5	3.0
6	21	6	3.5
7	28	7	4.0
8	36	8	4.5
9	45	9	5.0
10	55	10	5.5

The SAS System 2

Obs	x	a
1	1	1.0
2	2	1.5
3	3	2.0
4	4	2.5
5	5	3.0
6	6	4.0
7	7	5.0
8	8	6.0
9	9	7.0
10	10	8.0

The SAS System 3

Obs	patient	month	num	num_sum	count	last4	mov_aver
1	A	1	4	4	1	.	.
2	A	2	9	13	2	.	.
3	A	3	2	15	3	.	.
4	A	4	9	24	4	.	6.00
5	A	5	3	23	5	4	5.75
6	A	6	6	20	6	9	5.00
7	A	7	8	26	7	2	6.50
8	B	1	7	7	1	9	.
9	B	2	5	12	2	3	.
10	B	3	5	17	3	6	.
11	B	4	0	17	4	8	4.25
12	B	5	5	15	5	7	3.75
13	B	6	9	19	6	5	4.75
14	B	7	0	14	7	5	3.50
15	C	1	4	4	1	0	.
16	C	2	2	6	2	5	.
17	C	3	0	6	3	9	.
18	C	4	1	7	4	0	1.75
19	C	5	2	5	5	4	1.25
20	C	6	9	12	6	2	3.00
21	C	7	5	17	7	0	4.25

Example 4: Generating a Fibonacci Sequence of Numbers The following example generates a Fibonacci sequence of numbers. You start with 0 and 1, and then add the two previous Fibonacci numbers to generate the next Fibonacci number.

```
data _null_;
   put 'Fibonacci Sequence';
   n=1;
   f=1;
   put n= f=;
   do n=2 to 10;
      f=sum(f,lag(f));
      put n= f=;
   end;
run;
```

SAS writes the following output to the log:

```
Fibonacci Sequence
n=1 f=1
n=2 f=1
n=3 f=2
n=4 f=3
n=5 f=5
n=6 f=8
n=7 f=13
n=8 f=21
n=9 f=34
n=10 f=55
```

Example 5: Using Expressions for the LAG Function Argument The following program uses an expression for the value of *argument* and creates a data set that contains the values for X, Y, and Z. LAG dequeues the previous values of the expression and enqueues the current value.

```
options nodate pageno=1 ls=80 ps=85;

data one;
   input X @@;
   Y=lag1(x+10);
   Z=lag2(x);
   datalines;
1 2 3 4 5 6
;
proc print;
   title 'Lag Output: Using an Expression';
run;
```

Output 4.59 Output from the LAG Function: Using an Expression

```
                    Lag Output: Using an Expression                    1

                          Obs    X     Y     Z

                           1     1     .     .
                           2     2    11     .
                           3     3    12     1
                           4     4    13     2
                           5     5    14     3
                           6     6    15     4
```

See Also

Function:

"DIF Function" on page 653

LARGEST Function

Returns the *k*th largest non-missing value.

Category: Descriptive Statistics

Syntax

LARGEST (*k*, *value-1*<, *value-2 ...*>)

Arguments

k
is a numeric constant, variable, or expression that specifies which value to return.

value
specifies the value of a numeric constant, variable, or expression to be processed.

Details

If *k* is missing, less than zero, or greater than the number of values, the result is a missing value and _ERROR_ is set to 1. Otherwise, if *k* is greater than the number of non-missing values, the result is a missing value but _ERROR_ is not set to 1.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`k=1;`
`largest1=largest(k, 456, 789, .Q, 123);`
`put largest1;`
	`789`

`k=2;`
`largest2=largest(k, 456, 789, .Q, 123);`
`put largest2;`
	`456`

`k=3;`
`largest3=largest(k, 456, 789, .Q, 123);`
`put largest3;`
	`123`

`k=4;`
`largest4=largest(k, 456, 789, .Q, 123);`
`put largest4;`
	`.`

See Also

Functions:

"ORDINAL Function" on page 982
"PCTL Function" on page 985
"SMALLEST Function" on page 1128

LBOUND Function

Returns the lower bound of an array.

Category: Array

Syntax

LBOUND<*n*>(*array-name*)

LBOUND(*array-name*,*bound-n*)

Arguments

n
is an integer constant that specifies the dimension for which you want to know the lower bound. If no *n* value is specified, the LBOUND function returns the lower bound of the first dimension of the array.

array-name
is the name of an array that was defined previously in the same DATA step.

bound-n
is a numeric constant, variable, or expression that specifies the dimension for which you want to know the lower bound. Use *bound-n* only if *n* is not specified.

Details

The LBOUND function returns the lower bound of a one-dimensional array or the lower bound of a specified dimension of a multidimensional array. Use LBOUND in array processing to avoid changing the lower bound of an iterative DO group each time you change the bounds of the array. LBOUND and HBOUND can be used together to return the values of the lower and upper bounds of an array dimension.

Examples

Example 1: One-dimensional Array In this example, LBOUND returns the lower bound of the dimension, a value of 2. SAS repeats the statements in the DO loop five times.

```
array big{2:6} weight sex height state city;
do i=lbound(big) to hbound(big);
   ...more SAS statements...;
end;
```

Example 2: Multidimensional Array This example shows two ways of specifying the LBOUND function for multidimensional arrays. Both methods return the same value for LBOUND, as shown in the table that follows the SAS code example.

```
array mult{2:6,4:13,2} mult1-mult100;
```

Syntax	Alternative Syntax	Value
LBOUND(MULT)	LBOUND(MULT,1)	2
LBOUND2(MULT)	LBOUND(MULT,2)	4
LBOUND3(MULT)	LBOUND(MULT,3)	1

See Also

Functions:

"DIM Function" on page 655

"HBOUND Function" on page 802

Statements:

"ARRAY Statement" on page 1440

"Array Reference Statement" on page 1445

"Array Processing" in *SAS Language Reference: Concepts*

LCM Function

Returns the least common multiple.

Category: Mathematical

Syntax

LCM(*x1, x2, x3, …, xn*)

Arguments

x
specifies a numeric constant, variable, or expression that has an integer value.

Details

The LCM (least common multiple) function returns the smallest multiple that is exactly divisible by every member of a set of numbers. For example, the least common multiple of 12 and 18 is 36.

If any of the arguments are missing, then the returned value is a missing value.

Examples

The following example returns the smallest multiple that is exactly divisible by the integers 10 and 15.

```
data _null_;
   x=lcm(10,15);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=30
```

See Also

Functions:

"GCD Function" on page 785

LCOMB Function

Computes the logarithm of the COMB function, which is the logarithm of the number of combinations of *n* objects taken *r* at a time.

Category: Combinatorial

Syntax

LCOMB(*n*,*r*)

Arguments

n
is a non-negative integer that represents the total number of elements from which the sample is chosen.

r
is a non-negative integer that represents the number of chosen elements.

Restriction: $r \leq n$

Comparisons

The LCOMB function computes the logarithm of the COMB function.

Examples

The following statements produce these results:

SAS Statements	Results
x=lcomb(5000,500);	
put x;	1621.4411361
y=lcomb(100,10);	
put y;	30.482323362

See Also

Functions:

"COMB Function" on page 590

LEFT Function

Left-aligns a character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KLEFT in *SAS National Language Support (NLS): Reference Guide*. See "DBCS Compatibility" on page 882.

Syntax

LEFT(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

The Basics In a DATA step, if the LEFT function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the argument.

LEFT returns an argument with leading blanks moved to the end of the value. The argument's length does not change.

DBCS Compatibility The LEFT function left-aligns a character string. You can use the LEFT function in most cases. If an application can be executed in an ASCII environment, or if the application does not manipulate character strings, then using the LEFT function rather than the KLEFT function.

Examples

SAS Statements	Results
	`----+----1----+`

`a='  DUE DATE';`
`b=left(a);`
`put b;`
	`DUE DATE`

See Also

Functions:

"COMPRESS Function" on page 604
"RIGHT Function" on page 1097
"STRIP Function" on page 1138
"TRIM Function" on page 1173

LENGTH Function

Returns the length of a non-blank character string, excluding trailing blanks, and returns 1 for a blank character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KLENGTH in *SAS National Language Support (NLS): Reference Guide*.

Tip: The LENGTH function returns a length in bytes, while the KLENGTH function returns a length in a character based unit.

Syntax

LENGTH(*string*)

Arguments

string
specifies a character constant, variable, or expression.

Details

The LENGTH function returns an integer that represents the position of the rightmost non-blank character in *string*. If the value of *string* is blank, LENGTH returns a value of 1. If *string* is a numeric constant, variable, or expression (either initialized or uninitialized), SAS automatically converts the numeric value to a right-justified character string by using the BEST12. format. In this case, LENGTH returns a value of 12 and writes a note in the SAS log stating that the numeric values have been converted to character values.

Comparisons

- The LENGTH and LENGTHN functions return the same value for non-blank character strings. LENGTH returns a value of 1 for blank character strings, whereas LENGTHN returns a value of 0.
- The LENGTH function returns the length of a character string, excluding trailing blanks, whereas the LENGTHC function returns the length of a character string, including trailing blanks.
- The LENGTH function returns the length of a character string, excluding trailing blanks, whereas the LENGTHM function returns the amount of memory in bytes that is allocated for a character string.

Examples

SAS Statements	Results
`len=length('ABCDEF');` `put len;`	`6`
`len2=length(' ');` `put len2;`	`1`

See Also

Functions:

"LENGTHC Function" on page 884

"LENGTHM Function" on page 885

"LENGTHN Function" on page 887

LENGTHC Function

Returns the length of a character string, including trailing blanks.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

LENGTHC(*string*)

Arguments

string
specifies a character constant, variable, or expression.

Details

The LENGTHC function returns the number of characters, both blanks and non-blanks, in *string*. If *string* is a numeric constant, variable or expression (either initialized or uninitialized), SAS automatically converts the numeric value to a right-justified character string by using the BEST12. format. In this case, LENGTHC returns a value of 12 and writes a note in the SAS log stating that the numeric values have been converted to character values.

Comparisons

- The LENGTHC function returns the length of a character string, including trailing blanks, whereas the LENGTH and LENGTHN functions return the length of a character string, excluding trailing blanks. LENGTHC always returns a value that is greater than or equal to the value of LENGTHN.
- The LENGTHC function returns the length of a character string, including trailing blanks, whereas the LENGTHM function returns the amount of memory in bytes that is allocated for a character string. For fixed-length character strings, LENGTHC and LENGTHM always return the same value. For varying-length character strings, LENGTHC always returns a value that is less than or equal to the value returned by LENGTHM.

Examples

The following SAS statements produce these results:

SAS Statements	Results
`x=lengthc('variable with trailing blanks   ');` `put x;`	`32`

`length fixed $35;`
`fixed='variable with trailing blanks   ';`
`x=lengthc(fixed);`
`put x;`
	`35`

See Also

Functions:

"LENGTH Function" on page 883

"LENGTHM Function" on page 885

"LENGTHN Function" on page 887

LENGTHM Function

Returns the amount of memory (in bytes) that is allocated for a character string.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

LENGTHM(*string*)

Arguments

string
specifies a character constant, variable, or expression.

Details

The LENGTHM function returns an integer that represents the amount of memory in bytes that is allocated for *string*. If *string* is a numeric constant, variable, or expression (either initialized or uninitialized), SAS automatically converts the numeric value to a right-justified character string by using the BEST12. format. In this case, LENGTHM returns a value of 12 and writes a note in the SAS log stating that the numeric values have been converted to character values.

Comparisons

The LENGTHM function returns the amount of memory in bytes that is allocated for a character string, whereas the LENGTH, LENGTHC, and LENGTHN functions return the length of a character string. LENGTHM always returns a value that is greater than or equal to the values that are returned by LENGTH, LENGTHC, and LENGTHN.

Examples

Example 1: Determining the Amount of Allocated Memory for a Character Expression This example determines the amount of memory (in bytes) that is allocated for a buffer that stores intermediate results in a character expression. Because SAS does not know how long the value of the expression CAT(x,y) will be, SAS allocates memory for values up to 32767 bytes long.

```
data _null_;
   x='x';
   y='y';
   lc=lengthc(cat(x,y));
   lm=lengthm(cat(x,y));
   put lc= lm=;
run;
```

SAS writes the following output to the log:

```
lc=2 lm=32767
```

Example 2: Determining the Amount of Allocated Memory for a Variable from an External File This example determines the amount of memory (in bytes) that is allocated to a variable that is input into a SAS file from an external file.

```
data _null_;
   file 'test.txt';
   put 'trailing blanks   ';
run;

data test;
   infile 'test.txt';
   input;
   x=lengthm(_infile_);
   put x;
run;
```

The following line is written to the SAS log:

```
256
```

See Also

Functions:

"LENGTH Function" on page 883
"LENGTHC Function" on page 884
"LENGTHN Function" on page 887

LENGTHN Function

Returns the length of a character string, excluding trailing blanks.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

LENGTHN(*string*)

Arguments

string
specifies a character constant, variable, or expression.

Details

The LENGTHN function returns an integer that represents the position of the rightmost non-blank character in *string*. If the value of *string* is blank, LENGTHN returns a value of 0. If *string* is a numeric constant, variable, or expression (either initialized or uninitialized), SAS automatically converts the numeric value to a right-justified character string by using the BEST12. format. In this case, LENGTHN returns a value of 12 and writes a note in the SAS log stating that the numeric values have been converted to character values.

Comparisons

- The LENGTHN and LENGTH functions return the same value for non-blank character strings. LENGTHN returns a value of 0 for blank character strings, whereas LENGTH returns a value of 1.
- The LENGTHN function returns the length of a character string, excluding trailing blanks, whereas the LENGTHC function returns the length of a character string, including trailing blanks. LENGTHN always returns a value that is less than or equal to the value returned by LENGTHC.
- The LENGTHN function returns the length of a character string, excluding trailing blanks, whereas the LENGTHM function returns the amount of memory in bytes that is allocated for a character string. LENGTHN always returns a value that is less than or equal to the value returned by LENGTHM.

Examples

SAS Statements	Results
`len=lengthn('ABCDEF');` `put len;`	`6`
`len2=lengthn(' ');` `put len2;`	`0`

See Also

Functions:

"LENGTH Function" on page 883
"LENGTHC Function" on page 884
"LENGTHM Function" on page 885

LEXCOMB Function

Generates all distinct combinations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.

Category: Combinatorial

Restriction: The LEXCOMB function cannot be executed when you use the %SYSFUNC macro.

Syntax

LEXCOMB(*count*, *k*, *variable-1*, ..., *variable-n*)

Arguments

count
specifies an integer variable that is assigned values ffrom 1 to the number of combinations in a loop.

k
is a constant, variable, or expression between 1 and *n*, inclusive, that specifies the number of items in each combination.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before you execute the LEXCOMB function.

Tip: After executing the LEXCOMB function, the first *k* variables contain the values in one combination.

Details

The Basics Use the LEXCOMB function in a loop where the first argument to LEXCOMB takes each integral value from 1 to the number of distinct combinations of the non-missing values of the variables. In each execution of LEXCOMB within this loop, *k* should have the same value.

Number of Combinations When all of the variables have non-missing, unequal values, then the number of combinations is COMB(*n*,*k*). If the number of variables that have missing values is m, and all the non-missing values are unequal, then LEXCOMB produces COMB(*n-m*,*k*) combinations because the missing values are omitted from the combinations.

When some of the variables have equal values, the exact number of combinations is difficult to compute, but COMB(*n*,*k*) provides an upper bound. You do not need to compute the exact number of combinations, provided that your program leaves the loop when LEXCOMB returns a value that is less than zero.

LEXCOMB Processing On the first execution of the LEXCOMB function, the following actions occur:

- The argument types and lengths are checked for consistency.
- The m missing values are assigned to the last m arguments.
- The *n-m* non-missing values are assigned in ascending order to the first *n-m* arguments following *count*.
- LEXCOMB returns 1.

On subsequent executions, up to and including the last combination, the following actions occur:

- The next distinct combination of the non-missing values is generated in lexicographic order.
- If *variable-1* through *variable-i* did not change, but *variable-j* did change, where *j*=*i*+1, then LEXCOMB returns *j*.

If you execute the LEXCOMB function after generating all the distinct combinations, then LEXCOMB returns –1.

If you execute the LEXCOMB function with the first argument out of sequence, then the results are not useful. In particular, if you initialize the variables and then immediately execute the LEXCOMB function with a first argument of *j*, you will not get the j^{th} combination (except when *j* is 1). To get the j^{th} combination, you must execute the LEXCOMB function *j* times, with the first argument taking values from 1 through *j* in that exact order.

Comparisons

The LEXCOMB function generates all distinct combinations of the non-missing values of n variables taken *k* at a time in lexicographic order. The ALLCOMB function generates all combinations of the values of k variables taken *k* at a time in a minimum change order.

Examples

Example 1: Generating Distinct Combinations in Lexicographic Order The following example uses the LEXCOMB function to generate distinct combinations in lexicographic order.

```
data _null_;
   array x[5] $3 ('ant' 'bee' 'cat' 'dog' 'ewe');
   n=dim(x);
   k=3;
   ncomb=comb(n,k);
   do j=1 to ncomb+1;
      rc=lexcomb(j, k, of x[*]);
      put j 5. +3 x1-x3 +3 rc=;
      if rc<0 then leave;
   end;
run;
```

SAS writes the following output to the log:

```
    1   ant bee cat    rc=1
    2   ant bee dog    rc=3
    3   ant bee ewe    rc=3
    4   ant cat dog    rc=2
    5   ant cat ewe    rc=3
    6   ant dog ewe    rc=2
    7   bee cat dog    rc=1
    8   bee cat ewe    rc=3
    9   bee dog ewe    rc=2
   10   cat dog ewe    rc=1
   11   cat dog ewe    rc=-1
```

Example 2: Generating Distinct Combinations in Lexicographic Order: Another Example The following is another example of using the LEXCOMB function.

```
data _null_;
   array x[5] $3 ('X' 'Y' 'Z' 'Z' 'Y');
   n=dim(x);
   k=3;
   ncomb=comb(n,k);
   do j=1 to ncomb+1;
      rc=lexcomb(j, k, of x[*]);
      put j 5. +3 x1-x3 +3 rc=;
      if rc<0 then leave;
   end;
run;
```

SAS writes the following output to the log:

```
    1   X Y Y    rc=1
    2   X Y Z    rc=3
    3   X Z Z    rc=2
    4   Y Y Z    rc=1
    5   Y Z Z    rc=2
    6   Y Z Z    rc=-1
```

See Also

Functions and CALL Routines:

"CALL LEXCOMB Routine" on page 458

"ALLCOMB Function" on page 374

LEXCOMBI Function

Generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order.

Category: Combinatorial

Restriction: The LEXCOMBI function cannot be executed when you use the %SYSFUNC macro.

Syntax

LEXCOMBI(*n*, *k*, *index-1, …, k*)

Arguments

n
is a numeric constant, variable, or expression that specifies the total number of objects.

K
is a numeric constant, variable, or expression that specifies the number of objects in each combination.

index
is a numeric variable that contains indices of the objects in the combination that is returned. Indices are integers between 1 and *n* inclusive.

Tip: If *index-1* is missing or zero, then the LEXCOMBI function initializes the indices to **`index-1=1`** through **`index-k=k`**. Otherwise, LEXCOMBI creates a new combination by removing one index from the combination and adding another index.

Details

Before the first execution of the LEXCOMBI function, complete one of the following tasks:

- Set *index-1* equal to zero or to a missing value.
- Initialize *index-1* through *index-k* to distinct integers between 1 and *n* inclusive.

The number of combinations of *n* objects taken *k* at a time can be computed as COMB(*n*,*k*). To generate all combinations of *n* objects taken *k* at a time, execute the LEXCOMBI function in a loop that executes COMB(*n*,*k*) times.

In the LEXCOMBI function, the returned value indicates which, if any, indices changed. If *index-1* through *index-i* did not change, but *index-j* did change, where*j*=*i*+1, then LEXCOMBI returns*i*. If LEXCOMBI is called after the last combinations in lexicographic order have been generated, then LEXCOMBI returns –1.

Comparisons

The LEXCOMBI function generates all combinations of the indices of *n* objects taken *k* at a time in lexicographic order. The ALLCOMBI function generates all combinations of the indices of *n* objects taken *k* at a time in a minimum change order.

Examples

The following example uses the LEXCOMBI function to generate combinations of indices in lexicographic order.

```
data _null_;
   array x[5] $3 ('ant' 'bee' 'cat' 'dog' 'ewe');
   array c[3] $3;
   array i[3];
   n=dim(x);
   k=dim(i);
   i[1]=0;
   ncomb=comb(n,k);
   do j=1 to ncomb+1;
      rc=lexcombi(n, k, of i[*]);
      do h=1 to k;
         c[h]=x[i[h]];
      end;
      put @4 j= @10 'i= ' i[*] +3 'c= ' c[*] +3 rc=;
   end;
run;
```

SAS writes the following output to the log:

```
j=1   i= 1 2 3    c= ant bee cat    rc=1
j=2   i= 1 2 4    c= ant bee dog    rc=3
j=3   i= 1 2 5    c= ant bee ewe    rc=3
j=4   i= 1 3 4    c= ant cat dog    rc=2
j=5   i= 1 3 5    c= ant cat ewe    rc=3
j=6   i= 1 4 5    c= ant dog ewe    rc=2
j=7   i= 2 3 4    c= bee cat dog    rc=1
j=8   i= 2 3 5    c= bee cat ewe    rc=3
j=9   i= 2 4 5    c= bee dog ewe    rc=2
j=10  i= 3 4 5    c= cat dog ewe    rc=1
j=11  i= 3 4 5    c= cat dog ewe    rc=-1
```

See Also

Functions and CALL Routines:

"CALL LEXCOMBI Routine" on page 462

"CALL ALLCOMBI Routine" on page 434

LEXPERK Function

Generates all distinct permutations of the non-missing values of *n* variables taken *k* at a time in lexicographic order.

Category: Combinatorial

Restriction: The LEXPERK function cannot be executed when you use the %SYSFUNC macro.

Syntax

LEXPERK(*count*, *k*, *variable-1*, ..., *variable-n*)

Arguments

count
specifies an integer variable that ranges from 1 to the number of permutations.

k
is a numeric constant, variable, or expression with an integer value between 1 and *n* inclusive.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted.

Requirement: Initialize these variables before you execute the LEXPERK function.

Tip: After executing LEXPERK, the first *k* variables contain the values in one permutation.

Details

The Basics Use the LEXPERK function in a loop where the first argument to LEXPERK takes each integral value from 1 to the number of distinct permutations of *k* non-missing values of the variables. In each execution of LEXPERK within this loop, *k* should have the same value.

Number of Permutations When all of the variables have non-missing, unequal values, the number of permutations is PERM(*n*,*k*). If the number of variables that have missing values is *m*, and all the non-missing values are unequal, the LEXPERK function produces PERM(*n*-*m*,*k*) permutations because the missing values are omitted from the permutations. When some of the variables have equal values, the exact number of permutations is difficult to compute, but PERM(*n*,*k*) provides an upper bound. You do not need to compute the exact number of permutations, provided you exit the loop when the LEXPERK function returns a value that is less than zero.

LEXPERK Processing On the first execution of the LEXPERK function, the following actions occur:

- The argument types and lengths are checked for consistency.
- The *m* missing values are assigned to the last *m* arguments.

- The *n-m* non-missing values are assigned in ascending order to the first *n-m* arguments following *count*.
- LEXPERK returns 1.

On subsequent executions, up to and including the last permutation, the following actions occur:

- The next distinct permutation of *k* non-missing values is generated in lexicographic order.
- If *variable-1* through *variable-i* did not change, but *variable-i* did change, where *j*=*i*+1, then LEXPERK returns *j*.

If you execute the LEXPERK function after generating all the distinct permutations, then LEXPERK returns –1.

If you execute the LEXPERK function with the first argument out of sequence, then the results are not useful. In particular, if you initialize the variables and then immediately execute the LEXPERK function with a first argument of *j*, you will not get the j^{th} permutation (except when *j* is 1). To get the j^{th} permutation, you must execute the LEXPERK function *j* times, with the first argument taking values from 1 through *j* in that exact order.

Comparisons

The LEXPERK function generates all distinct permutations of the non-missing values of *n* variables taken *k* at a time in lexicographic order. The LEXPERM function generates all distinct permutations of the non-missing values of *n* variables in lexicographic order. The ALLPERM function generates all permutations of the values of several variables in a minimal change order.

Examples

The following is an example of the LEXPERK function.

```
data _null_;
   array x[5] $3 ('X' 'Y' 'Z' 'Z' 'Y');
   n=dim(x);
   k=3;
   nperm=perm(n,k);
   do j=1 to nperm+1;
      rc=lexperk(j, k, of x[*]);
      put j 5. +3 x1-x3 +3 rc=;
      if rc<0 then leave;
   end;
run;
```

SAS writes the following output to the log:

```
    1   X Y Y    rc=1
    2   X Y Z    rc=3
    3   X Z Y    rc=2
    4   X Z Z    rc=3
    5   Y X Y    rc=1
    6   Y X Z    rc=3
    7   Y Y X    rc=2
    8   Y Y Z    rc=3
    9   Y Z X    rc=2
   10   Y Z Y    rc=3
```

```
11   Y Z Z    rc=3
12   Z X Y    rc=1
13   Z X Z    rc=3
14   Z Y X    rc=2
15   Z Y Y    rc=3
16   Z Y Z    rc=3
17   Z Z X    rc=2
18   Z Z Y    rc=3
19   Z Z Y    rc=-1
```

See Also

Functions and CALL Routines:

"ALLPERM Function" on page 376
"LEXPERM Function" on page 896
"CALL RANPERK Routine" on page 503
"CALL RANPERM Routine" on page 505

LEXPERM Function

Generates all distinct permutations of the non-missing values of several variables in lexicographic order.

Category: Combinatorial

Syntax

LEXPERM(*count*, *variable-1* <, …, *variable-N*>)

Arguments

count
specifies an integer variable that ranges from 1 to the number of permutations.

variable
specifies either all numeric variables, or all character variables that have the same length. The values of these variables are permuted by LEXPERM.

Requirement: Initialize these variables before you execute the LEXPERM function.

Details

Determine the Number of Distinct Permutations The following variables are defined for use in the equation that follows:

N specifies the number of variables that are being permuted—that is, the number of arguments minus one.

M	specifies the number of missing values among the variables that are being permuted.
d	specifies the number of distinct non-missing values among the arguments.
N_i	for i=1, through i=d, N_i specifies the number of instances of the i*th* distinct value.

The number of distinct permutations of non-missing values of the arguments is expressed as follows:

$$P = \frac{(N_1 + N_2 + ... + N_d)!}{N_1!N_2!...N_d!} <= N!$$

Note: The LEXPERM function cannot be executed with the %SYSFUNC macro. △

LEXPERM Processing Use the LEXPERM function in a loop where the argument *count* takes each integral value from 1 to P. You do not need to compute P provided you exit the loop when LEXPERM returns a value that is less than zero.

For 1=*count*<P, the following actions occur:

- The argument types and lengths are checked for consistency.
- The M missing values are assigned to the last M arguments.
- The N-M non-missing values are assigned in ascending order to the first N-M arguments following *count*.
- LEXPERM returns 1.

For 1<*count*<=P, the following actions occur:

- The next distinct permutation of the non-missing values is generated in lexicographic order.
- If *variable-1* through *variable-I* did not change, but *variable-J* did change, where J=I+1, then LEXPERM returns J.

For *count*>P, LEXPERM returns –1.

If the LEXPERM function is executed with the first argument out of sequence, the results might not be useful. In particular, if you initialize the variables and then immediately execute LEXPERM with a first argument of K, you will not get the K*th* permutation (except when K is 1). To get the K*th* permutation, you must execute LEXPERM K times, with the first argument accepting values from 1 through K in that exact order.

Comparisons

SAS provides three functions or CALL routines for generating all permutations:

- ALLPERM generates all *possible* permutations of the values, *missing or non-missing*, of several variables. Each permutation is formed from the previous permutation by interchanging two consecutive values.
- LEXPERM generates all *distinct* permutations of the *non-missing* values of several variables. The permutations are generated in lexicographic order.
- LEXPERK generates all *distinct* permutations of K of the *non-missing* values of N variables. The permutations are generated in lexicographic order.

ALLPERM is the fastest of these functions and CALL routines. LEXPERK is the slowest.

Examples

The following is an example of the LEXPERM function.

```
data _null_;
   array x[6] $1 ('X' 'Y' 'Z' ' ' 'Z' 'Y');
   nfact=fact(dim(x));
   put +3 nfact=;
   do i=1 to nfact;
      rc=lexperm(i, of x[*]);
      put i 5. +2 rc= +2 x[*];
      if rc<0 then leave;
   end;
run;
```

SAS writes the following output to the log:

```
   nfact=720
    1  rc=1   X Y Y Z Z
    2  rc=3   X Y Z Y Z
    3  rc=4   X Y Z Z Y
    4  rc=2   X Z Y Y Z
    5  rc=4   X Z Y Z Y
    6  rc=3   X Z Z Y Y
    7  rc=1   Y X Y Z Z
    8  rc=3   Y X Z Y Z
    9  rc=4   Y X Z Z Y
   10  rc=2   Y Y X Z Z
   11  rc=3   Y Y Z X Z
   12  rc=4   Y Y Z Z X
   13  rc=2   Y Z X Y Z
   14  rc=4   Y Z X Z Y
   15  rc=3   Y Z Y X Z
   16  rc=4   Y Z Y Z X
   17  rc=3   Y Z Z X Y
   18  rc=4   Y Z Z Y X
   19  rc=1   Z X Y Y Z
   20  rc=4   Z X Y Z Y
   21  rc=3   Z X Z Y Y
   22  rc=2   Z Y X Y Z
   23  rc=4   Z Y X Z Y
   24  rc=3   Z Y Y X Z
   25  rc=4   Z Y Y Z X
   26  rc=3   Z Y Z X Y
   27  rc=4   Z Y Z Y X
   28  rc=2   Z Z X Y Y
   29  rc=3   Z Z Y X Y
   30  rc=4   Z Z Y Y X
   31  rc=-1   Z Z Y Y X
```

See Also

Functions and CALL Routines:

"CALL ALLPERM Routine" on page 437
"ALLPERM Function" on page 376

"CALL RANPERK Routine" on page 503
"CALL RANPERM Routine" on page 505

LFACT Function

Computes the logarithm of the FACT (factorial) function.

Category: Combinatorial

Syntax

LFACT(*n*)

Arguments

n
is an integer that represents the total number of elements from which the sample is chosen.

Comparisons

The LFACT function computes the logarithm of the FACT function.

Examples

The following statements produce these results:

SAS Statements	Results
x=lfact(5000);	
put x;	37591.143509
y=lfact(100);	
put y;	363.73937556

See Also

Functions:
"FACT Function" on page 678

LGAMMA Function

Returns the natural logarithm of the Gamma function.

Category: Mathematical

Syntax

LGAMMA(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Range: must be positive.

Examples

SAS Statements	Results
`x=lgamma(2);`	`0`
`x=lgamma(1.5);`	`-0.120782238`

LIBNAME Function

Assigns or deassigns a libref for a SAS library.

Category: SAS File I/O

See: LIBNAME Function in the documentation for your operating environment.

Syntax

LIBNAME(*libref*<,*SAS-library*<,*engine*<,*options*>>>)

Arguments

libref
is a character constant, variable, or expression that specifies the libref that is assigned to a SAS library.

Tip: The maximum length of *libref* is eight characters.

SAS-library
is a character constant, variable, or expression that specifies the physical name of the SAS library that is associated with the libref. Specify this name as required by the host operating environment. This argument can be null.

engine
is a character constant, variable, or expression that specifies the engine that is used to access SAS files opened in the data library. If you are specifying a SAS/SHARE server, then the value of engine should be REMOTE. This argument can be null.

options
is a character constant, variable, or expression that specifies one or more valid options for the specified engine, delimited with blanks. This argument can be null.

Details

Basic Information about Return Codes

The LIBNAME function assigns or deassigns a libref from a SAS library. When you use the LIBNAME function with two or more arguments, SAS attempts to assign the libref. When you use one argument, SAS attempts to deassign the libref. Return codes are generated depending on the value of the arguments that are used in the LIBNAME function and whether the libref is assigned.

When assigning a libref, the return code will be 0 if the libref is successfully assigned. If the return code is nonzero and the SYSMSG function returns a warning message or a note, then the assignment was successful. If the SYSMSG function returns an error, then the assignment was unsuccessful.

If a library is already assigned, and you attempt to assign a different name to the library, the libref is assigned, the LIBNAME function returns a nonzero return code, and the SYSMSG function returns a note.

When LIBNAME Has One Argument When LIBNAME has one argument, the following rules apply:

- If the libref is not assigned, a nonzero return code is returned and the SYSMSG function returns a warning message.
- If the libref is successfully assigned, a 0 return code is returned and the SYSMSG function returns a blank value.

When LIBNAME Has Two Arguments When LIBNAME has two arguments, the following rules apply:

- If the second argument is NULL, all blanks, or zero length, SAS attempts to deassign the libref.
- If the second argument is not NULL, not all blanks, and not zero length, SAS attempts to assign the specified path (the second argument) to the libref.
- If the libref is not assigned, a nonzero return code is returned and the SYSMSG function returns an error message.
- If the libref is successfully assigned, a 0 return code is returned and the SYSMSG function returns a blank value.

When LIBNAME Has Three or Four Arguments

- If the second argument is NULL, all blanks, or zero length, the results depend on your operating environment.
- If the second argument is NULL and the libref is not already assigned, then a nonzero return code is returned and the SYSMSG function returns an error message.

- If the second argument is NULL and the libref has already been assigned, then LIBNAME returns a value of 0 and the SYSMSG function returns a blank value.
- If at least one of the previous conditions is not met, then SAS attempts to assign the specified path (second argument) to the libref.

Note: In the DATA step, a character constant that consists of two consecutive quotation marks without any intervening spaces is interpreted as a single space, not as a string with a length of 0. To specify a string with a length of 0, use the "TRIMN Function" on page 1175. △

Operating Environment Information: Some systems allow a *SAS-library* value of ' '(a space between the single quotation marks) to assign a libref to the current directory. Other systems disassociate the libref from the SAS library when the *SAS-library* value contains only blanks. The behavior of LIBNAME when a single space is specified for *SAS-library* is dependent on your operating environment.

Under some operating environments, you can assign librefs by using system commands that are outside the SAS session. △

Examples

Example 1: Assigning a Libref

This example attempts to assign the libref NEW to the SAS library MYLIB. If an error or warning occurs, the message is written to the SAS log. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%if (%sysfunc(libname(new,MYLIB))) %then
  %put %sysfunc(sysmsg());
```

Example 2: Deassigning a Libref This example deassigns the libref NEW that was previously associated with the data library MYLIB in the preceding example. If an error or warning occurs, the message is written to the SAS log. In a macro statement, you do not enclose character strings in quotation marks.

```
%if (%sysfunc(libname(new))) %then
  %put %sysfunc(sysmsg());
```

Example 3: Compressing a Library This example assigns the libref NEW to the MYLIB data library and uses the COMPRESS option to compress the library. This example uses the default SAS engine. In a DATA step, you enclose character strings in quotation marks.

```
data test;
   rc=libname('new','MYLIB',,'compress=yes');
run;
```

See Also

Functions:

"LIBREF Function" on page 903

"SYSMSG Function" on page 1154

LIBREF Function

Verifies that a libref has been assigned.

Category: SAS File I/O

See: LIBREF Functionin the documentation for your operating environment.

Syntax

LIBREF(*libref*)

Arguments

libref
specifies the libref to be verified. In a DATA step, *libref* can be a character expression, a string enclosed in quotation marks, or a DATA step variable whose value contains the libref. In a macro, *libref* can be any expression.

Range: 1 to 8 characters

Details

The LIBREF function returns 0 if the libref has been assigned, or returns a nonzero value if the libref has not been assigned.

Examples

This example verifies a libref. If an error or warning occurs, the message is written to the log. Under some operating environments, the user can assign librefs by using system commands outside the SAS session.

```
%if (%sysfunc(libref(sashelp))) %then
  %put %sysfunc(sysmsg());
```

See Also

Function:

"LIBNAME Function" on page 900

LOG Function

Returns the natural (base e) logarithm.

Category: Mathematical

Syntax

LOG(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.
Range: must be positive.

Examples

SAS Statements	Results
`x=log(1.0);`	`0`
`x=log(10.0);`	`2.302585093`

LOG1PX Function

Returns the log of 1 plus the argument.

Category: Mathematical

Syntax

LOG1PX(*x*)

Arguments

x
specifies a numeric variable, constant, or expression.

Details

The LOG1PX function computes the log of 1 plus the argument. The LOG1PX function is mathematically defined by the following equation, where –1 < x:

$$LOG1PX\,(x) = log\,(1 + x)$$

When x is close to 0, **`LOG1PX(x)`** can be more accurate than **`LOG(1+x)`**.

Examples

Example 1: Computing the Log with the LOG1PX Function The following example computes the log of 1 plus the value 0.5.

```
data _null_;
   x=log1px(0.5);
   put x=;
run;
```

SAS writes the following output to the Log:

```
x=0.4054651081
```

Example 2: Comparing the LOG1PX Function with the LOG Function In the following example, the value of X is computed by using the LOG1PX function. The value of Y is computed by using the LOG function.

```
data _null_;
   x=log1px(1.e-5);
   put x= hex16.;

   y=log(1+1.e-5);
   put y= hex16.;
run;
```

SAS writes the following output to the Log:

```
x=3EE4F8AEA9AE7317
y=3EE4F8AEA9AF0A25
```

See Also

Functions:

"LOG Function" on page 903

LOG10 Function

Returns the logarithm to the base 10.

Category: Mathematical

Syntax

LOG10(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Range: must be positive.

Examples

SAS Statements	Results
`x=log10(1.0);`	`0`
`x=log10(10.0);`	`1`
`x=log10(100.0);`	`2`

LOG2 Function

Returns the logarithm to the base 2.

Category: Mathematical

Syntax

LOG2(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Range: must be positive.

Examples

SAS Statements	Results
`x=log2(2.0);`	`1`
`x=log2(0.5);`	`-1`

LOGBETA Function

Returns the logarithm of the beta function.

Category: Mathematical

Syntax

LOGBETA(*a*,*b*)

Arguments

a
is the first shape parameter, where $a>0$.

b
is the second shape parameter, where $b>0$.

Details

The LOGBETA function is mathematically given by the equation

$$log\left(\beta\left(a,b\right)\right)=log\left(\Gamma\left(a\right)\right)+log\left(\Gamma\left(b\right)\right)-log\left(\Gamma\left(a+b\right)\right)$$

where $\Gamma\left(.\right)$ is the gamma function.
If the expression cannot be computed, LOGBETA returns a missing value.

Examples

SAS Statements	Results
`LOGBETA(5,3);`	`-4.653960350`

See Also

Function:
"BETA Function" on page 416

LOGCDF Function

Returns the logarithm of a left cumulative distribution function.

Category: Probability
See: "CDF Function" on page 558

Syntax

LOGCDF(*'dist',quantile<,parm-1,...,parm-k>*)

Arguments

'dist'
is a character constant, variable, or expression that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	**'BERNOULLI'**
Beta	**'BETA'**
Binomial	**'BINOMIAL'**
Cauchy	**'CAUCHY'**
Chi-Square	**'CHISQUARE'**
Exponential	**'EXPONENTIAL'**
F	**'F'**
Gamma	**'GAMMA'**
Geometric	**'GEOMETRIC'**
Hypergeometric	**'HYPERGEOMETRIC'**
Laplace	**'LAPLACE'**
Logistic	**'LOGISTIC'**
Lognormal	**'LOGNORMAL'**
Negative binomial	**'NEGBINOMIAL'**
Normal	**'NORMAL'|'GAUSS'**
Normal mixture	**'NORMALMIX'**
Pareto	**'PARETO'**
Poisson	**'POISSON'**
T	**'T'**
Uniform	**'UNIFORM'**
Wald (inverse Gaussian)	**'WALD'|'IGAUSS'**
Weibull	**'WEIBULL'**

Note: Except for T, F, and NORMALMIX, you can minimally identify any distribution by its first four characters. △

quantile
is a numeric variable, constant, or expression that specifies the value of a random variable.

parm-1,...,parm-k
are optional *shape*, *location*, or *scale* parameters appropriate for the specific distribution.

The LOGCDF function computes the logarithm of a left cumulative distribution function (logarithm of the left side) from various continuous and discrete distributions. For more information, see the "CDF Function" on page 558.

See Also

Functions:

"CDF Function" on page 558
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120
"QUANTILE Function" on page 1064

LOGPDF Function

Returns the logarithm of a probability density (mass) function.

Category: Probability
Alias: LOGPMF
See: "PDF Function" on page 986

Syntax

LOGPDF(*'dist',quantile,parm-1,...,parm-k*)

Arguments

'dist'
is a character constant, variable, or expression that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	**'BERNOULLI'**
Beta	**'BETA'**
Binomial	**'BINOMIAL'**
Cauchy	**'CAUCHY'**
Chi-Square	**'CHISQUARE'**
Exponential	**'EXPONENTIAL'**
F	**'F'**
Gamma	**'GAMMA'**
Geometric	**'GEOMETRIC'**
Hypergeometric	**'HYPERGEOMETRIC'**
Laplace	**'LAPLACE'**
Logistic	**'LOGISTIC'**

Distribution	Argument
Lognormal	**'LOGNORMAL'**
Negative binomial	**'NEGBINOMIAL'**
Normal	**'NORMAL'|'GAUSS'**
Normal mixture	**'NORMALMIX'**
Pareto	**'PARETO'**
Poisson	**'POISSON'**
T	**'T'**
Uniform	**'UNIFORM'**
Wald (inverse Gaussian)	**'WALD'|'IGAUSS'**
Weibull	**'WEIBULL'**

Note: Except for T, F, and NORMALMIX, you can minimally identify any distribution by its first four characters. △

quantile
is a numeric constant, variable, or expression that specifies the value of a random variable.

parm-1,...,parm-k
are optional *shape*, *location*, or *scale* parameters appropriate for the specific distribution.

The LOGPDF function computes the logarithm of the probability density (mass) function from various continuous and discrete distributions. For more information, see the "PDF Function" on page 986.

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120
"QUANTILE Function" on page 1064

LOGSDF Function

Returns the logarithm of a survival function.

Category: Probability

See: "SDF Function" on page 1120

Syntax

LOGSDF(*'dist',quantile,parm-1,...,parm-k*)

Arguments

'dist'
is a character constant, variable, or expression that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	**'BERNOULLI'**
Beta	**'BETA'**
Binomial	**'BINOMIAL'**
Cauchy	**'CAUCHY'**
Chi-Square	**'CHISQUARE'**
Exponential	**'EXPONENTIAL'**
F	**'F'**
Gamma	**'GAMMA'**
Geometric	**'GEOMETRIC'**
Hypergeometric	**'HYPERGEOMETRIC'**
Laplace	**'LAPLACE'**
Logistic	**'LOGISTIC'**
Lognormal	**'LOGNORMAL'**
Negative binomial	**'NEGBINOMIAL'**
Normal	**'NORMAL'|'GAUSS'**
Normal mixture	**'NORMALMIX'**
Pareto	**'PARETO'**
Poisson	**'POISSON'**
T	**'T'**
Uniform	**'UNIFORM'**
Wald (inverse Gaussian)	**'WALD'|'IGAUSS'**
Weibull	**'WEIBULL'**

Note: Except for T, F, and NORMALMIX, you can minimally identify any distribution by its first four characters. △

quantile
is a numeric constant, variable, or expression that specifies the value of a random variable.

parm-1,...,parm-k
are optional *shape*, *location*, or *scale* parameters appropriate for the specific distribution.

The LOGSDF function computes the logarithm of the survival function from various continuous and discrete distributions. For more information, see the "SDF Function" on page 1120.

See Also

Functions:

"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"CDF Function" on page 558
"PDF Function" on page 986
"SDF Function" on page 1120
"QUANTILE Function" on page 1064

LOWCASE Function

Converts all letters in an argument to lowercase.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

LOWCASE(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

In a DATA step, if the LOWCASE function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the argument.

The LOWCASE function copies the character argument, converts all uppercase letters to lowercase letters, and returns the altered value as a result.

The results of the LOWCASE function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

Examples

SAS Statements	Results

`x='INTRODUCTION';`
`y=lowcase(x);`
`put y;`
	`introduction`

See Also

Functions:

"UPCASE Function" on page 1177
"PROPCASE Function" on page 1037

LPERM Function

Computes the logarithm of the PERM function, which is the logarithm of the number of permutations of *n* objects, with the option of including *r* number of elements.

Category: Combinatorial

Syntax

LPERM(*n*<,*r*>)

Arguments

n
is an integer that represents the total number of elements from which the sample is chosen.

r
is an optional integer value that represents the number of chosen elements. If *r* is omitted, the function returns the factorial of *n*.

Restriction: $r \leq n$

Comparisons

The LPERM function computes the logarithm of the PERM function.

Examples

The following statements produce these results:

SAS Statements	Results
x=lperm(5000,500);	
put x;	4232.7715946
y=lperm(100,10);	
put y;	45.586735935

See Also

Functions:

"PERM Function" on page 1008

LPNORM Function

Returns the L*p* norm of the second argument and subsequent non-missing arguments.

Category: Descriptive Statistics

Syntax

LPNORM(*p*, *value-1* <,*value-2* ...>)

Arguments

p
specifies a numeric constant, variable, or expression that is greater than or equal to 1, which is used as the power for computing the L_p norm.

value
specifies a numeric constant, variable, or expression.

Details

If all arguments have missing values, then the result is a missing value. Otherwise, the result is the L_p norm of the non-missing values of the second and subsequent arguments.

In the following example, *p* is the value of the first argument, and $x1, x2, ..., xn$ are the values of the other non-missing arguments.

$$LPNORM\left(p, x1, x2, ..., xn\right) = \left(abs\left(x1\right)^p + abs\left(x2\right)^p + ... + abs\left(xn\right)^p\right)^{1/p}$$

Examples

Example 1: Calculating the L*p* Norm The following example returns the L_p norm of the second and subsequent non-missing arguments.

```
data _null_;
   x1 = lpnorm(1, ., 3, 0, .q, -4);
   x2 = lpnorm(2, ., 3, 0, .q, -4);
   x3 = lpnorm(3, ., 3, 0, .q, -4);
   x999 = lpnorm(999, ., 3, 0, .q, -4);
   put x1= / x2= / x3= / x999=;
run;
```

SAS writes the following output to the log:

```
x1=7
x2=5
x3=4.4979414453
x999=4
```

Example 2: Calculating the L*p* Norm When You Use a Variable List The following example uses a variable list and returns the L_p norm.

```
data _null_;
   x1 = 1;
   x2 = 3;
   x3 = 4;
   x4 = 3;
   x5 = 1;
   x = lpnorm(of x1-x5);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=11
```

See Also

Functions:

"SUMABS Function" on page 1149 (L_1 norm)

"EUCLID Function" on page 673 (L_2 norm)

"MAX Function" on page 921 ($L_{infinity}$ norm)

MAD Function

Returns the median absolute deviation from the median.

Category: Descriptive Statistics

Syntax

MAD(*value-1* <, *value-2...*>)

Arguments

value
specifies a numeric constant, variable, or expression of which the median absolute deviation from the median is to be computed.

Details

If all arguments have missing values, the result is a missing value. Otherwise, the result is the median absolute deviation from the median of the non-missing values. The formula for the median is the same as the one that is used in the UNIVARIATE procedure. For more information, see *Base SAS Procedures Guide*.

Examples

SAS Statements	Results
`mad=mad(2,4,1,3,5,999999);` `put mad;`	`1.5`

See Also

Functions:

"IQR Function" on page 864
"MEDIAN Function" on page 925
"PCTL Function" on page 985

MARGRCLPRC Function

Calculates call prices for European options on stocks, based on the Margrabe model.

Category: Financial

Syntax

MARGRCLPRC(*X_1*, *t*, *X_2*, *sigma1*, *sigma2*, *rho12*)

Arguments

X_1
is a non-missing, positive value that specifies the price of the first asset.

Requirement: Specify X_1 and X_2 in the same units.

t
is a non-missing value that specifies the time to expiration.

X_2
is a non-missing, positive value that specifies the price of the second asset.

Requirement: Specify X_2 and X_1 in the same units.

sigma1
is a non-missing, positive fraction that specifies the volatility of the first asset.

Requirement: *sigma1* must be for the same time period as the unit of *t*.

sigma2
is a non-missing, positive fraction that specifies the volatility of the second asset.

Requirement: Specify a value for *sigma2* for the same time period as the unit of *t*.

rho12
specifies the correlation between the first and second assets, $\rho_{x_1 x_2}$.

Range: between –1 and 1

Details

The MARGRCLPRC function calculates the call price for European options on stocks, based on the Margrabe model. The function is based on the following relationship:

$$\texttt{CALL} = X_1 N(d_1) - X_2 N(d_2)$$

where

X_1	specifies the price of the first asset.
X_2	specifies the price of the second asset.
N	specifies the cumulative normal density function.

$$d_1 = \frac{\left(\ln\left(\frac{X_1}{X_2}\right) + \left(\frac{\sigma^2}{2}\right) t\right)}{\sigma\sqrt{t}}$$
$$d_2 = d_1 - \sigma\sqrt{t}$$
$$\sigma^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 - 2\rho_{x_1,x_2}\sigma_{x_1}\sigma_{x_2}$$

where

t	specifies the time to expiration.
$\sigma_{x_1}^2$	specifies the variance of the first asset.
$\sigma_{x_2}^2$	specifies the variance of the second asset.
σ_{x_1}	specifies the volatility of the first asset.
σ_{x_2}	specifies the volatility of the second asset.
ρ_{x_1,x_2}	specifies the correlation between the first and second assets.

For the special case of t=0, the following equation is true:

$$\texttt{CALL} = \max((X_1 - X_2), 0)$$

Note: This function assumes that there are no dividends from the two assets. △

For information about the basics of pricing, see "Using Pricing Functions" on page 311.

Comparisons

The MARGRCLPRC function calculates the call price for European options on stocks, based on the Margrabe model. The MARGRPTPRC function calculates the put price for European options on stocks, based on the Margrabe model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=margrclprc(500, .5, 950, 4, 5, 1);` `put a;`	`46.441283642`
`b=margrclprc(850, 1.2, 125, 5, 3, 1);` `put b;`	`777.67008185`
`c=margrclprc(7500, .9, 950, 3, 2, 1);` `put c;`	`6562.0354886`
`d=margrclprc(5000, -.5, 237, 3, 3, 1);` `put d;`	`0`

See Also

Functions:

"MARGRPTPRC Function" on page 919

MARGRPTPRC Function

Calculates put prices for European options on stocks, based on the Margrabe model.

Category: Financial

Syntax

MARGRPTPRC(X_1, t, X_2, *sigma1*, *sigma2*, *rho12*)

Arguments

***X*$_1$**
is a non-missing, positive value that specifies the price of the first asset.
Requirement: Specify X_1 and X_2 in the same units.

t
is a non-missing value that specifies the time to expiration.

***X*$_2$**
is a non-missing, positive value that specifies the price of the second asset.
Requirement: Specify X_2 and X_1 in the same units.

sigma1
is a non-missing, positive fraction that specifies the volatility of the first asset.
Requirement: *sigma1* must be for the same time period as the unit of t.

sigma2
is a non-missing, positive fraction that specifies the volatility of the second asset.

Requirement: Specify a value for *sigma2* for the same time period as the unit of *t*.

rho12
specifies the correlation between the first and second assets, $\rho_{x_1 x_2}$.

Range: between –1 and 1

Details

The MARGRPTPRC function calculates the put price for European options on stocks, based on the Margrabe model. The function is based on the following relationship:

$$\texttt{PUT} = X_2 N\left(pd_1\right) - X_1 N\left(pd_2\right)$$

where

X_1	specifies the price of the first asset.
X_2	specifies the price of the second asset.
N	specifies the cumulative normal density function.

$$pd_1 = \frac{\left(\ln\left(\frac{X_1}{X_2}\right) + \left(\frac{\sigma^2}{2}\right)t\right)}{\sigma\sqrt{t}}$$
$$pd_2 = pd_1 - \sigma\sqrt{t}$$
$$\sigma^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 - 2\rho_{x_1,x_2}\sigma_{x_1}\sigma_{x_2}$$

where

t	is a non-missing value that specifies the time to expiration.
$\sigma_{x_1}^2$	specifies the variance of the first asset.
$\sigma_{x_2}^2$	specifies the variance of the second asset.
σ_{x_1}	specifies the volatility of the first asset.
σ_{x_2}	specifies the volatility of the second asset.
ρ_{x_1,x_2}	specifies the correlation between the first and second assets.

To view the corresponding CALL relationship, see the "MARGRCLPRC Function" on page 917.

For the special case of t=0, the following equation is true:

$$\texttt{PUT} = \max\left(\left(X_2 - X_1\right), 0\right)$$

Note: This function assumes that there are no dividends from the two assets. △

For basic information about pricing, see "Using Pricing Functions" on page 311.

Comparisons

The MARGRPTPRC function calculates the put price for European options on stocks, based on the Margrabe model. The MARGRCLPRC function calculates the call price for

European options on stocks, based on the Margrabe model. These functions return a scalar value.

Examples

SAS Statements	Results
	`----+----1----+-----2--`
`a=margrptprc(500, .5, 950, 4, 5, 1);` `put a;`	`496.44128364`
`b=margrptprc(850, 1.2, 125, 5, 3, 1);` `put b;`	`52.670081846`
`c=margrptprc(7500, .9, 950, 3, 2, 1);` `put c;`	`12.035488581`
`d=margrptprc(5000, -.5, 237, 3, 3, 1);` `put d;`	`0`

See Also

Functions:

"MARGRCLPRC Function" on page 917

MAX Function

Returns the largest value.

Category: Descriptive Statistics

Syntax

MAX(*argument-1,argument-2<,...argument-n>*)

Arguments

argument
specifies a numeric constant, variable, or expression. At least two arguments are required. The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The MAX function returns a missing value (.) only if all arguments are missing.

The MAX operator (<>) returns a missing value only if both operands are missing. In this case, it returns the value of the operand that is higher in the sort order for missing values.

Examples

SAS Statements	Results
`x=max(8,3);`	8
`x1=max(2,6,.);`	6
`x2=max(2,-3,1,-1);`	2
`x3=max(3,.,-3);`	3
`x4=max(of x1-x3);`	6

MD5 Function

Returns the result of the message digest of a specified string.

Category: Character

Syntax

MD5(*string*)

Arguments

string
specifies a character constant, variable, or expression.

Tip: Enclose a literal string of characters in quotation marks.

Details

Length of Returned Variable In a DATA step, if the MD5 function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The MD5 function converts a string, based on the MD5 algorithm, into a 128-bit hash value. This hash value is referred to as a message digest (digital signature), which is nearly unique for each string that is passed to the function.

The MD5 function does not format its own output. You must specify a valid format (such as hex32. or binary128.) to view readable results.

Operating Environment Information: In the z/OS operating environment, the MD5 function produces output in EBCDIC rather than in ASCII. Therefore, the output will differ. △

The Message Digest Algorithm A message digest results from manipulating and compacting an arbitrarily long stream of binary data. An ideal message digest algorithm never generates the same result for two different sets of input. However, generating such a unique result would require a message digest as long as the input itself. Therefore, MD5 generates a message digest of modest size (16 bytes), created with an algorithm that is designed to make a nearly unique result.

Using the MD5 Function You can use the MD5 function to track changes in your data sets. The MD5 function can generate a digest of a set of column values in a record in a table. This digest could be treated as the signature of the record, and be used to keep track of changes that are made to the record. If the digest from the new record matches the existing digest of a record in a table, then the two records are the same. If the digest is different, then a column value in the record has changed. The new changed record could then be added to the table along with a new surrogate key because it represents a change to an existing keyed value.

The MD5 function can be useful when developing shell scripts or Perl programs for software installation, for file comparison, and for detection of file corruption and tampering.

You can also use the MD5 function to create a unique identifier for observations to be used as the key of a hash object. For information about hash objects, see "Introduction to DATA Step Component Objects" in *SAS Language Reference: Concepts*.

Examples

The following is an example of how to generate results that are returned by the MD5 function.

```
data _null_;
   y = md5('abc');
   z = md5('access method');
   put y= / y = hex32.;
   put z= / z = hex32.;
run;
```

The output from this program contains unprintable characters.

MDY Function

Returns a SAS date value from month, day, and year values.

Category: Date and Time

Syntax

MDY(*month,day,year*)

Arguments

month
specifies a numeric constant, variable, or expression that represents an integer from 1 through 12.

day
specifies a numeric constant, variable, or expression that represents an integer from 1 through 31.

year
specifies a numeric constant, variable, or expression with a value of a two-digit or four-digit integer that represents the year. The YEARCUTOFF= system option defines the year value for two-digit dates.

Examples

SAS Statements	Results
`birthday=mdy(8,27,90);`	
`put birthday;`	`11196`
`put birthday= worddate.;`	`birthday=August 27, 1990`
`anniversary=mdy(7,11,2001);`	
`put anniversary;`	`15167`
`put anniversary=date9.;`	`anniversary=11JUL2001`

See Also

Functions:

"DAY Function" on page 637
"MONTH Function" on page 936
"YEAR Function" on page 1233

MEAN Function

Returns the arithmetic mean (average).

Category: Descriptive Statistics

Syntax

MEAN(*argument-1*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least one non-missing argument is required. Otherwise, the function returns a missing value.

Tip: The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The GEOMEAN function returns the geometric mean, the HARMEAN function returns the harmonic mean, and the MEDIAN function returns the median of the non-missing values, whereas the MEAN function returns the arithmetic mean (average).

Examples

SAS Statements	Results
`x1=mean(2,.,.,6);`	`4`
`x2=mean(1,2,3,2);`	`2`
`x3=mean(of x1-x2);`	`3`

See Also

Function:

"GEOMEAN Function" on page 788
"GEOMEANZ Function" on page 790
"HARMEAN Function" on page 799
"HARMEANZ Function" on page 801
"MEDIAN Function" on page 925

MEDIAN Function

Returns the median value.

Category: Descriptive Statistics

Syntax

MEDIAN(*value1*<, *value2, ...*>)

Arguments

value
is a numeric constant, variable, or expression.

Details

The MEDIAN function returns the median of the nonmissing values. If all arguments have missing values, the result is a missing value.

Note: The formula that is used in the MEDIAN function is the same as the formula that is used in PROC UNIVARIATE. For more information, see "SAS Elementary Statistics Procedures" in *Base SAS Procedures Guide*. △

Comparisons

The MEDIAN function returns the median of nonmissing values, whereas the MEAN function returns the arithmetic mean (average).

Examples

SAS Statements	Results
`x=median(2,4,1,3);`	`2.5`
`y=median(5,8,0,3,4);`	`4`

See Also

Function:

"MEAN Function" on page 925

MIN Function

Returns the smallest value.

Category: Descriptive Statistics

Syntax

MIN(*argument-1,argument-2*<,*...argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least two arguments are required. The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The MIN function returns a missing value (.) only if all arguments are missing.

The MIN operator (><) returns a missing value if either operand is missing. In this case, it returns the value of the operand that is lower in the sort order for missing values.

Examples

SAS Statements	Results
`x=min(7,4);`	`4`
`x1=min(2,.,6);`	`2`
`x2=min(2,-3,1,-1);`	`-3`
`x3=min(0,4);`	`0`
`x4=min(of x1-x3);`	`-3`

MINUTE Function

Returns the minute from a SAS time or datetime value.

Category: Date and Time

Syntax

MINUTE(*time* | *datetime*)

Arguments

time
is a numeric constant, variable, or expression that specifies a SAS time value.

datetime
is a numeric constant, variable, or expression that specifies a SAS datetime value.

Details

The MINUTE function returns an integer that represents a specific minute of the hour. MINUTE always returns a positive number in the range of 0 through 59.

Examples

SAS Statements	Results

`time='3:19:24't;`
`m=minute(time);`
`put m;`
	`19`

See Also

Functions:

"HOUR Function" on page 807
"SECOND Function" on page 1122

MISSING Function

Returns a numeric result that indicates whether the argument contains a missing value.

Category: Descriptive Statistics
Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

MISSING(*numeric-expression* | *character-expression*)

Arguments

numeric-expression
specifies a numeric constant, variable, or expression.

character-expression
specifies a character constant, variable, or expression.

Details

- The MISSING function checks a numeric or character expression for a missing value, and returns a numeric result. If the argument does not contain a missing value, SAS returns a value of 0. If the argument contains a missing value, SAS returns a value of 1.
- A numeric expression is considered missing if it evaluates to a numeric missing value: ., ._, .A, ..., .Z.
- A character expression is considered missing if it evaluates to a string that contains all blanks or has a length of zero.

Comparisons

The MISSING function can have only one argument. The CMISS function can have multiple arguments and returns a count of the missing values. The NMISS function requires numeric arguments and returns the number of missing values in the list of arguments.

Examples

This example uses the MISSING function to check whether the input variables contain missing values.

```
data values;
   input @1 var1 3. @5 var2 3.;
   if missing(var1) then
      do;
         put 'Variable 1 is Missing.';
      end;
   else if missing(var2) then
      do;
```

```
         put 'Variable 2 is Missing.';
      end;
   datalines;
127
988 195
;

run;
```

SAS writes the following output to the log:

```
Variable 2 is Missing.
```

See Also

Functions and CALL Routines:

"CMISS Function" on page 584
"NMISS Function" on page 947
"CALL MISSING Routine" on page 473

MOD Function

Returns the remainder from the division of the first argument by the second argument, fuzzed to avoid most unexpected floating-point results.

Category: Mathematical

Syntax

MOD (*argument-1, argument-2*)

Arguments

argument-1
is a numeric constant, variable, or expression that specifies the dividend.

argument-2
is a numeric constant, variable, or expression that specifies the divisor.

Restriction: cannot be 0

Details

The MOD function returns the remainder from the division of *argument-1* by *argument-2*. When the result is non-zero, the result has the same sign as the first argument. The sign of the second argument is ignored.

The computation that is performed by the MOD function is exact if both of the following conditions are true:

□ Both arguments are exact integers.

- All integers that are less than either argument have exact 8-byte floating-point representations.

  To determine the largest integer for which the computation is exact, execute the following DATA step:

  ```
  data _null_;
     exactint = constant('exactint');
     put exactint=;
  run;
  ```

  Operating Environment Information: You can also refer to the SAS documentation for your operating environment for information about the largest integer. △

If either of the above conditions is not true, a small amount of numerical error can occur in the floating-point computation. In this case

- MOD returns zero if the remainder is very close to zero or very close to the value of the second argument.
- MOD returns a missing value if the remainder cannot be computed to a precision of approximately three digits or more. In this case, SAS also writes an error message to the log.

Note: Before SAS 9, the MOD function did not perform the adjustments to the remainder that were described in the previous paragraph. For this reason, the results of the MOD function in SAS 9 might differ from previous versions. △

Comparisons

Here are some comparisons between the MOD and MODZ functions:

- The MOD function performs extra computations, called fuzzing, to return an exact zero when the result would otherwise differ from zero because of numerical error.
- The MODZ function performs no fuzzing.
- Both the MOD and MODZ functions return a missing value if the remainder cannot be computed to a precision of approximately three digits or more.

Examples

The following SAS statements produce results for MOD and MODZ.

SAS Statements	Results
`x1=mod(10,3);` `put x1 9.4;`	`1.0000`
`xa=modz(10,3);` `put xa 9.4;`	`1.0000`
`x2=mod(.3,-.1);` `put x2 9.4;`	`0.0000`
`xb=modz(.3,-.1);` `put xb 9.4;`	`0.1000`
`x3=mod(1.7,.1);` `put x3 9.4;`	`0.0000`
`xc=modz(1.7,.1);` `put xc 9.4;`	`0.0000`

SAS Statements	Results
`x4=mod(.9,.3);` `put x4 24.20;`	`0.00000000000000000000`
`xd=modz(.9,.3);` `put xd 24.20;`	`0.00000000000000005551`

See Also

Functions:

"INT Function" on page 829
"INTZ Function" on page 861
"MODZ Function" on page 934

MODEXIST Function

Determines whether a software image exists in the version of SAS that you have installed.

Category: Numeric

Syntax

MODEXIST('*product-name*')

Arguments

'*product-name*'
specifies a character constant, variable, or expression that is the name of the product image you are checking.

Details

The MODEXIST function determines whether a software image exists in the version of SAS that you have installed. If an image exists, then MODEXIST returns a value of 1. If an image does not exist, then MODEXIST returns a value of 0.

Comparisons

The MODEXIST function determines whether a software image exists in the version of SAS that you have installed. The SYSPROD function determines whether a product is licensed.

Examples

Example 1: Determining Whether a Product Is Licensed and the Image Is Installed This example returns a value of 1 if a SAS/GRAPH image is installed in

your version of SAS, and returns a value of 0 if the image is not installed. The SYSPROD function determines whether the product is licensed.

```
data _null_;
   rc1 = sysprod('graph');
   rc2 = modexist('sasgplot');
   put rc1= rc2=;
run;
```

Output 4.60 Output from MODEXIST

```
rc1=1 rc2=1
```

MODULEC Function

Calls an external routine and returns a character value.

Category: External Routines
See: "CALL MODULE Routine" on page 475

Syntax

MODULEC(<*cntl-string*,>*module-name*<,*argument-1, ..., argument-n*>)

Details

For details on the MODULEC function, see "CALL MODULE Routine" on page 475.

See Also

Functions and CALL Routines:
"CALL MODULE Routine" on page 475
"MODULEN Function" on page 933

MODULEN Function

Calls an external routine and returns a numeric value.

Category: External Routines
See: "CALL MODULE Routine" on page 475

Syntax

MODULEN(<*cntl-string*,>*module-name*<,*argument-1, ..., argument-n*>)

Details

For details about the MODULEN function, see “CALL MODULE Routine” on page 475.

See Also

Functions and CALL Routines:

“CALL MODULE Routine” on page 475
“MODULEC Function” on page 933

MODZ Function

Returns the remainder from the division of the first argument by the second argument, using zero fuzzing.

Category: Mathematical

Syntax

MODZ (*argument-1*, *argument-2*)

Arguments

argument-1
is a numeric constant, variable, or expression that specifies the dividend.

argument-2
is a non-zero numeric constant, variable, or expression that specifies the divisor.

Details

The MODZ function returns the remainder from the division of *argument-1* by *argument-2*. When the result is non-zero, the result has the same sign as the first argument. The sign of the second argument is ignored.

The computation that is performed by the MODZ function is exact if both of the following conditions are true:

- Both arguments are exact integers.
- All integers that are less than either argument have exact 8-byte floating-point representation.

  To determine the largest integer for which the computation is exact, execute the following DATA step:

```
data _null_;
   exactint = constant('exactint');
   put exactint=;
run;
```

Operating Environment Information: You can also refer to the SAS documentation for your operating environment for information about the largest integer. △

If either of the above conditions is not true, a small amount of numerical error can occur in the floating-point computation. For example, when you use exact arithmetic and the result is zero, MODZ might return a very small positive value or a value slightly less than the second argument.

Comparisons

Here are some comparisons between the MODZ and MOD functions:

- The MODZ function performs no fuzzing.
- The MOD function performs extra computations, called fuzzing, to return an exact zero when the result would otherwise differ from zero because of numerical error.
- Both the MODZ and MOD functions return a missing value if the remainder cannot be computed to a precision of approximately three digits or more.

Examples

The following SAS statements produce results for MOD and MODZ.

SAS Statements	Results
`x1=mod(10,3);` `put x1 9.4;`	`1.0000`
`xa=modz(10,3);` `put xa 9.4;`	`1.0000`
`x2=mod(.3,-.1);` `put x2 9.4;`	`0.0000`
`xb=modz(.3,-.1);` `put xb 9.4;`	`0.1000`
`x3=mod(1.7,.1);` `put x3 9.4;`	`0.0000`
`xc=modz(1.7,.1);` `put xc 9.4;`	`0.0000`
`x4=mod(.9,.3);` `put x4 24.20;`	`0.00000000000000000000`
`xd=modz(.9,.3);` `put xd 24.20;`	`0.00000000000000005551`

See Also

Functions:

"INT Function" on page 829

"INTZ Function" on page 861

"MOD Function" on page 930

MONTH Function

Returns the month from a SAS date value.

Category: Date and Time

Syntax

MONTH(*date*)

Arguments

date
specifies a numeric constant, variable, or expression that represents a SAS date value.

Details

The MONTH function returns a numeric value that represents the month from a SAS date value. Numeric values can range from 1 through 12.

Examples

SAS Statements	Results

`date='25jan94'd;`
`m=month(date);`
`put m;`
	`1`

See Also

Functions:

"DAY Function" on page 637
"YEAR Function" on page 1233

MOPEN Function

Opens a file by directory ID and member name, and returns either the file identifier or a 0.

Category: External Files
See: MOPEN Function in the documentation for your operating environment.

Syntax

MOPEN(*directory-id,member-name<,open-mode<,record-length<,record-format>>>*)

Arguments

directory-id
is a numeric variable that specifies the identifier that was assigned when the directory was opened, generally by the DOPEN function.

member-name
is a character constant, variable, or expression that specifies the member name in the directory.

open-mode
is a character constant, variable, or expression that specifies the type of access to the file:

A	APPEND mode allows writing new records after the current end of the file.
I	INPUT mode allows reading only (default).
O	OUTPUT mode defaults to the OPEN mode specified in the operating environment option in the FILENAME statement or function. If no operating environment option is specified, it allows writing new records at the beginning of the file.
S	Sequential input mode is used for pipes and other sequential devices such as hardware ports.
U	UPDATE mode allows both reading and writing.
W	Sequential update mode is used for pipes and other sequential devices such as ports.

Default: I

record-length
is a numeric variable, constant, or expression that specifies a new logical record length for the file. To use the existing record length for the file, specify a length of 0, or do not provide a value here.

record-format
is a character constant, variable, or expression that specifies a new record format for the file. To use the existing record format, do not specify a value here. The following values are valid:

B	specifies that data is to be interpreted as binary data.
D	specifies the default record format.
E	specifies the record format that you can edit.
F	specifies that the file contains fixed-length records.
P	specifies that the file contains printer carriage control in operating environment-dependent record format.
V	specifies that the file contains variable-length records.

Note: If an argument is invalid, then MOPEN returns 0. You can obtain the text of the corresponding error message from the SYSMSG function. Invalid arguments do not produce a message in the SAS log and do not set the _ERROR_ automatic variable. △

Details

MOPEN returns the identifier for the file, or 0 if the file could not be opened. You can use a *file-id* that is returned by the MOPEN function as you would use a *file-id* returned by the FOPEN function.

CAUTION:
Use OUTPUT mode with care. Opening an existing file for output might overwrite the current contents of the file without warning. △

The member is identified by *directory-id* and *member-name* instead of by a fileref. You can also open a directory member by using FILENAME to assign a fileref to the member, followed by a call to FOPEN. However, when you use MOPEN, you do not have to use a separate fileref for each member.

If the file already exists, the output and update modes default to the operating environment option (append or replace) specified with the FILENAME statement or function. For example,

```
%let rc=%sysfunc(filename(file,physical-name,,mod));
%let did=%sysfunc(dopen(&file));
%let fid=%sysfunc(mopen(&did,member-name,o,0,d));
%let rc=%sysfunc(fput(&fid,This is a test.));
%let rc=%sysfunc(fwrite(&fid));
%let rc=%sysfunc(fclose(&fid));
```

If **'file'** already exists, FWRITE appends the new record instead of writing it at the beginning of the file. However, if no operating environment option is specified with the FILENAME function, the output mode implies that the record be replaced.

If the open fails, use SYSMSG to retrieve the message text.

Operating Environment Information: The term *directory* in this description refers to an aggregate grouping of files that are managed by the operating environment. Different host operating environments identify such groupings with different names, such as directory, subdirectory, folder, MACLIB, or partitioned data set. For details, see the SAS documentation for your operating environment.

Opening a directory member for output or append is not possible in some operating environments. △

Examples

This example assigns the fileref MYDIR to a directory. Then it opens the directory, determines the number of members, retrieves the name of the first member, and opens that member. The last three arguments to MOPEN are the defaults. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let filrf=mydir;
%let rc=%sysfunc(filename(filrf,physical-name));
%let did=%sysfunc(dopen(&filrf));
%let frstname=' ';
%let memcount=%sysfunc(dnum(&did));
%if (&memcount > 0) %then
   %do;
      %let frstname =
         %sysfunc(dread(&did,1));
      %let fid =
         %sysfunc(mopen(&did,&frstname,i,0,d));
      macro statements to process the member
```

```
      %let rc=%sysfunc(fclose(&fid));
   %end;
%else
   %put %sysfunc(sysmsg());
%let rc=%sysfunc(dclose(&did));
```

See Also

Functions:

"DCLOSE Function" on page 638
"DNUM Function" on page 660
"DOPEN Function" on page 661
"DREAD Function" on page 665
"FCLOSE Function" on page 680
"FILENAME Function" on page 690
"FOPEN Function" on page 762
"FPUT Function" on page 771
"FWRITE Function" on page 778
"SYSMSG Function" on page 1154

MORT Function

Returns amortization parameters.

Category: Financial

Syntax

MORT(*a,p,r,n*)

Arguments

a
is numeric, and specifies the initial amount.

p
is numeric, and specifies the periodic payment.

r
is numeric, and specifies the periodic interest rate that is expressed as a fraction.

n
is an integer, and specifies the number of compounding periods.

Range: $n \geq 0$

Details

Calculating Results The MORT function returns the missing argument in the list of four arguments from an amortization calculation with a fixed interest rate that is compounded each period. The arguments are related by the following equation:

$$p = \frac{ar\left(1+r\right)^{n}}{\left(1+r\right)^{n}-1}$$

One missing argument must be provided. The value is then calculated from the remaining three. No adjustment is made to convert the results to round numbers.

Restrictions in Calculating Results The MORT function returns an invalid argument note to the SAS log and sets _ERROR_ to 1 if one of the following argument combinations is true:

- rate < –1 or n < 0
- principal <= 0 or payment <= 0 or n <= 0
- principal <= 0 or payment <= 0 or rate <= –1
- principal * rate > payment
- principal > payment * n

Examples

In the following statement, an amount of $50,000 is borrowed for 30 years at an annual interest rate of 10 percent compounded monthly. The monthly payment can be expressed as follows:

```
payment=mort(50000, . , .10/12,30*12);
```

The value that is returned is 438.79 (rounded). The second argument has been set to missing, which indicates that the periodic payment is to be calculated. The 10 percent nominal annual rate has been converted to a monthly rate of 0.10/12. The rate is the fractional (not the percentage) interest rate per compounding period. The 30 years are converted to 360 months.

MSPLINT Function

Returns the ordinate of a monotonicity-preserving interpolating spline.

Category: Mathematical

Syntax

MSPLINT(*X*, *n*, X_1 <, X_2, …, X_n>, Y_1 <,Y_2, …, Y_n> <, D_1, D_n>)

Arguments

X
is a numeric constant, variable, or expression that specifies the abscissa for which the ordinate of the spline is to be computed.

n
is a numeric constant, variable, or expression that specifies the number of knots. *N* must be a positive integer.

***X*$_1$, ..., *X*$_n$**
are numeric constants, variables, or expressions that specify the abscissas of the knots. These values must be non-missing and listed in nondecreasing order. Otherwise, the result is undefined. MSPLINT does not check the order of the X_1 through X_n arguments.

***Y*$_1$, ..., *Y*$_n$**
are numeric constants, variables, or expressions that specify the ordinates of the knots. The number of Y_1 through Y_n arguments must be the same as the number of X_1 throughX_n arguments.

***D*$_1$, *D*$_n$**
are optional numeric constants, variables, or expressions that specify the derivatives of the spline at X_1 and X_n. These derivatives affect only abscissas that are less than X_2 or greater than X_{n-1}.

Details

The MSPLINT function returns the ordinate of a monotonicity-preserving cubic interpolating spline for a single abscissa, *X*.

An interpolating spline is a function that passes through each point that is specified by the ordered pairs (X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n). These points are called knots.

A spline preserves monotonicity if both of the following conditions are true:

- For any two or more consecutive knots with nondecreasing ordinates, all interpolated values within that interval are also nondecreasing.
- For any two or more consecutive knots with nonincreasing ordinates, all interpolated values within that interval are also nonincreasing.

However, if you specify values of D_1 or D_n with the wrong sign, monotonicity will not be preserved for values that are less than X_2 or greater than X_{n-1}.

If the arguments D_1 and D_n are omitted or missing, then the following actions occur:

- For *n*=1, MSPLINT returns Y_1.
- For *n*=2, MSPLINT uses linear interpolation or extrapolation.

If the arguments D_1 and D_n have non-missing values, or if *n* >= 3, then the following actions occur:

- If $X < X_1$ or $X > X_n$, MSPLINT uses linear extrapolation.
- If $X_1 <= X <= X_n$, MSPLINT uses cubic spline interpolation.

If two knots have equal abscissas but different ordinates, then the spline will be discontinuous at that abscissa. If two knots have equal abscissas and equal ordinates, then the spline will be continuous at that abscissa, but the first derivative will usually be discontinuous at that abscissa. Otherwise, the spline is continuous and has a continuous first derivative.

If X is missing, or if any other arguments required to compute the result are missing, then MSPLINT returns a missing value. MSPLINT does not check all of the arguments for missing values. Because the arguments D_1 and D_n are optional, and they are not required to compute the result, if one or both are missing and no errors occur, then MSPLINT returns a non-missing result.

Examples

The following is an example of the MSPLINT function.

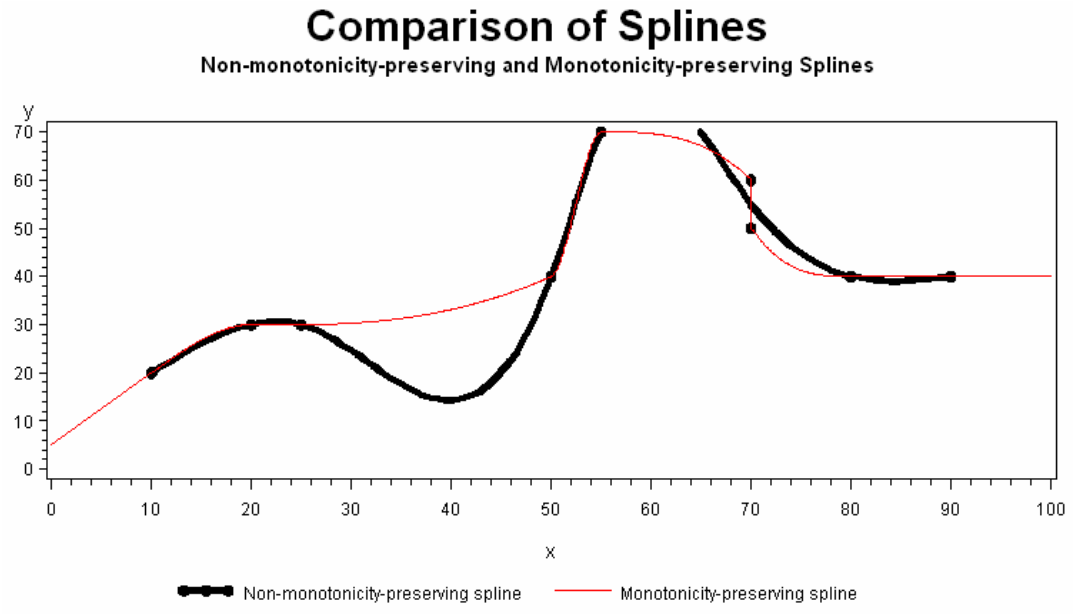
```
data msplint;
   do x=0 to 100 by .1;
      msplint=msplint(x, 9,
         10, 20, 25, 50, 55, 70, 70, 80, 90,
         20, 30, 30, 40, 70, 60, 50, 40, 40);
      output;
   end;
run;

data knots;
   input x y;
   datalines;
10 20
20 30
25 30
50 40
55 70
70 60
70 50
80 40
90 40
;

data plot;
   merge knots msplint;
   by x;
run;

title "Comparison of Splines";
title2 "Non-monotonicity-preserving and Monotonicity-preserving
        Splines";
legend1 value=('Non-monotonicity-preserving spline'
               'Monotonicity-preserving spline') label=none;
symbol1 value=dot  interpol=spline color=black width=5;
symbol2 value=none interpol=join   color=red;
proc gplot data=plot;
   plot y*x=1 msplint*x=2/overlay legend=legend1;
run;
quit;
```

Display 4.8 Results of Using the MSPLINT Function



Reference

Fritsch, F. N., and J. Butland. 1984. "A method for constructing local monotone piecewise cubic interpolants." *Siam Journal of Scientific and Statistical Computing* 5:2, 300-304.

N Function

Returns the number of non-missing numeric values.

Category: Descriptive Statistics

Syntax

N(*argument-1*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least one argument is required. The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The N function counts nonmissing values, whereas the NMISS and the CMISS functions count missing values. N requires numeric arguments, whereas CMISS works with both numeric and character values.

Examples

SAS Statements	Results
`x1=n(1,0,.,2,5,.);`	`4`
`x2=n(1,2);`	`2`
`x3=n(of x1-x2);`	`2`

NETPV Function

Returns the net present value as a fraction.

Category: Financial

Syntax

NETPV(*r,freq,c0,c1,...,cn*)

r
is numeric, the interest rate over a specified base period of time expressed as a fraction.

Range: $r >= 0$

freq
is numeric, the number of payments during the base period of time that is specified with the rate *r*.

Range: $freq > 0$

Exception: The case $freq = 0$ is a flag to allow continuous discounting.

c0,c1,...,cn
are numeric cash flows that represent cash outlays (payments) or cash inflows (income) occurring at times 0, 1, ...n. These cash flows are assumed to be equally spaced, beginning-of-period values. Negative values represent payments, positive values represent income, and values of 0 represent no cash flow at a given time. The *c0* argument and the *c1* argument are required.

Details

The NETPV function returns the net present value at time 0 for the set of cash payments *c0*,*c1*, ...,*cn*, with a rate *r* over a specified base period of time. The argument *freq*>0 describes the number of payments that occur over the specified base period of time.

The net present value is given by

$$\mathrm{NETPV}\left(r, freq, c_0, c_1, ..., c_n\right) = \sum_{i=0}^{n} c_i x^i$$

where

$$x = \begin{cases} \frac{1}{(1+r)^{(1/freq)}} & freq > 0 \\ e^{-r} & freq = 0 \end{cases}$$

Missing values in the payments are treated as 0 values. When *freq*>0, the rate *r* is the effective rate over the specified base period. To compute with a quarterly rate (the base period is three months) of 4 percent with monthly cash payments, set *freq* to 3 and set *r* to .04.

If *freq* is 0, continuous discounting is assumed. The base period is the time interval between two consecutive payments, and the rate *r* is a nominal rate.

To compute with a nominal annual interest rate of 11 percent discounted continuously with monthly payments, set *freq* to 0 and set *r* to .11/12.

Examples

For an initial investment of $500 that returns biannual payments of $200, $300, and $400 over the succeeding 6 years and an annual discount rate of 10 percent, the net present value of the investment can be expressed as follows:

```
value=netpv(.10,.5,-500,200,300,400);
```

The value returned is 95.98.

NLITERAL Function

Converts a character string that you specify to a SAS name literal.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NLITERAL(*string*)

Arguments

string

specifies a character constant, variable, or expression that is to be converted to a SAS name literal.

Tip: Enclose a literal string of characters in quotation marks.

Restriction: If the string is a valid SAS variable name, it is not changed.

Details

Length of Returned Variable In a DATA step, if the NLITERAL function returns a value to a variable that has not previously been assigned a length, then the variable is given a length of 200 bytes.

The Basics *String* will be converted to a name literal, unless it qualifies under the default rules for a SAS variable name. These default rules are in effect when the SAS system option VALIDVARNAME=V7:

- It begins with an English letter or an underscore.
- All subsequent characters are English letters, underscores, or digits.
- The length is 32 or fewer alphanumeric characters.

String qualifies as a SAS variable name, when all of these rules are true.

The NLITERAL function encloses the value of *string* in single or double quotation marks, based on the contents of *string*.

Value in *string*	Result
an ampersand (&)	enclosed in single quotation marks
a percent sign (%)	enclosed in single quotation marks
more double quotation marks than single quotation marks	enclosed in single quotation marks
none of the above	enclosed in double quotation marks

If insufficient space is available for the resulting n-literal, NLITERAL returns a blank string, prints an error message, and sets _ERROR_ to 1.

Examples

This example demonstrates multiple uses of NLITERAL.

```
data test;
   input string $32.;
   length result $ 67;
   result = nliteral(string);
   datalines;
abc_123
This and That
cats & dogs
Company's profits (%)
"Double Quotes"
'Single Quotes'
;

proc print;
title 'Strings Converted to N-Literals or Returned Unchanged';
run;
```

Output 4.61 Converting Strings to Name Literals with NLITERAL

```
          Strings Converted to N-Literals or Returned Unchanged            1

        Obs    string                   result

         1     abc_123                  abc_123
         2     This and That            "This and That"N
         3     cats & dogs              'cats & dogs'N
         4     Company's profits (%)    'Company''s profits (%)'N
         5     "Double Quotes"          '"Double Quotes"'N
         6     'Single Quotes'          "'Single Quotes'"N
```

See Also

Functions:

"COMPARE Function" on page 591
"DEQUOTE Function" on page 646
"NVALID Function" on page 975

System Option:

"VALIDVARNAME= System Option" on page 2049

"Rules for Words and Names in the SAS Language" in *SAS Language Reference: Concepts*

NMISS Function

Returns the number of missing numeric values.

Category: Descriptive Statistics

Syntax

NMISS(*argument-1*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least one argument is required. The argument list can consist of a variable list, which is preceded by OF.

Comparisons

The NMISS function returns the number of missing values, whereas the N function returns the number of nonmissing values. NMISS requires numeric values, whereas CMISS works with both numeric and character values. NMISS works with multiple numeric values, whereas MISSING works with only one value that can be either numeric or character.

Examples

SAS Statements	Results
`x1=nmiss(1,0,.,2,5,.);`	`2`
`x2=nmiss(1,0);`	`0`
`x3=nmiss(of x1-x2);`	`0`

NORMAL Function

Returns a random variate from a normal, or Gaussian, distribution.

Category: Random Number
Alias: RANNOR
See: "RANNOR Function" on page 1086

NOTALNUM Function

Searches a character string for a non-alphanumeric character, and returns the first position at which the character is found.

Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

NOTALNUM(*string* <,*start*>)

Arguments

string
specifies a character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTALNUM function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System

Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The NOTALNUM function searches a string for the first occurrence of any character that is not a digit or an uppercase or lowercase letter. If such a character is found, NOTALNUM returns the position in the string of that character. If no such character is found, NOTALNUM returns a value of 0.

If you use only one argument, NOTALNUM begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTALNUM returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTALNUM function searches a character string for a non-alphanumeric character. The ANYALNUM function searches a character string for an alphanumeric character.

Examples

The following example uses the NOTALNUM function to search a string from left to right for non-alphanumeric characters.

```
data _null_;
   string='Next = Last + 1;';
   j=0;
   do until(j=0);
      j=notalnum(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=5 c=
   j=6 c==
   j=7 c=
   j=12 c=
   j=13 c=+
   j=14 c=
   j=16 c=;
   That's all
```

See Also

Function:

"ANYALNUM Function" on page 379

NOTALPHA Function

Searches a character string for a nonalphabetic character, and returns the first position at which the character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTALPHA(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTALPHA function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in the *SAS National Language Support (NLS): Reference Guide*.

The NOTALPHA function searches a string for the first occurrence of any character that is not an uppercase or lowercase letter. If such a character is found, NOTALPHA returns the position in the string of that character. If no such character is found, NOTALPHA returns a value of 0.

If you use only one argument, NOTALPHA begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTALPHA returns a value of zero when one of the following is true:

- The character that you are searching for is not found.

- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTALPHA function searches a character string for a nonalphabetic character. The ANYALPHA function searches a character string for an alphabetic character.

Examples

Example 1: Searching a String for Nonalphabetic Characters The following example uses the NOTALPHA function to search a string from left to right for nonalphabetic characters.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notalpha(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=5 c=
   j=6 c==
   j=7 c=
   j=8 c=_
   j=10 c=_
   j=11 c=
   j=12 c=+
   j=13 c=
   j=14 c=1
   j=15 c=2
   j=17 c=3
   j=18 c=;
   That's all
```

Example 2: Identifying Control Characters by Using the NOTALPHA Function You can execute the following program to show the control characters that are identified by the NOTALPHA function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   notalpha=notalpha(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"ANYALPHA Function" on page 381

NOTCNTRL Function

Searches a character string for a character that is not a control character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTCNTRL(*string*<,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTCNTRL function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in the *SAS National Language Support (NLS): Reference Guide*.

The NOTCNTRL function searches a string for the first occurrence of a character that is not a control character. If such a character is found, NOTCNTRL returns the position in the string of that character. If no such character is found, NOTCNTRL returns a value of 0.

If you use only one argument, NOTCNTRL begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTCNTRL returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTCNTRL function searches a character string for a character that is not a control character. The ANYCNTRL function searches a character string for a control character.

Examples

You can execute the following program to show the control characters that are identified by the NOTCNTRL function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   notcntrl=notcntrl(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"ANYCNTRL Function" on page 383

NOTDIGIT Function

Searches a character string for any character that is not a digit, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTDIGIT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTDIGIT function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in the *SAS National Language Support (NLS): Reference Guide*.

The NOTDIGIT function searches a string for the first occurrence of any character that is not a digit. If such a character is found, NOTDIGIT returns the position in the string of that character. If no such character is found, NOTDIGIT returns a value of 0.

If you use only one argument, NOTDIGIT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTDIGIT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTDIGIT function searches a character string for any character that is not a digit. The ANYDIGIT function searches a character string for a digit.

Examples

The following example uses the NOTDIGIT function to search for a character that is not a digit.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notdigit(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=5 c=
   j=6 c==
   j=7 c=
   j=8 c=_
   j=9 c=n
   j=10 c=_
   j=11 c=
   j=12 c=+
   j=13 c=
   j=16 c=E
   j=18 c=;
   That's all
```

See Also

Function:

"ANYDIGIT Function" on page 384

NOTE Function

Returns an observation ID for the current observation of a SAS data set.

Category: SAS File I/O

Syntax

NOTE(*data-set-id*)

Arguments

data-set-id
is a numeric variable that specifies the data set identifier that the OPEN function returns.

Details

You can use the observation ID value to return to the current observation by using POINT. Observations can be marked by using NOTE and then returned to later by using POINT. Each observation ID is a unique numeric value.

To free the memory that is associated with an observation ID, use DROPNOTE.

Examples

This example calls CUROBS to display the observation number, calls NOTE to mark the observation, and calls POINT to point to the observation that corresponds to NOTEID.

```
%let dsid=%sysfunc(open(sasuser.fitness,i));
  /* Go to observation 10 in data set */
%let rc=%sysfunc(fetchobs(&dsid,10));
%if %sysfunc(abs(&rc)) %then
  %put FETCHOBS FAILED;
%else
  %do;
      /* Display observation number     */
      /* in the Log                     */
    %let cur=%sysfunc(curobs(&dsid));
    %put CUROBS=&cur;
      /* Mark observation 10 using NOTE */
    %let noteid=%sysfunc(note(&dsid));
      /* Rewind pointer to beginning    */
      /* of data                        */
      /* set using REWIND               */
    %let rc=%sysfunc(rewind(&dsid));
      /* FETCH first observation into DDV */
    %let rc=%sysfunc(fetch(&dsid));
      /* Display first observation number */
    %let cur=%sysfunc(curobs(&dsid));
    %put CUROBS=&cur;
```

```
      /* POINT to observation 10 marked  */
      /* earlier by NOTE                 */
    %let rc=%sysfunc(point(&dsid,&noteid));
      /* FETCH observation into DDV */
    %let rc=%sysfunc(fetch(&dsid));
      /* Display observation number 10   */
      /* marked by NOTE                  */
    %let cur=%sysfunc(curobs(&dsid));
    %put CUROBS=&cur;
  %end;
%if (&dsid > 0) %then
  %let rc=%sysfunc(close(&dsid));
```

The output produced by this program is:

```
CUROBS=10
CUROBS=1
CUROBS=10
```

See Also

Functions:

"DROPNOTE Function" on page 667
"OPEN Function" on page 980
"POINT Function" on page 1010
"REWIND Function" on page 1096

NOTFIRST Function

Searches a character string for an invalid first character in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTFIRST(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The NOTFIRST function does not depend on the TRANTAB, ENCODING, or LOCALE options.

The NOTFIRST function searches a string for the first occurrence of any character that is not valid as the first character in a SAS variable name under VALIDVARNAME=V7. These characters are any except the underscore (_) and uppercase or lowercase English letters. If such a character is found, NOTFIRST returns the position in the string of that character. If no such character is found, NOTFIRST returns a value of 0.

If you use only one argument, NOTFIRST begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTFIRST returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTFIRST function searches a string for the first occurrence of any character that is not valid as the first character in a SAS variable name under VALIDVARNAME=V7. The ANYFIRST function searches a string for the first occurrence of any character that is valid as the first character in a SAS variable name under VALIDVARNAME=V7.

Examples

The following example uses the NOTFIRST function to search a string for any character that is not valid as the first character in a SAS variable name under VALIDVARNAME=V7.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notfirst(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
j=5 c=
j=6 c==
j=7 c=
j=11 c=
j=12 c=+
j=13 c=
j=14 c=1
j=15 c=2
j=17 c=3
j=18 c=;
That's all
```

See Also

Function:

"ANYFIRST Function" on page 386

NOTGRAPH Function

Searches a character string for a non-graphical character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTGRAPH(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTGRAPH function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The NOTGRAPH function searches a string for the first occurrence of a non-graphical character. A graphical character is defined as any printable character other than white space. If such a character is found, NOTGRAPH returns the position in the string of that character. If no such character is found, NOTGRAPH returns a value of 0.

If you use only one argument, NOTGRAPH begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTGRAPH returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTGRAPH function searches a character string for a non-graphical character. The ANYGRAPH function searches a character string for a graphical character.

Examples

Example 1: Searching a String for Non-Graphical Characters The following example uses the NOTGRAPH function to search a string for a non-graphical character.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notgraph(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
```

```
   end;
run;
```

The following lines are written to the SAS log:

```
j=5 c=
j=7 c=
j=11 c=
j=13 c=
That's all
```

Example 2: Identifying Control Characters by Using the NOTGRAPH Function You can execute the following program to show the control characters that are identified by the NOTGRAPH function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   notgraph=notgraph(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"ANYGRAPH Function" on page 388

NOTLOWER Function

Searches a character string for a character that is not a lowercase letter, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTLOWER(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTLOWER function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The NOTLOWER function searches a string for the first occurrence of any character that is not a lowercase letter. If such a character is found, NOTLOWER returns the position in the string of that character. If no such character is found, NOTLOWER returns a value of 0.

If you use only one argument, NOTLOWER begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTLOWER returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTLOWER function searches a character string for a character that is not a lowercase letter. The ANYLOWER function searches a character string for a lowercase letter.

Examples

The following example uses the NOTLOWER function to search a string for any character that is not a lowercase letter.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notlower(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
j=1 c=N
j=5 c=
j=6 c==
j=7 c=
j=8 c=_
j=10 c=_
j=11 c=
j=12 c=+
j=13 c=
j=14 c=1
j=15 c=2
j=16 c=E
j=17 c=3
j=18 c=;
That's all
```

See Also

Function:

"ANYLOWER Function" on page 390

NOTNAME Function

Searches a character string for an invalid character in a SAS variable name under VALIDVARNAME=V7, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTNAME(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The NOTNAME function does not depend on the TRANTAB, ENCODING, or LOCALE options.

The NOTNAME function searches a string for the first occurrence of any character that is not valid in a SAS variable name under VALIDVARNAME=V7. These characters are any except underscore (_), digits, and uppercase or lowercase English letters. If such a character is found, NOTNAME returns the position in the string of that character. If no such character is found, NOTNAME returns a value of 0.

If you use only one argument, NOTNAME begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTNAME returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTNAME function searches a string for the first occurrence of any character that is not valid in a SAS variable name under VALIDVARNAME=V7. The ANYNAME function searches a string for the first occurrence of any character that is valid in a SAS variable name under VALIDVARNAME=V7.

Examples

The following example uses the NOTNAME function to search a string for any character that is not valid in a SAS variable name under VALIDVARNAME=V7.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notname(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
j=5 c=
j=6 c==
j=7 c=
j=11 c=
j=12 c=+
j=13 c=
j=18 c=;
That's all
```

See Also

Function:
"ANYNAME Function" on page 392

NOTPRINT Function

Searches a character string for a nonprintable character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTPRINT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTPRINT function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The NOTPRINT function searches a string for the first occurrence of a non-printable character. If such a character is found, NOTPRINT returns the position in the string of that character. If no such character is found, NOTPRINT returns a value of 0.

If you use only one argument, NOTPRINT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTPRINT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTPRINT function searches a character string for a non-printable character. The ANYPRINT function searches a character string for a printable character.

Examples

You can execute the following program to show the control characters that are identified by the NOTPRINT function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   notprint=notprint(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"ANYPRINT Function" on page 394

NOTPUNCT Function

Searches a character string for a character that is not a punctuation character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTPUNCT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTPUNCT function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The NOTPUNCT function searches a string for the first occurrence of a character that is not a punctuation character. If such a character is found, NOTPUNCT returns the position in the string of that character. If no such character is found, NOTPUNCT returns a value of 0.

If you use only one argument, NOTPUNCT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTPUNCT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTPUNCT function searches a character string for a character that is not a punctuation character. The ANYPUNCT function searches a character string for a punctuation character.

Examples

Example 1: Searching a String for Characters That Are Not Punctuation Characters The following example uses the NOTPUNCT function to search a string for characters that are not punctuation characters.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notpunct(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=5 c=
   j=7 c=
   j=9 c=n
   j=11 c=
   j=13 c=
   j=14 c=1
   j=15 c=2
   j=16 c=E
   j=17 c=3
   That's all
```

Example 2: Identifying Control Characters by Using the NOTPUNCT Function You can execute the following program to show the control characters that are identified by the NOTPUNCT function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
   notpunct=notpunct(byte);
   output;
 end;

 proc print data=test;
 run;
```

See Also

Function:

"ANYPUNCT Function" on page 396

NOTSPACE Function

Searches a character string for a character that is not a white-space character (blank, horizontal and vertical tab, carriage return, line feed, and form feed), and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTSPACE(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTSPACE function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The NOTSPACE function searches a string for the first occurrence of a character that is not a blank, horizontal tab, vertical tab, carriage return, line feed, or form feed. If such a character is found, NOTSPACE returns the position in the string of that character. If no such character is found, NOTSPACE returns a value of 0.

If you use only one argument, NOTSPACE begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTSPACE returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTSPACE function searches a character string for the first occurrence of a character that is not a blank, horizontal tab, vertical tab, carriage return, line feed, or form feed. The ANYSPACE function searches a character string for the first occurrence of a character that is a blank, horizontal tab, vertical tab, carriage return, line feed, or form feed.

Examples

Example 1: Searching a String for a Character That Is Not a White-Space Character The following example uses the NOTSPACE function to search a string for a character that is not a white-space character.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notspace(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=6 c==
   j=8 c=_
   j=9 c=n
   j=10 c=_
   j=12 c=+
   j=14 c=1
   j=15 c=2
   j=16 c=E
   j=17 c=3
   j=18 c=;
   That's all
```

Example 2: Identifying Control Characters by Using the NOTSPACE Function You can execute the following program to show the control characters that are identified by the NOTSPACE function.

```
data test;
do dec=0 to 255;
   byte=byte(dec);
   hex=put(dec,hex2.);
```

```
  notspace=notspace(byte);
  output;
end;

proc print data=test;
run;
```

See Also

Function:
"ANYSPACE Function" on page 397

NOTUPPER Function

Searches a character string for a character that is not an uppercase letter, and returns the first position at which that character is found.

Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

NOTUPPER(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The results of the NOTUPPER function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

The NOTUPPER function searches a string for the first occurrence of a character that is not an uppercase letter. If such a character is found, NOTUPPER returns the position in the string of that character. If no such character is found, NOTUPPER returns a value of 0.

If you use only one argument, NOTUPPER begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTUPPER returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTUPPER function searches a character string for a character that is not an uppercase letter. The ANYUPPER function searches a character string for an uppercase letter.

Examples

The following example uses the NOTUPPER function to search a string for any character that is not an uppercase letter.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notupper(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=2 c=e
   j=3 c=x
   j=4 c=t
   j=5 c=
   j=6 c==
   j=7 c=
   j=8 c=_
   j=9 c=n
   j=10 c=_
   j=11 c=
   j=12 c=+
   j=13 c=
   j=14 c=1
   j=15 c=2
   j=17 c=3
   j=18 c=;
   That's all
```

See Also

Function:

"ANYUPPER Function" on page 399

NOTXDIGIT Function

Searches a character string for a character that is not a hexadecimal character, and returns the first position at which that character is found.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

NOTXDIGIT(*string* <,*start*>)

Arguments

string
is the character constant, variable, or expression to search.

start
is an optional numeric constant, variable, or expression with an integer value that specifies the position at which the search should start and the direction in which to search.

Details

The NOTXDIGIT function searches a string for the first occurrence of any character that is not a digit or an uppercase or lowercase A, B, C, D, E, or F. If such a character is found, NOTXDIGIT returns the position in the string of that character. If no such character is found, NOTXDIGIT returns a value of 0.

If you use only one argument, NOTXDIGIT begins the search at the beginning of the string. If you use two arguments, the absolute value of the second argument, *start*, specifies the position at which to begin the search. The direction in which to search is determined in the following way:

- If the value of *start* is positive, the search proceeds to the right.
- If the value of *start* is negative, the search proceeds to the left.
- If the value of *start* is less than the negative length of the string, the search begins at the end of the string.

NOTXDIGIT returns a value of zero when one of the following is true:

- The character that you are searching for is not found.
- The value of *start* is greater than the length of the string.
- The value of *start* = 0.

Comparisons

The NOTXDIGIT function searches a character string for a character that is not a hexadecimal character. The ANYXDIGIT function searches a character string for a character that is a hexadecimal character.

Examples

The following example uses the NOTXDIGIT function to search a string for a character that is not a hexadecimal character.

```
data _null_;
   string='Next = _n_ + 12E3;';
   j=0;
   do until(j=0);
      j=notxdigit(string,j+1);
      if j=0 then put +3 "That's all";
      else do;
         c=substr(string,j,1);
         put +3 j= c=;
      end;
   end;
run;
```

The following lines are written to the SAS log:

```
   j=1 c=N
   j=3 c=x
   j=4 c=t
   j=5 c=
   j=6 c==
   j=7 c=
   j=8 c=_
   j=9 c=n
   j=10 c=_
   j=11 c=
   j=12 c=+
   j=13 c=
   j=18 c=;
   That's all
```

See Also

Function:

"ANYXDIGIT Function" on page 401

NPV Function

Returns the net present value with the rate expressed as a percentage.

Category: Financial

Syntax

NPV(*r,freq,c0,c1,...,cn*)

Arguments

r
is numeric, the interest rate over a specified base period of time expressed as a percentage.

freq
is numeric, the number of payments during the base period of time specified with the rate *r*.

Range: *freq* > 0

Exception: The case *freq* = 0 is a flag to allow continuous discounting.

c0,c1,...,cn
are numeric cash flows that represent cash outlays (payments) or cash inflows (income) occurring at times 0, 1, ...n. These cash flows are assumed to be equally spaced, beginning-of-period values. Negative values represent payments, positive values represent income, and values of 0 represent no cash flow at a given time. The *c0* argument and the *c1* argument are required.

Comparisons

The NPV function is identical to NETPV, except that the *r* argument is provided as a percentage.

NVALID Function

Checks the validity of a character string for use as a SAS variable name.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

NVALID(*string*<,*validvarname*>)

Arguments

string
specifies a character constant, variable, or expression which will be checked to determine whether its value can be used as a SAS variable name.

Note: Trailing blanks are ignored. △

Tip: Enclose a literal string of characters in quotation marks.

validvarname
is a character constant, variable, or expression that specifies one of the following values:

V7 determines that *string* is a valid SAS variable name when all three of the following are true:

- It begins with an English letter or an underscore.
- All subsequent characters are English letters, underscores, or digits.
- The length is 32 or fewer alphanumeric characters.

ANY determines that *string* is a valid SAS variable name if it contains 32 or fewer characters of any type, including blanks.

NLITERAL determines that *string* is a valid SAS variable name if it is in the form of a SAS name literal ('name'N) or if it is a valid SAS variable name when VALIDVARNAME=V7.

See: V7 above in this same list.

Default: If no value is specified, the NVALID function determines that *string* is a valid SAS variable name based on the value of the SAS system option VALIDVARNAME=.

Details

The NVALID function checks the value of *string* to determine whether it can be used as a SAS variable name.

The NVALID function returns a value of 1 or 0.

Condition	Returned Value
string can be used as a SAS variable name	1
string cannot be used as a SAS variable name	0

Examples

This example determines the validity of specified strings as SAS variable names. The value that is returned by the NVALID function varies with the validvarname argument. The value of 1 is returned when the string is determined to be a valid SAS variable

name under the rules for the specified validvarname argument. Otherwise, the value of 0 is returned.

```
options validvarname=v7 ls=64;
data string;
   input string $char40.;
   v7=nvalid(string,'v7');
   any=nvalid(string,'any');
   nliteral=nvalid(string,'nliteral');
   default=nvalid(string);
   datalines;
Tooooooooooooooooooooooooo Long

OK
Very_Long_But_Still_OK_for_V7
1st_char_is_a_digit
Embedded blank
!@#$%^&*
"Very Loooong N-Literal with """N
'No closing quotation mark
;

proc print noobs;
title1 'NLITERAL and Validvarname Arguments Determine';
title2 'Invalid (0) and Valid (1) SAS Variable Names';
run;
```

Output 4.62 Determining the Validity of SAS Variable Names with NLITERAL

```
       NLITERAL and Validvarname Arguments Determine          1
         Invalid (0) and Valid (1) SAS Variable Names

string                               v7  any  nliteral  default

Tooooooooooooooooooooooooo Long       0   0       0        0
                                      0   0       0        0
OK                                    1   1       1        1
Very_Long_But_Still_OK_for_V7         1   1       1        1
1st_char_is_a_digit                   0   1       1        0
Embedded blank                        0   1       1        0
!@#$%^&*                              0   1       1        0
"Very Loooong N-Literal with """N     0   0       1        0
'No closing quotation mark            0   1       0        0
```

See Also

Functions:

"COMPARE Function" on page 591

"NLITERAL Function" on page 945

System Option:

"VALIDVARNAME= System Option" on page 2049

"Rules for Words and Names in the SAS Language" in *SAS Language Reference: Concepts*

NWKDOM Function

Returns the date for the *n*th occurrence of a weekday for the specified month and year.

Category: Date and Time

Syntax

NWKDOM(*n*, *weekday*, *month*, *year*)

Arguments

n
specifies the numeric week of the month that contains the specified day.

Range: 1–5

Tip: *N*=5 indicates that the specified day occurs in the last week of that month. Sometimes *n*=4 and *n*=5 produce the same results.

weekday
specifies the number that corresponds to the day of the week.

Range: 1–7

Tip: Sunday is considered the first day of the week and has a *weekday* value of 1.

month
specifies the number that corresponds to the month of the year.

Range: 1–12

year
specifies a four-digit calendar year.

Details

The NWKDOM function returns a SAS date value for the *n*th weekday of the month and year that you specify. Use any valid SAS date format, such as the DATE9. format, to display a calendar date. You can specify *n*=5 for the last occurrence of a particular weekday in the month.

Sometimes *n*=5 and *n*=4 produce the same result. These results occur when there are only four occurrences of the requested weekday in the month. For example, if the month of January begins on a Sunday, there will be five occurrences of Sunday, Monday, and Tuesday, but only four occurrences of Wednesday, Thursday, Friday, and Saturday. In this case, specifying *n*=5 or *n*=4 for Wednesday, Thursday, Friday, or Saturday will produce the same result.

If February is not a leap year, the month has 28 days and there are four occurrences of each day of the week. In this case, *n*=5 and *n*=4 produce the same results for every day.

Comparisons

In the NWKDOM function, the value for *weekday* corresponds to the numeric day of the week beginning on Sunday. This value is the same value that is used in the WEEKDAY function, where Sunday =1, and so on. The value for *month* corresponds to the numeric

month of the year beginning in January. This value is the same value that is used in the MONTH function, where January =1, and so on.

You can use the NWKDOM function to calculate events that are not defined by the HOLIDAY function. For example, if a university always schedules graduation on the first Saturday in June, then you can use the following statement to calculate the date:

```
UnivGrad = nwkdom(1, 7, 6, year);
```

Examples

Example 1: Returning Date Values The following example uses the NWKDOM function and returns the date for specific occurrences of a weekday for a specified month and year.

```
data _null_;
      /* Return the date of the third Monday in May 2000. */
   a=nwkdom(3, 2, 5, 2000);
      /* Return the date of the fourth Wednesday in November 2007. */
   b=nwkdom(4, 4, 11, 2007);
      /* Return the date of the fourth Saturday in November 2007. */
   c=nwkdom(4, 7, 11, 2007);
      /* Return the date of the first Sunday in January 2007. */
   d=nwkdom(1, 1, 1, 2007);
      /* Return the date of the second Tuesday in September 2007. */
   e=nwkdom(2, 3, 9, 2007);
      /* Return the date of the fifth Thursday in December 2007. */
   f=nwkdom(5, 5, 12, 2007);
   put a= weekdatx.;
   put b= weekdatx.;
   put c= weekdatx.;
   put d= weekdatx.;
   put e= weekdatx.;
   put f= weekdatx.;
run;
```

Output 4.63 Output from Returning Date Values

```
a=Monday, 15 May 2000
b=Wednesday, 28 November 2007
c=Saturday, 24 November 2007
d=Sunday, 7 January 2007
e=Tuesday, 11 September 2007
f=Thursday, 27 December 2007
```

Example 2: Returning the Date of the Last Monday in May The following example returns the date that corresponds to the last Monday in the month of May in the year 2007.

```
data _null_;
      /* The last Monday in May. */
   x=nwkdom(5,2,5,2007);
   put x date9.;
run;
```

Output 4.64 Output from Calculating the Date of the Last Monday in May

```
28MAY2007
```

See Also

Functions:

"HOLIDAY Function" on page 804
"INTNX Function" on page 848
"MONTH Function" on page 936
"WEEKDAY Function" on page 1229

OPEN Function

Opens a SAS data set.

Category: SAS File I/O

Syntax

OPEN(<*data-set-name* <,*mode* <,*generation-number* <,*type*>>>>)

Arguments

data-set-name
is a character constant, variable, or expression that specifies the name of the SAS data set or SAS SQL view to be opened. The value of this character string should be of the form

<*libref.*>*member-name*<(*data-set-options*)>

Default: The default value for *data-set-name* is _LAST_.

Restriction: If you specify the FIRSTOBS= and OBS= data set options, they are ignored. All other data set options are valid.

mode
is a character constant, variable, or expression that specifies the type of access to the data set:

I opens the data set in INPUT mode (default). Values can be read but not modified. **'I'** uses the strongest access mode available in the engine. That is, if the engine supports random access, OPEN defaults to random access. Otherwise, the file is opened in **'IN'** mode automatically. Files are opened with sequential access and a system level warning is set.

IN opens the data set in INPUT mode. Observations are read sequentially, and you are allowed to revisit an observation.

IS opens the data set in INPUT mode. Observations are read sequentially, but you are not allowed to revisit an observation.

Default: I

generation-number
specifies a consistently increasing number that identifies one of the historical versions in a generation group.

Tip: The *generation-number* argument is ignored if *type* = F.

type
is a character constant and can be one of the following values:

D specifies that the first argument, *data-set-name*, is a one-level or two-level data set name.

The following example shows how the D *type* value can be used:

```
rc = open('lib.mydata', , , 'D');
```

Tip: D is the default if there is no fourth argument.

F specifies that the first argument, *data-set-name*, is a filename, a physical path to a file.

The following examples show how the F *type* value can be used:

```
rc = open('c:\data\mydata.sas7bdat', , , 'F');
rc = open('c:\data\mydata', , , 'F');
```

Tip: If you use the F value, then the third argument, *generation-number*, is ignored.

Note: If an argument is invalid, OPEN returns 0. You can obtain the text of the corresponding error message from the SYSMSG function. Invalid arguments do not produce a message in the SAS log and do not set the _ERROR_ automatic variable. △

Details

The OPEN function opens a SAS data set, DATA step, or a SAS SQL view and returns a unique numeric data set identifier, which is used in most other data set access functions. OPEN returns 0 if the data set could not be opened.

If you call the OPEN function from a macro, then the result of the call is valid only when the result is passed to functions in a macro. If you call the OPEN function from the DATA step, then the result is valid only when the result is passed to functions in the same DATA step.

By default, a SAS data set is opened with a control level of RECORD. For details, see the "CNTLLEV= Data Set Option" on page 18 . An open SAS data set should be closed when it is no longer needed. If you open a data set within a DATA step, it will be closed automatically when the DATA step ends.

OPEN defaults to the strongest access mode available in the engine. That is, if the engine supports random access, OPEN defaults to random access. Otherwise, data sets are opened with sequential access, and a system-level warning is set.

Examples

- This example opens the data set PRICES in the library MASTER using INPUT mode. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let dsid=%sysfunc(open(master.prices,i));
%if (&dsid = 0) %then
     %put %sysfunc(sysmsg());
%else
     %put PRICES data set has been opened;
```

- This example passes values from macro or DATA step variables to be used on data set options. It opens the data set SASUSER.HOUSES, and uses the WHERE= data set option to apply a permanent WHERE clause. Note that in a macro statement you do not enclose character strings in quotation marks.

```
%let choice = style="RANCH";
%let dsid=%sysfunc(open(sasuser.houses
                        (where=(&choice)),i));
```

- This example shows how to check the returned value for errors and to write an error message from the SYSMSG function.

```
data _null_;
   d=open('bad','?');
   if not d then do;
      m=sysmsg();
      put m;
      abort;
   end;
   ... more SAS statements ...;
run;
```

See Also

Functions:

"CLOSE Function" on page 583

"SYSMSG Function" on page 1154

ORDINAL Function

Returns the k*th* smallest of the missing and nonmissing values.

Category: Descriptive Statistics

Syntax

ORDINAL(*k,argument-1,argument-2*<,...*argument-n*>)

Arguments

k
is a numeric constant, variable, or expression with an integer value that is less than or equal to the number of subsequent elements in the list of arguments.

argument
specifies a numeric constant, variable, or expression. At least two arguments are required. An argument can consist of a variable list, preceded by OF.

Details

The ORDINAL function returns the k*th* smallest value, either missing or nonmissing, among the second through the last arguments.

Comparisons

The ORDINAL function counts both missing and nonmissing values, whereas the SMALLEST function counts only nonmissing values.

Examples

SAS Statements	Results
`x1=ordinal(4,1,2,3,-4,5,6,7);`	`3`

PATHNAME Function

Returns the physical name of an external file or a SAS library, or returns a blank.

Category: SAS File I/O

Category: External Files

See: PATHNAME Function in the documentation for your operating environment.

Syntax

PATHNAME((*fileref* | *libref*) <,*search-ref*>)

Arguments

fileref
is a character constant, variable, or expression that specifies the fileref that is assigned to an external file.

libref
is a character constant, variable, or expression that specifies the libref that is assigned to a SAS library.

search-ref
is a character constant, variable, or expression that specifies whether to search for a fileref or a libref.

F	specifies a search for a fileref.
L	specifies a search for a libref.

Details

PATHNAME returns the physical name of an external file or SAS library, or blank if *fileref* or *libref* is invalid.

If the name of a fileref is identical to the name of a libref, you can use the *search-ref* argument to choose which reference you want to search. If you specify a value of F, SAS searches for a fileref. If you specify a value of L, SAS searches for a libref.

If you do not specify a *search-ref* argument, and the name of a fileref is identical to the name of a libref, PATHNAME searches first for a fileref. If a fileref does not exist, PATHNAME then searches for a libref.

The default length of the target variable in the DATA step is 200 characters.

You can assign a fileref to an external file by using the FILENAME statement or the FILENAME function.

You can assign a libref to a SAS library using the LIBNAME statement or the LIBNAME function. Some operating environments allow you to assign a libref using system commands.

Operating Environment Information: Under some operating environments, filerefs can also be assigned by using system commands. For details, see the SAS documentation for your operating environment. △

Examples

This example uses the FILEREF function to verify that the fileref MYFILE is associated with an external file. Then it uses PATHNAME to retrieve the actual name of the external file:

```
data _null_;
   length fname $ 100;
   rc=fileref('myfile');
   if (rc=0) then
   do;
      fname=pathname('myfile');
      put fname=;
   end;
run;
```

See Also

Functions:

"FEXIST Function" on page 686
"FILEEXIST Function" on page 689
"FILENAME Function" on page 690

"FILEREF Function" on page 692

Statements:

"LIBNAME Statement" on page 1656
"FILENAME Statement" on page 1520

PCTL Function

Returns the percentile that corresponds to the percentage.

Category: Descriptive Statistics

Syntax

PCTL<*n*>(*percentage*, *value1*<, *value2, ...*>)

Arguments

n
is a digit from 1 to 5 which specifies the definition of the percentile to be computed.
Default: definition 5

percentage
is a numeric constant, variable, or expression that specifies the percentile to be computed.
Requirement: is numeric where, 0 $\leq$*percentage* $\leq$100.

value
is a numeric variable, constant, or expression.

Details

The PCTL function returns the percentile of the nonmissing values corresponding to the percentage. If *percentage* is missing, less than zero, or greater than 100, the PCTL function generates an error message.

Note: The formula that is used in the PCTL function is the same formula that used in PROC UNIVARIATE. For more information, see "SAS Elementary Statistics Procedures" in *Base SAS Procedures Guide*. △

Examples

SAS Statements	Results
`lower_quartile=PCTL(25,2,4,1,3);` `put lower_quartile;`	`1.5`
`percentile_def2=PCTL2(25,2,4,1,3);` `put percentile_def2;`	`1`
`lower_tertile=PCTL(100/3,2,4,1,3);` `put lower_tertile;`	`2`

SAS Statements	Results
`percentile_def3=PCTL3(100/3,2,4,1,3);` `put percentile_def3;`	`2`
`median=PCTL(50,2,4,1,3);` `put median;`	`2.5`
`upper_tertile=PCTL(200/3,2,4,1,3);` `put upper_tertile;`	`3`
`upper_quartile=PCTL(75,2,4,1,3);` `put upper_quartile;`	`3.5`

PDF Function

Returns a value from a probability density (mass) distribution.

Category: Probability

Alias: PMF

Syntax

PDF (*dist,quantile<,parm-1, ... ,parm-k>*)

Arguments

dist

is a character constant, variable, or expression that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	`BERNOULLI`
Beta	`BETA`
Binomial	`BINOMIAL`
Cauchy	`CAUCHY`
Chi-Square	`CHISQUARE`
Exponential	`EXPONENTIAL`
F	`F`
Gamma	`GAMMA`
Geometric	`GEOMETRIC`
Hypergeometric	`HYPERGEOMETRIC`
Laplace	`LAPLACE`
Logistic	`LOGISTIC`

Distribution	Argument
Lognormal	**LOGNORMAL**
Negative binomial	**NEGBINOMIAL**
Normal	**NORMAL|GAUSS**
Normal mixture	**NORMALMIX**
Pareto	**PARETO**
Poisson	**POISSON**
T	**T**
Uniform	**UNIFORM**
Wald (inverse Gaussian)	**WALD|IGAUSS**
Weibull	**WEIBULL**

Note: Except for T, F, and NORMALMIX, you can minimally identify any distribution by its first four characters. △

quantile
is a numeric constant, variable, or expression that specifies the value of the random variable.

parm-1,...,parm-k
are optional numeric constants, variables, or expressions that specify the values of *shape*, *location*, or *scale* parameters that are appropriate for the specific distribution.

See: "Details" on page 987 for complete information about these parameters

Details

Bernoulli Distribution

PDF('BERNOULLI',*x*,*p*)

where

x
is a numeric random variable.

p
is a numeric probability of success.

Range: $0 \leq p \leq 1$

The PDF function for the Bernoulli distribution returns the probability density function of a Bernoulli distribution, with probability of success equal to p. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('BERN', x, p\right) = \begin{cases} 0 & x < 0 \\ 1-p & x = 0 \\ 0 & 0 < x < 1 \\ p & x = 1 \\ 0 & x > 1 \end{cases}$$

Note: There are no *location* or *scale* parameters for this distribution. △

Beta Distribution

PDF('BETA',*x*,*a*,*b*<,*l*,*r*>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.
Range: $a > 0$

b
is a numeric shape parameter.
Range: $b > 0$

l
is the numeric left location parameter.
Default: 0

r
is the right location parameter.
Default: 0
Range: $r > l$

The PDF function for the beta distribution returns the probability density function of a beta distribution, with shape parameters a and b. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('BETA', x, a, b, l, r\right) = \begin{cases} 0 & x < l \\ \frac{1}{\beta(a,b)} \frac{(x-l)^{a-1}(r-x)^{b-1}}{(r-l)^{a+b-1}} & l \le x \le r \\ 0 & x > r \end{cases}$$

Note: The quantity $\frac{x-l}{r-l}$ is forced to be $\epsilon \le \frac{x-l}{r-l} \le 1 - 2\epsilon$. △

Binomial Distribution

PDF('BINOMIAL',*m*,*p*,*n*)

where

m
is an integer random variable that counts the number of successes.

Range: $m = 0, 1, ...$

p
is a numeric probability of success.
Range: $0 \leq p \leq 1$

n
is an integer parameter that counts the number of independent Bernoulli trials.
Range: $n = 0, 1, ...$

The PDF function for the binomial distribution returns the probability density function of a binomial distribution, with parameters p and n, which is evaluated at the value m. The equation follows:

$$PDF\left('BINOM', m, p, n\right) = \begin{cases} 0 & m < 0 \\ \binom{n}{m} p^m \left(1-p\right)^{n-m} & 0 \leq m \leq n \\ 0 & m > n \end{cases}$$

Note: There are no *location* or *scale* parameters for the binomial distribution. △

Cauchy Distribution

PDF('CAUCHY',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter.
Default: 0

λ
is a numeric scale parameter.
Default: 1
Range: $\lambda > 0$

The PDF function for the Cauchy distribution returns the probability density function of a Cauchy distribution, with the location parameter θ and the scale parameter λ. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('CAUCHY', x, \theta, \lambda\right) = \frac{1}{\pi}\left(\frac{\lambda}{\lambda^2 + \left(x-\theta\right)^2}\right)$$

Chi-Square Distribution

PDF('CHISQUARE',*x*,*df* <,*nc*>)

where

x
is a numeric random variable.

df
is a numeric degrees of freedom parameter.

Range: *df* > 0

nc
is an optional numeric non-centrality parameter.

Range: *nc* ≥ 0

The PDF function for the chi-square distribution returns the probability density function of a chi-square distribution, with *df* degrees of freedom and non-centrality parameter *nc*. The PDF function is evaluated at the value *x*. This function accepts non-integer degrees of freedom. If *nc* is omitted or equal to zero, the value returned is from the central chi-square distribution. The following equation describes the PDF function of the chi–square distribution,

$$PDF\left('CHISQ', x, v, \lambda\right) = \begin{cases} 0 & x < 0 \\ \sum_{j=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^j}{j!} p_c\left(x, v + 2j\right) & x \geq 0 \end{cases}$$

where $p_c(.,.)$ denotes the density from the central chi-square distribution:

$$p_c\left(x, a\right) = \frac{1}{2} p_g\left(\frac{x}{2}, \frac{a}{2}\right)$$

and where $p_g(y,b)$ is the density from the gamma distribution, which is given by

$$p_g\left(y, b\right) = \frac{1}{\Gamma\left(b\right)} e^{-y} y^{b-1}$$

Exponential Distribution

PDF('EXPONENTIAL',*x* <,λ>)

where

x
is a numeric random variable.

λ
is a scale parameter.

Default: 1

Range: $\lambda > 0$

The PDF function for the exponential distribution returns the probability density function of an exponential distribution, with the scale parameter λ. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('EXPO', x, \lambda\right) = \begin{cases} 0 & x < 0 \\ \frac{1}{\lambda} exp\left(-\frac{x}{\lambda}\right) & x \geq 0 \end{cases}$$

F Distribution

PDF('F',*x*,*ndf*,*ddf*<,*nc*>)

where

x
is a numeric random variable.

ndf
is a numeric numerator degrees of freedom parameter.

Range: $ndf > 0$

ddf
is a numeric denominator degrees of freedom parameter.

Range: $ddf > 0$

nc
is a numeric non-centrality parameter.

Range: $nc \geq 0$

The PDF function for the *F* distribution returns the probability density function of an *F* distribution, with *ndf* numerator degrees of freedom, *ddf* denominator degrees of freedom, and non-centrality parameter *nc*. The PDF function is evaluated at the value *x*. This PDF function accepts non-integer degrees of freedom for *ndf* and *ddf*. If *nc* is omitted or equal to zero, the value returned is from a central *F* distribution. In the following equation, let ν_1 = *ndf*, let ν_2 = *ddf*, and let λ = *nc*. The following equation describes the PDF function of the *F* distribution.

$$PDF\left('F', x, v_1, v_2, \lambda\right) = \begin{cases} 0 & x < 0 \\ \sum_{j=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^j}{j!} p_f\left(f, v_1 + 2j, v_2\right) & x \geq 0 \end{cases}$$

where $p_f(f,u_1,u_2)$ is the density from the central *F* distribution with

$$p_f\left(f, u_1, u_2\right) = p_B\left(\frac{u_1 f}{uf + u_2}, \frac{u_1}{2}, \frac{u_2}{2}\right) \frac{u_1 u_2}{\left(u_2 + u_1 f\right)^2}$$

and where $p_B(x,a,b)$ is the density from the standard beta distribution.

Note: There are no *location* or *scale* parameters for the *F* distribution. △

Gamma Distribution

PDF('GAMMA',*x*,*a*<,λ>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.
Range: $a > 0$

λ
is a numeric scale parameter.
Default: 1
Range: $\lambda > 0$

The PDF function for the gamma distribution returns the probability density function of a gamma distribution, with the shape parameter *a* and the scale parameter λ. The PDF function is evaluated at the value *x*. The equation follows:

$$PDF\left('GAMMA', x, a, \lambda\right) = \begin{cases} 0 & x < 0 \\ \frac{1}{\lambda^a \Gamma(a)} x^{a-1} exp\left(-\frac{x}{\lambda}\right) & x \geq 0 \end{cases}$$

Geometric Distribution

PDF('GEOMETRIC',*m,p*)

where

m
is a numeric random variable that denotes the number of failures before the first success.

Range: $m \geq 0$

p
is a numeric probability of success.

Range: $0 \leq p \leq 1$

The PDF function for the geometric distribution returns the probability density function of a geometric distribution, with parameter p. The PDF function is evaluated at the value m. The equation follows:

$$PDF\left('GEOM', m, p\right) = \begin{cases} 0 & m < 0 \\ p\left(1-p\right)^{m} & m \geq 0 \end{cases}$$

Note: There are no *location* or *scale* parameters for this distribution. △

Hypergeometric Distribution

PDF('HYPER',*x,N,R,n*<,*o*>)

where

x
is an integer random variable.

N
is an integer population size parameter.

Range: N = 1, 2, ...

R
is an integer number of items in the category of interest.

Range: R = 0, 1, ..., N

n
is an integer sample size parameter.

Range: n = 1, 2, ..., N

o
is an optional numeric odds ratio parameter.

Range: $o > 0$

The PDF function for the hypergeometric distribution returns the probability density function of an extended hypergeometric distribution, with population size *N*, number of items *R*, sample size *n*, and odds ratio *o*. The PDF function is evaluated at the value *x*. If *o* is omitted or equal to 1, the value returned is from the usual hypergeometric distribution. The equation follows:

$$PDF\left('HYPER', x, N, R, n, o\right) =$$
$$\begin{cases} 0 & x < max\left(0, R+n-N\right) \\ \dfrac{\binom{R}{x}\binom{N-R}{n-x}o^x}{\sum_{j=max(0,R+n-N)}^{min(R,n)}\binom{R}{j}\binom{N-R}{n-j}o^j} & max\left(0, R+n-N\right) \le x \le min\left(R, n\right) \\ 0 & x > min\left(R, n\right) \end{cases}$$

Laplace Distribution

PDF('LAPLACE',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter.
Default: 0

λ
is a numeric scale parameter.
Default: 1
Range: $\lambda > 0$

The PDF function for the Laplace distribution returns the probability density function of the Laplace distribution, with the location parameter θ and the scale parameter λ. The PDF function is evaluated at the value *x*. The equation follows:

$$PDF\left('LAPLACE', x, \theta, \lambda\right) = \frac{1}{2\lambda} exp\left(-\frac{|x-\theta|}{\lambda}\right)$$

Logistic Distribution

PDF('LOGISTIC',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter.
Default: 0

λ
is a numeric scale parameter.
Default: 1

Range: $\lambda > 0$

The PDF function for the logistic distribution returns the probability density function of a logistic distribution, with the location parameter θ and the scale parameter λ. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('LOGISTIC', x, \theta, \lambda\right) = \frac{exp\left(\frac{x-\theta}{\lambda}\right)}{\lambda\left(1 + exp\left(\frac{x-\theta}{\lambda}\right)\right)^2}$$

Lognormal Distribution

PDF('LOGNORMAL',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
specifies a numeric log scale parameter. (exp(θ) is a scale parameter.)

Default: 0

λ
specifies a numeric shape parameter.

Default: 1

Range: $\lambda > 0$

The PDF function for the lognormal distribution returns the probability density function of a lognormal distribution, with the log scale parameter θ and the shape parameter λ. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('LOGN', x, \theta, \lambda\right) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{x\lambda\sqrt{2\pi}}\ exp\left(-\frac{(log(x)-\theta)^2}{2\lambda^2}\right) & x > 0 \end{cases}$$

Negative Binomial Distribution

PDF('NEGBINOMIAL',*m,p,n*)

where

m
is a positive integer random variable that counts the number of failures.

Range: m= 0, 1, ...

p
is a numeric probability of success.

Range: $0 \leq p \leq 1$

n
is a numeric value that counts the number of successes.

Range: $n > 0$

The PDF function for the negative binomial distribution returns the probability density function of a negative binomial distribution, with probability of success p and number of successes n. The PDF function is evaluated at the value m. The equation follows:

$$PDF\left('NEGB', m, p, n\right) = \begin{cases} 0 & m < 0 \\ \binom{n+m-1}{n-1} p^n (1-p)^m & m \geq 0 \end{cases}$$

Note: There are no *location* or *scale* parameters for the negative binomial distribution. △

Normal Distribution

PDF('NORMAL',*x*<,θ,λ>)

where

x
is a numeric random variable.

θ
is a numeric location parameter.
Default: 0

λ
is a numeric scale parameter.
Default: 1
Range: $\lambda > 0$

The PDF function for the normal distribution returns the probability density function of a normal distribution, with the location parameter θ and the scale parameter λ. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('NORMAL', x, \theta, \lambda\right) = \frac{1}{\lambda\sqrt{2\pi}} \; exp\left(-\frac{(x-\theta)^2}{2\lambda^2}\right)$$

Normal Mixture Distribution

PDF('NORMALMIX',*x,n,p,m,s*)

where

x
is a numeric random variable.

n
is the integer number of mixtures.

Range: *n* = 1, 2, ...

p
is the n proportions, $p_1, p_2, \ldots, p_n$, where $\sum_{i=1}^{i=n} p_i = 1$.

Range: *p* = 0, 1, ...

m
is the *n* means $m_1, m_2, \ldots, m_n$.

s
is the *n* standard deviations $s_1, s_2, \ldots, s_n$.

Range: *s* > 0

The PDF function for the normal mixture distribution returns the probability that an observation from a mixture of normal distribution is less than or equal to *x*. The equation follows:

$$PDF\left('NORMALMIX', x, n, p, m, s\right) = \sum_{i=1}^{i=n} p_i \, PDF\left('NORMAL', x, m_i, s_i\right)$$

Note: There are no *location* or *scale* parameters for the normal mixture distribution. △

Pareto Distribution

PDF('PARETO',*x,a*<,*k*>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.

Range: *a* > 0

k
is a numeric scale parameter.

Default: 1

Range: *k* > 0

The PDF function for the Pareto distribution returns the probability density function of a Pareto distribution, with the shape parameter a and the scale parameter k. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('PARETO', x, a, k\right) = \begin{cases} 0 & x < k \\ \frac{a}{k}\left(\frac{k}{x}\right)^{a+1} & x \geq k \end{cases}$$

Poisson Distribution

PDF('POISSON',*n,m*)

where

n
is an integer random variable.

Range: *n*= 0, 1, ...

m
is a numeric mean parameter.

Range: *m* > 0

The PDF function for the Poisson distribution returns the probability density function of a Poisson distribution, with mean m. The PDF function is evaluated at the value n. The equation follows:

$$PDF\left('POISSON', n, m\right) = \begin{cases} 0 & n < 0 \\ e^{-m}\frac{m^n}{n!} & n \geq 0 \end{cases}$$

Note: There are no *location* or *scale* parameters for the Poisson distribution. △

T Distribution

PDF('T',*t,df*<,*nc*>)

where

t
is a numeric random variable.

df
is a numeric degrees of freedom parameter.

Range: *df* > 0

nc
is an optional numeric non-centrality parameter.

The PDF function for the *T* distribution returns the probability density function of a *T* distribution, with degrees of freedom *df* and non-centrality parameter *nc*. The PDF function is evaluated at the value *x*. This PDF function accepts non-integer degrees of freedom. If *nc* is omitted or equal to zero, the value returned is from the central *T* distribution. In the following equation, let ν = *df* and let δ = *nc*.

$$PDF\left('T', t, v, \delta\right) = \frac{1}{2^{\left(\frac{1}{2}v-1\right)}\Gamma\left(\frac{1}{2}v\right)}\int_0^\infty x^{v-1}e^{-\frac{1}{2}x^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{tx}{\sqrt{v}}-\delta\right)^2}\frac{x}{\sqrt{v}}dx$$

Note: There are no *location* or *scale* parameters for the *T* distribution. △

Uniform Distribution

PDF('UNIFORM',*x*<,*l*,*r*>)

where

x
is a numeric random variable.

l
is the numeric left location parameter.

Default: 0

r
is the numeric right location parameter.

Default: 1

Range: $r > l$

The PDF function for the uniform distribution returns the probability density function of a uniform distribution, with the left location parameter l and the right location parameter r. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('UNIFORM', x, l, r\right) = \begin{cases} 0 & x < l \\ \frac{1}{r-l} & l \le x \le r \\ 0 & x > r \end{cases}$$

Wald (Inverse Gaussian) Distribution

PDF('WALD',*x*,*d*)

PDF('IGAUSS',*x*,*d*)

where

x
is a numeric random variable.

d
is a numeric shape parameter.

Range: $d > 0$

The PDF function for the Wald distribution returns the probability density function of a Wald distribution, with shape parameter d, which is evaluated at the value x. The equation follows:

$$PDF\left('WALD', x, d\right) = \begin{cases} 0 & x \le 0 \\ \sqrt{\frac{d}{2\pi x^3}}\ exp\left(-\frac{d}{2}x + d - \frac{d}{2x}\right) & x > 0 \end{cases}$$

Note: There are no *location* or *scale* parameters for the Wald distribution. △

Weibull Distribution

PDF('WEIBULL',*x*,*a*<,λ>)

where

x
is a numeric random variable.

a
is a numeric shape parameter.
Range: $a > 0$

λ
is a numeric scale parameter.
Default: 1
Range: $\lambda > 0$

The PDF function for the Weibull distribution returns the probability density function of a Weibull distribution, with the shape parameter a and the scale parameter λ. The PDF function is evaluated at the value x. The equation follows:

$$PDF\left('WEIBULL', x, a, \lambda\right) = \begin{cases} 0 & x < 0 \\ exp\left(-\left(\frac{x}{\lambda}\right)^{a}\right) \frac{a}{\lambda}\left(\frac{x}{\lambda}\right)^{a-1} & x \geq 0 \end{cases}$$

Examples

SAS Statements	Results
`y=pdf('BERN',0,.25);`	0.75
`y=pdf('BERN',1,.25);`	0.25
`y=pdf('BETA',0.2,3,4);`	1.2288
`y=pdf('BINOM',4,.5,10);`	0.20508
`y=pdf('CAUCHY',2);`	0.063662
`y=pdf('CHISQ',11.264,11);`	0.081686
`y=pdf('EXPO',1);`	0.36788
`y=pdf('F',3.32,2,3);`	0.054027
`y=pdf('GAMMA',1,3);`	0.18394
`y=pdf('HYPER',2,200,50,10);`	0.28685
`y=pdf('LAPLACE',1);`	0.18394
`y=pdf('LOGISTIC',1);`	0.19661
`y=pdf('LOGNORMAL',1);`	0.39894
`y=pdf('NEGB',1,.5,2);`	0.25
`y=pdf('NORMAL',1.96);`	0.058441
`y=pdf('NORMALMIX',2.3,3,.33,.33,.34, .5,1.5,2.5,.79,1.6,4.3);`	0.1166
`y=pdf('PARETO',1,1);`	1
`y=pdf('POISSON',2,1);`	0.18394
`y=pdf('T',.9,5);`	0.24194
`y=pdf('UNIFORM',0.25);`	1

SAS Statements	Results
`y=pdf('WALD',1,2);`	`0.56419`
`y=pdf('WEIBULL',1,2);`	`0.73576`

See Also

Functions:

"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"CDF Function" on page 558
"SDF Function" on page 1120
"QUANTILE Function" on page 1064

PEEK Function

Stores the contents of a memory address in a numeric variable on a 32–bit platform.

Category: Special
Restriction: Use on 32–bit platforms only.

Syntax

PEEK(*address*<,*length*>)

Arguments

address
is a numeric constant, variable, or expression that specifies the memory address.

length
is a numeric constant, variable, or expression that specifies the data length.

Default: a 4-byte address pointer

Range: 2 to 8

Details

If you do not have access to the memory storage location that you are requesting, the PEEK function returns an "Invalid argument" error.

You cannot use the PEEK function on 64-bit platforms. If you attempt to use it, SAS writes a message to the log stating that this restriction applies. If you have legacy applications that use PEEK, change the applications and use PEEKLONG instead. You can use PEEKLONG on both 32–bit and 64–bit platforms.

Comparisons

The PEEK function stores the contents of a memory address into a *numeric* variable. The PEEKC function stores the contents of a memory address into a *character* variable.

Note: SAS recommends that you use PEEKLONG instead of PEEK because PEEKLONG can be used on both 32–bit and 64–bit platforms. △

Examples

The following example, specific to the z/OS operating environment, returns a numeric value that represents the address of the Communication Vector Table (CVT).

```
data _null_;
      /* 16 is the location of the CVT address */
   y=16;
   x=peek(y);
   put 'x= ' x hex8.;
run;
```

See Also

Functions:

"ADDR Function" on page 371

"PEEKC Function" on page 1002

CALL Routine:

"CALL POKE Routine" on page 477

PEEKC Function

Stores the contents of a memory address in a character variable on a 32–bit platform.

Category: Special

Restriction: Use on 32–bit platforms only.

Syntax

PEEKC(*address*<,*length*>)

Arguments

address
is a numeric constant, variable, or expression that specifies the memory address.

length
is a numeric constant, variable, or expression that specifies the data length.

Default: 8, unless the variable length has already been set (by the LENGTH statement, for example)

Range: 1 to 32,767

Details

If you do not have access to the memory storage location that you are requesting, the PEEKC function returns an "Invalid argument" error.

You cannot use the PEEKC function on 64-bit platforms. If you attempt to use it, SAS writes a message to the log stating that this restriction applies. If you have legacy applications that use PEEKC, change the applications and use PEEKCLONG instead. You can use PEEKCLONG on both 32–bit and 64–bit platforms.

Comparisons

The PEEKC function stores the contents of a memory address into a *character* variable. The PEEK function stores the contents of a memory address into a *numeric* variable.

Note: SAS recommends that you use PEEKCLONG instead of PEEKC because PEEKCLONG can be used on both 32–bit and 64–bit platforms. △

Examples

Example 1: Listing ASCB Bytes The following example, specific to the z/OS operating environment, uses both PEEK and PEEKC, and prints the first four bytes of the Address Space Control Block (ASCB).

```
data _null_;
   length y $4;
      /* 220x is the location of the ASCB pointer */
   x=220x;
   y=peekc(peek(x));
   put 'y= ' y;
run;
```

Example 2: Creating a DATA Step View This example, specific to the z/OS operating environment, also uses both the PEEK and PEEKC functions. It creates a DATA step view that accesses the entries in the Task Input Output Table (TIOT). The PRINT procedure is then used to print the entries. Entries in the TIOT include the three components outlined in the following list. In this example, TIOT represents the starting address of the TIOT entry.

TIOT+4 is the ddname. This component takes up 8 bytes.

TIOT+12 is a 3-byte pointer to the Job File Control Block (JFCB).

TIOT+134 is the volume serial number (volser) of the data set. This component takes up 6 bytes.

Here is the program:

```
   /* Create a DATA step view of the contents    */
   /* of the TIOT. The code steps through each   */
   /* TIOT entry to extract the ddname, JFCB,    */
   /* and volser of each ddname that has been    */
   /* allocated for the current task. The data   */
   /* set name is also extracted from the JFCB.  */

data save.tiot/view=save.tiot;
   length ddname $8 volser $6 dsname $44;
      /* Get the TCB (Task Control Block)address */
      /* from the PSATOLD variable in the PSA    */
      /* (Prefixed Save Area).  The address of   */
      /* the PSA is 21CX. Add 12 to the address  */
      /* of the TCB to get the address of the    */
      /* TIOT. Add 24 to bypass the 24-byte      */
      /* header, so that TIOTVAR represents the  */
      /* start of the TIOT entries.              */

   tiotvar=peek(peek(021CX)+12)+24;

      /* Loop through all TIOT entries until the */
      /* TIOT entry length (indicated by the     */
      /* value of the first byte) is 0.          */

   do while(peek(tiotvar,1));

      /* Check to see whether the current TIOT   */
      /* entry is a freed TIOT entry (indicated  */
      /* by the high order bit of the second     */
      /* byte of the TIOT entry). If it is not   */
      /* freed, then proceed.                    */

      if peek(tiotvar+1,1)NE'l.......'B then do;
         ddname=peekc(tiotvar+4);
         jfcb=peek(tiotvar+12,3);
         volser=peekc(jfcb+134);
            /* Add 16 to the JFCB value to get  */
            /* the data set name.  The data set */
            /* name is 44 bytes.                */

         dsname=peekc(jfcb+16);
```

```
         output;
         end;

         /* Increment the TIOTVAR value to point */
         /* to the next TIOT entry. This is done */
         /* by adding the length of the current  */
         /* TIOT entry (indicated by first byte  */
         /* of the entry) to the current value   */
         /* of TIOTVAR.                          */

      tiotvar+peek(tiotvar,1);
   end;

      /* The final DATA step view does not       */
      /* contain the TIOTVAR and JFCB variables. */

   keep ddname volser dsname;
run;

     /* Print the TIOT entries.                  */
proc print data=save.tiot uniform width=minimum;
run;
```

In the PROC PRINT statement, the UNIFORM option ensures that each page of the output is formatted exactly the same way. WIDTH=MINIMUM causes the PRINT procedure to use the minimum column width for each variable on the page. The column width is defined by the longest data value in that column.

See Also

CALL Routine:

"CALL POKE Routine" on page 477

Functions:

"ADDR Function" on page 371
"PEEK Function" on page 1001

PEEKCLONG Function

Stores the contents of a memory address in a character variable on 32-bit and 64-bit platforms.

Category: Special

See: PEEKCLONG Function in the documentation for your operating environment.

Syntax

PEEKCLONG(*address*<,*length*>)

Arguments

address
specifies a character constant, variable, or expression that contains the binary pointer address.

length
is a numeric constant, variable, or expression that specifies the length of the character data.

Default: 8

Range: 1 to 32,767

Details

If you do not have access to the memory storage location that you are requesting, the PEEKCLONG function returns an "Invalid argument" error.

Comparisons

The PEEKCLONG function stores the contents of a memory address in a *character* variable.

The PEEKLONG function stores the contents of a memory address in a *numeric* variable. It assumes that the input address refers to an integer in memory.

Examples

Example 1: Example for a 32-bit Platform The following example returns the pointer address for the character variable Z.

```
data _null_;
   x='ABCDE';
   y=addrlong(x);
   z=peekclong(y,2);
   put z=;
run;
```

The output from the SAS log is: **z=AB**

Example 2: Example for a 64-bit Platform The following example, specific to the z/OS operating environment, returns the pointer address for the character variable Y.

```
data _null_;
   length y $4;
   x220addr=put(220x,pib4.);
   ascb=peeklong(x220addr);
   ascbaddr=put(ascb,pib4.);
   y=peekclong(ascbaddr);
run;
```

The output from the SAS log is: **y='ASCB'**

See Also

Function:

"PEEKLONG Function" on page 1007

PEEKLONG Function

Stores the contents of a memory address in a numeric variable on 32-bit and 64-bit platforms.

Category: Special

See: PEEKLONG Function in the documentation for your operating environment

Syntax

PEEKLONG(*address*<,*length*>)

Arguments

address

specifies a character constant, variable, or expression that contains the binary pointer address.

length

is a numeric constant, variable, or expression that specifies the length of the character data.

Default: 4 on 32-bit computers; 8 on 64-bit computers.

Range: 1-4 on 32-bit computers; 1-8 on 64-bit computers.

Details

If you do not have access to the memory storage location that you are requesting, the PEEKLONG function returns an "Invalid argument" error.

Comparisons

The PEEKLONG function stores the contents of a memory address in a *numeric* variable. It assumes that the input address refers to an integer in memory.

The PEEKCLONG function stores the contents of a memory address in a *character* variable. It assumes that the input address refers to character data.

Examples

Example 1: Example for a 32-bit Platform The following example returns the pointer address for the numeric variable Z.

```
data _null_;
   length y $4;
   y=put(1,IB4.);
   addry=addrlong(y);
   z=peeklong(addry,4);
   put z=;
run;
```

The output from the SAS log is: **z=1**

Example 2: Example for a 64-bit Platform The following example, specific to the z/OS operating environment, returns the pointer address for the numeric variable X.

```
data _null_;
   x=peeklong(put(16,pib4.));
   put x=hex8.;
run;
```

The output from the SAS log is: **x=00FCFCB0**

See Also

Function:

"PEEKCLONG Function" on page 1005

PERM Function

Computes the number of permutations of *n* items that are taken *r* at a time.

Category: Combinatorial

Syntax

PERM(*n*<,*r*>)

Arguments

n
is an integer that represents the total number of elements from which the sample is chosen.

r
is an integer value that represents the number of chosen elements. If *r* is omitted, the function returns the factorial of *n*.

Restriction: $r \leq n$

Details

The mathematical representation of the PERM function is given by the following equation:

$$PERM\,(n,r) = \frac{n!}{(n-r)!}$$

with $n \geq 0$, $r \geq 0$, and $n \geq r$.

If the expression cannot be computed, a missing value is returned. For moderately large values, it is sometimes not possible to compute the PERM function.

Examples

SAS Statements	Results
`x=perm(5,1);`	`5`
`x=perm(5);`	`120`
`x=perm(5,2)`	`20`

See Also

Functions:

"COMB Function" on page 590
"FACT Function" on page 678
"LPERM Function" on page 913

POINT Function

Locates an observation that is identified by the NOTE function.

Category: SAS File I/O

Syntax

POINT(*data-set-id*,*note-id*)

Arguments

data-set-id
is a numeric variable that specifies the data set identifier that the OPEN function returns.

note-id
is a numeric variable that specifies the identifier assigned to the observation by the NOTE function.

Details

POINT returns 0 if the operation was successful, ≠0 if it was not successful. POINT prepares the program to read from the SAS data set. The Data Set Data Vector is not updated until a read is done using FETCH or FETCHOBS.

Examples

This example calls NOTE to obtain an observation ID for the most recently read observation of the SAS data set MYDATA. It calls POINT to point to that observation, and calls FETCH to return the observation marked by the pointer.

```
%let dsid=%sysfunc(open(mydata,i));
%let rc=%sysfunc(fetch(&dsid));
%let noteid=%sysfunc(note(&dsid));
   ...more macro statements...
%let rc=%sysfunc(point(&dsid,&noteid));
%let rc=%sysfunc(fetch(&dsid));
   ...more macro statements...
%let rc=%sysfunc(close(&dsid));
```

See Also

Functions:

"DROPNOTE Function" on page 667
"NOTE Function" on page 956
"OPEN Function" on page 980

POISSON Function

Returns the probability from a Poisson distribution.

Category: Probability

See: "CDF Function" on page 558

Syntax

POISSON(*m,n*)

Arguments

m
is a numeric mean parameter.

Range: $m \geq 0$

n
is an integer random variable.

Range: $n \geq 0$

Details

The POISSON function returns the probability that an observation from a Poisson distribution, with mean *m*, is less than or equal to *n*. To compute the probability that an observation is equal to a given value, *n*, compute the difference of two probabilities from the Poisson distribution for *n* and *n*–1.

Examples

SAS Statements	Results
`x=poisson(1,2);`	`0.9196986029`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBBETA Function

Returns the probability from a beta distribution.

Category: Probability

See: "CDF Function" on page 558

Syntax

PROBBETA(*x,a,b*)

Arguments

x
is a numeric random variable.
Range: $0 \le x \le 1$

a
is a numeric shape parameter.
Range: $a > 0$

b
is a numeric shape parameter.
Range: $b > 0$

Details

The PROBBETA function returns the probability that an observation from a beta distribution, with shape parameters a and b, is less than or equal to x.

Examples

SAS Statements	Results
`x=probbeta(.2,3,4);`	`0.09888`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBBNML Function

Returns the probability from a binomial distribution.

Category: Probability

See: "CDF Function" on page 558, "PDF Function" on page 986

Syntax

PROBBNML(*p*,*n*,*m*)

Arguments

p
is a numeric probability of success parameter.
RANGE: $0 \leq p \leq 1$

n
is an integer number of independent Bernoulli trials parameter.
RANGE: $n > 0$

m
is an integer number of successes random variable.
RANGE: $0 \leq m \leq n$

Details

The PROBBNML function returns the probability that an observation from a binomial distribution, with probability of success p, number of trials n, and number of successes m, is less than or equal to m. To compute the probability that an observation is equal to a given value m, compute the difference of two probabilities from the binomial distribution for m and $m-1$ successes.

Examples

SAS Statements	Results
`x=probbnml(0.5,10,4);`	`0.376953125`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910

"PDF Function" on page 986
"SDF Function" on page 1120

PROBBNRM Function

Returns a probability from a bivariate normal distribution.

Category: Probability

Syntax

PROBBNRM(*x, y, r*)

Arguments

x
specifies a numeric constant, variable, or expression.

y
specifies a numeric constant, variable, or expression.

r
is a numeric correlation coefficient.
Range: $-1 \le r \le 1$

Details

The PROBBNRM function returns the probability that an observation (X, Y) from a standardized bivariate normal distribution with mean 0, variance 1, and a correlation coefficient r, is less than or equal to (x, y). That is, it returns the probability that $X \le x$ and $Y \le y$. The following equation describes the PROBBNRM function, where u and v represent the random variables x and y, respectively:

$$\mathrm{PROBBNRM}(x, y, r) = \frac{1}{2\pi\sqrt{1 - \mathrm{r}^2}} \int_{-\infty}^{x} \int_{-\infty}^{y} \exp\left[-\frac{u^2 - 2ruv + v^2}{2(1 - r^2)} \right] dv\, du$$

Examples

SAS Statements	Result
`p=probbnrm(.4, -.3, .2);` `put p;`	`0.2783183345`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBCHI Function

Returns the probability from a chi-square distribution.

Category: Probability
See: "CDF Function" on page 558

Syntax

PROBCHI(*x*,*df*<,*nc*>)

Arguments

x
is a numeric random variable.
Range: $x \geq 0$

df
is a numeric degrees of freedom parameter.
Range: $df > 0$

nc
is an optional numeric noncentrality parameter.
Range: $nc \geq 0$

Details

The PROBCHI function returns the probability that an observation from a chi-square distribution, with degrees of freedom *df* and noncentrality parameter *nc*, is less than or equal to *x*. This function accepts a noninteger degrees of freedom parameter *df*. If the optional parameter *nc* is not specified or has the value 0, the value returned is from the central chi-square distribution.

Examples

SAS Statements	Results
`x=probchi(11.264,11);`	`0.5785813293`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBF Function

Returns the probability from an *F* distribution.

Category: Probability
See: "CDF Function" on page 558

Syntax

PROBF(*x,ndf,ddf*<,nc>)

Arguments

x
is a numeric random variable.
Range: $x \geq 0$

ndf
is a numeric numerator degrees of freedom parameter.
Range: $ndf > 0$

ddf
is a numeric denominator degrees of freedom parameter.
Range: $ddf > 0$

nc
is an optional numeric noncentrality parameter.
Range: $nc \geq 0$

Details

The PROBF function returns the probability that an observation from an *F* distribution, with numerator degrees of freedom *ndf*, denominator degrees of freedom *ddf*, and noncentrality parameter *nc*, is less than or equal to *x*. The PROBF function accepts noninteger degrees of freedom parameters *ndf* and *ddf*. If the optional parameter *nc* is not specified or has the value 0, the value returned is from the central *F* distribution.

The significance level for an *F* test statistic is given by

```
p=1-probf(x,ndf,ddf);
```

Examples

SAS Statements	Results
`x=probf(3.32,2,3);`	`0.8263933602`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBGAM Function

Returns the probability from a gamma distribution.

Category: Probability
See: "CDF Function" on page 558

Syntax

PROBGAM(*x*,*a*)

Arguments

x
is a numeric random variable.
Range: $x \geq 0$

a
is a numeric shape parameter.
Range: $a > 0$

Details

The PROBGAM function returns the probability that an observation from a gamma distribution, with shape parameter a, is less than or equal to x.

Examples

SAS Statements	Results
`x=probgam(1,3);`	`0.0803013971`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBHYPR Function

Returns the probability from a hypergeometric distribution.

Category: Probability

See: "CDF Function" on page 558

Syntax

PROBHYPR(*N,K,n,x<,r>*)

Arguments

N
is an integer population size parameter, with $N \geq 1$.
Range:

K
is an integer number of items in the category of interest parameter.
Range: $0 \leq K \leq N$

n
is an integer sample size parameter.
Range: $0 \leq n \leq N$

x
is an integer random variable.
Range: $max(0, K + n–N) \leq x \leq min(K,n)$

r
is an optional numeric odds ratio parameter.

Range: $r \geq 0$

Details

The PROBHYPR function returns the probability that an observation from an extended hypergeometric distribution, with population size N, number of items K, sample size n, and odds ratio r, is less than or equal to x. If the optional parameter r is not specified or is set to 1, the value returned is from the usual hypergeometric distribution.

Examples

SAS Statements	Results
`x=probhypr(200,50,10,2);`	`0.5236734081`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBIT Function

Returns a quantile from the standard normal distribution.

Category: Quantile

Syntax

PROBIT(*p*)

Arguments

p
is a numeric probability.

Range: $0 < p < 1$

Details

The PROBIT function returns the p^{th} quantile from the standard normal distribution. The probability that an observation from the standard normal distribution is less than or equal to the returned quantile is p.

CAUTION:
The result could be truncated to lie between -8.222 and 7.941. △

Note: PROBIT is the inverse of the PROBNORM function. △

Examples

SAS Statements	Results
`x=probit(.025);`	`-1.959963985`
`x=probit(1.e-7);`	`-5.199337582`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBMC Function

Returns a probability or a quantile from various distributions for multiple comparisons of means.

Category: Probability

Syntax

PROBMC(*distribution, q, prob, df, nparms*<, *parameters*>)

Arguments

distribution
is a character constant, variable, or expression that identifies the distribution. The following are valid distributions:

Distribution	Argument
Analysis of Means	**ANOM**
One-sided Dunnett	**DUNNETT1**
Two-sided Dunnett	**DUNNETT2**
Maximum Modulus	**MAXMOD**
Partitioned Range	**PARTRANGE**
Studentized Range	**RANGE**
Williams	**WILLIAMS**

q
is the quantile from the distribution.

Restriction: Either *q* or *prob* can be specified, but not both.

prob
is the left probability from the distribution.

Restriction: Either *prob* or *q* can be specified, but not both.

df
is the degrees of freedom.

Note: A missing value is interpreted as an infinite value. △

nparms
is the number of treatments.

Note: For DUNNETT1 and DUNNETT2, the control group is not counted. △

parameters
is an optional set of *nparms* parameters that must be specified to handle the case of unequal sample sizes. The meaning of *parameters* depends on the value of *distribution*. If *parameters* is not specified, equal sample sizes are assumed, which is usually the case for a null hypothesis.

Details

The PROBMC function returns the probability or the quantile from various distributions with finite and infinite degrees of freedom for the variance estimate.

The *prob* argument is the probability that the random variable is less than *q*. Therefore, *p*-values can be computed as 1– *prob*. For example, to compute the critical value for a 5% significance level, set *prob*= 0.95. The precision of the computed probability is $O(10^{-8})$ (absolute error); the precision of computed quantile is $O(10^{-5})$.

Note: The studentized range is not computed for finite degrees of freedom and unequal sample sizes. △

Note: Williams' test is computed only for equal sample sizes. △

Formulas and Parameters The equations listed here define expressions used in equations that relate the probability, *prob*, and the quantile, *q*, for different distributions and different situations within each distribution. For these equations, let ν be the degrees of freedom, *df*.

$$d\mu_\nu(x) = \frac{\nu^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)2^{\frac{\nu}{2}-1}}x^{\nu-1}e^{-\frac{\nu x^2}{2}}dx$$

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

$$\Phi(x) = \int_{-\infty}^{x}\phi(u)\,du$$

Computing the Analysis of Means Analysis of Means (ANOM) applies to data that is organized as k (Gaussian) samples, the i^{th} sample being of size n_i. Let $I = \sqrt{-1}$. The distribution function [1, 2, 3, 4, 5] is the CDF for the maximum absolute of a k-dimensional multivariate T vector, with ν degrees of freedom, and an associated correlation matrix $\rho_{ij} = -\alpha_i\alpha_j$. This equation can be written as

$$\begin{aligned} prob &= \Pr\left(|t_1| < h, |t_2| < h, ..., |t_k| < h\right) \\ &= \int_0^\infty \left\{ \int_0^\infty \prod_{j=0}^{j=k} g\left(sh, y, \alpha_j\right)\phi(y)\,dy \right\} d\mu_\nu(s) \end{aligned}$$

where

$$g\left(sh, y, \alpha_j\right) = \Phi\left(\frac{sh - \mathrm{I}y\alpha_j}{\sqrt{1+\alpha_j^2}}\right) - \Phi\left(\frac{-sh - \mathrm{I}y\alpha_j}{\sqrt{1+\alpha_j^2}}\right)$$

where $\Gamma(.)$, $\phi(.)$, and $\mathbf{\Phi}(.)$, are the gamma function, the density, and the CDF from the standard normal distribution, respectively.

For $\nu = \infty$, the distribution reduces to:

$$\Pr\left(|t_1| < h, |t_2| < h, ..., |t_k| < h\right) = \int_0^\infty \prod_{j=0}^{j=k} g\left(h, y, \alpha_j\right)\phi(y)\,dy$$

where

$$g\left(h, y, \alpha_j\right) = \Phi\left(\frac{h - \mathrm{I}y\alpha_j}{\sqrt{1+\alpha_j^2}}\right) - \Phi\left(\frac{-h - \mathrm{I}y\alpha_j}{\sqrt{1+\alpha_j^2}}\right)$$

For the balanced case, the distribution reduces to:

$$\Pr\left(|t_1| < h, |t_2| < h, ..., |t_n| < h\right) = \int_0^\infty f\left(h, y, \rho\right)^n \phi\left(y\right) dy$$

where

$$f\left(h, y, \rho\right) = \Phi\left(\frac{h - \mathrm{I}y\sqrt{\rho}}{\sqrt{1+\rho}}\right) - \Phi\left(\frac{-h - \mathrm{I}y\sqrt{\rho}}{\sqrt{1+\rho}}\right)$$

and $\rho = \frac{1}{n-1}$

The syntax for this distribution is:

x=probmc('anom', q,p,nu,n,<alpha$_1$,..., alpha$_n$>);

where

x	is a numeric value with the returned result.
q	is a numeric value denoting the quantile.
p	is a numeric value denoting the probability. One of *p* and *q* must be missing.
nu	is a numeric value denoting the degrees of freedom.
n	is a numeric value denoting the number of samples.
alpha$_i$, i=1,...,*k*	are optional numeric values denoting the alpha values from the first equation of this distribution (see "Computing the Analysis of Means" on page 1022.

Many-One *t*-Statistics: Dunnett's One-Sided Test

□ This case relates the probability, *prob*, and the quantile, *q*, for the unequal case with finite degrees of freedom. The *parameters* are λ_1, ..., λ_k, the value of *nparms* is set to *k*, and the value of *df* is set to ν. The equation follows:

$$prob = \int_0^\infty \int_{-\infty}^\infty \phi\left(y\right) \prod_{i=1}^k \Phi\left(\frac{\lambda_i y + qx}{\sqrt{1-\lambda_i^2}}\right) dy\, du_\nu\left(x\right)$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with finite degrees of freedom. No *parameters* are passed $\left(\lambda = \sqrt{\frac{1}{2}}\right)$, the value of *nparms* is set to *k*, and the value of *df* is set to ν. The equation follows:

$$prob = \int_0^{\infty} \int_{-\infty}^{\infty} \phi(y) \left[\Phi\left(y + \sqrt{2qx}\right)\right]^k dy\, du_\nu(x)$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the unequal case with infinite degrees of freedom. The *parameters* are λ_1, ..., λ_k, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = \int_{-\infty}^{\infty} \phi(y) \prod_{i=1}^{k} \Phi\left(\frac{\lambda_i y + q}{\sqrt{1-\lambda_i^2}}\right) dy$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with infinite degrees of freedom. No *parameters* are passed $\left(\lambda = \sqrt{\frac{1}{2}}\right)$, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = \int_{-\infty}^{\infty} \phi(y) \left[\Phi\left(y + \sqrt{2q}\right)\right]^k dy$$

Many-One *t*-Statistics: Dunnett's Two-sided Test

□ This case relates the probability, *prob*, and the quantile, *q*, for the unequal case with finite degrees of freedom. The *parameters* are λ_1, ..., λ_k, the value of *nparms* is set to *k*, and the value of *df* is set to ν. The equation follows:

$$prob = \int_0^{\infty} \int_{-\infty}^{\infty} \phi(y) \prod_{i=1}^{k} \left[\Phi\left(\frac{\lambda_i y + qx}{\sqrt{1-\lambda_i^2}}\right) - \Phi\left(\frac{\lambda_i y - qx}{\sqrt{1-\lambda_i^2}}\right)\right] dy\, du_\nu(x)$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with finite degrees of freedom. No *parameters* are passed, the value of *nparms* is set to *k*, and the value of *df* is set to ν. The equation follows:

$$prob = \int_0^{\infty} \int_{-\infty}^{\infty} \phi(y) \left[\Phi\left(y + \sqrt{2qx}\right) - \Phi\left(y - \sqrt{2qx}\right)\right]^k dy\, du_\nu(x)$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the unequal case with infinite degrees of freedom. The *parameters* are λ_1, ..., λ_k, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = \int_{-\infty}^{\infty} \phi(y) \prod_{i=1}^{k} \left[\Phi\left(\frac{\lambda_i y + q}{\sqrt{1-\lambda_i^2}} \right) - \Phi\left(\frac{\lambda_i y - q}{\sqrt{1-\lambda_i^2}} \right) \right] dy$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with infinite degrees of freedom. No *parameters* are passed, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = \int_{-\infty}^{\infty} \phi(y) \left[\Phi\left(y + \sqrt{2q} \right) - \Phi\left(y - \sqrt{2q} \right) \right]^k dy$$

Computing the Partitioned Range RANGE applies to the distribution of the studentized range for *n* group means. PARTRANGE applies to the distribution of the partitioned studentized range. Let the *n* groups be partitioned into *k* subsets of size n_1+ ...+ n_k= *n*. Then the partitioned range is the maximum of the studentized ranges in the respective subsets, with the studentization factor being the same in all cases.

$$prob = \int_{0}^{\infty} \prod_{i=1}^{i=k} \left(\int_{-\infty}^{\infty} k\phi(y) \left(\mathbf{\Phi}(y) - \mathbf{\Phi}(y - qx) \right)^{k-1} dy \right)^{n_i} d\mu_\nu(x)$$

The syntax for this distribution is:

x=probmc('partrange', q,p,nu,k,n_1*,...,*n_k*);*

where

x	is a numeric value with the returned result (either the probability or the quantile).
q	is a numeric value denoting the quantile.
p	is a numeric value denoting the probability. One of *p* and *q* must be missing.
nu	is a numeric value denoting the degrees of freedom.
k	is a numeric value denoting the number of groups.
n_i *i*=1,...,*k*	are optional numeric values denoting the *n* values from the equation in this distribution (see "Computing the Partitioned Range" on page 1025.

The Studentized Range

Note: The studentized range is not computed for finite degrees of freedom and unequal sample sizes. △

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with finite degrees of freedom. No *parameters* are passed, the value of *nparms* is set to *k*, and the value of *df* is set to ν. The equation follows:

$$prob = \int_{0}^{\infty} \int_{-\infty}^{\infty} k\phi(y) \left[\Phi(y) - \Phi(y - qx)\right]^{k-1} dy \, du_{\nu}(x)$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the unequal case with infinite degrees of freedom. The *parameters* are σ_1, ..., σ_k, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = \int_{-\infty}^{\infty} \sum_{j=1}^{k} \left\{ \prod_{i=1}^{k} \left[\Phi\left(\frac{y}{\sigma_i}\right) - \Phi\left(\frac{y-q}{\sigma_i}\right) \right] \right\} \phi\left(\frac{y}{\sigma_j}\right) \frac{1}{\sigma_j} dy$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with infinite degrees of freedom. No *parameters* are passed, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = \int_{-\infty}^{\infty} k\phi(y) \left[\Phi(y) - \Phi(y - q)\right]^{k-1} dy$$

The Studentized Maximum Modulus

□ This case relates the probability, *prob*, and the quantile, *q*, for the unequal case with finite degrees of freedom. The *parameters* are σ_1, ..., σ_k, the value of *nparms* is set to *k*, and the value of *df* is set to ν. The equation follows:

$$prob = \int_{0}^{\infty} \prod_{i=1}^{k} \left[2\Phi\left(\frac{qx}{\sigma_i}\right) - 1 \right] d\mu_{\nu}(x)$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with finite degrees of freedom. No *parameters* are passed, the value of *nparms* is set to *k*, and the value of *df* is set to ν. The equation follows:

$$prob = \int_{0}^{\infty} \left[2\Phi(qx) - 1\right]^{k} d\mu_{\nu}(x)$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the unequal case with infinite degrees of freedom. The *parameters* are σ_1, ..., σ_k, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = \prod_{i=1}^{k} \left[2\Phi\left(\frac{q}{\sigma_i}\right) - 1 \right]$$

□ This case relates the probability, *prob*, and the quantile, *q*, for the equal case with infinite degrees of freedom. No *parameters* are passed, the value of *nparms* is set to *k*, and the value of *df* is set to missing. The equation follows:

$$prob = [2\Phi(q) - 1]^k$$

Williams' Test PROBMC computes the probabilities or quantiles from the distribution defined in Williams (1971, 1972) (See "References" on page 1255). It arises when you compare the dose treatment means with a control mean to determine the lowest effective dose of treatment.

Note: Williams' Test is computed only for equal sample sizes. △

Let X_1, X_2, ..., X_k be identical independent N(0,1) random variables. Let Y_k denote their average given by

$$Y_k = \frac{X_1 + X_2 + ... + X_k}{k}$$

It is required to compute the distribution of

$$(Y_k - Z)/S$$

where

Y_k	is as defined previously
Z	is an N(0,1) independent random variable
S	is such that $\frac{1}{2}\nu S^2$ is a χ^2 variable with ν degrees of freedom.

As described in Williams (1971) (See "References" on page 1255), the full computation is extremely lengthy and is carried out in three stages.

1 Compute the distribution of Y_k. It is the fundamental (expensive) part of this operation and it can be used to find both the density and the probability of Y_k. Let U_i be defined as

$$U_i = \frac{X_1 + X_2 + ... + X_i}{i},\ i = 1, 2, ..., k$$

You can write a recursive expression for the probability of $Y_k > d$, with d being any real number.

$$\begin{aligned}\Pr(Y_k > d) &= \Pr(U_1 > d) \\ &\quad + \Pr(U_2 > d, U_1 < d) \\ &\quad + \Pr(U_3 > d, U_2 < d, U_1 < d) \\ &\quad + \ldots \\ &\quad + \Pr(U_k > d, U_{k-1} < d, \ldots, U_1 < d) \\ &= \Pr(Y_{k-1} > d) + \Pr(X_k + (k-1)U_{k-1} > kd)\end{aligned}$$

To compute this probability, start from an N(0,1) density function

$$D\left(U_1 = x\right) = \phi\left(x\right)$$

and recursively compute the convolution

$$D\left(U_k = x, U_{k-1} < d, \ldots, U_1 < d\right) =$$
$$\int_{-\infty}^{d} D\left(U_{k-1} = y, U_{k-2} < d, \ldots, U_1 < d\right)\left(k-1\right)\phi\left(kx - \left(k-1\right)y\right)dy$$

From this sequential convolution, it is possible to compute all the elements of the recursive equation for $\Pr\left(Y_k < d\right)$, shown previously.

2 Compute the distribution of $Y_k - Z$. This computation involves another convolution to compute the probability

$$\Pr\left(\left(Y_k - Z\right) > d\right) = \int_{-\infty}^{\infty} \Pr\left(Y_k > \sqrt{2d} + y\right)\phi\left(y\right)dy$$

3 Compute the distribution of $(Y_k - Z)/S$. This computation involves another convolution to compute the probability

$$\Pr\left(\left(Y_k - Z\right) > tS\right) = \int_{0}^{\infty} \Pr\left(\left(Y_k - Z\right) > ty\right)d\mu_\nu\left(y\right)$$

The third stage is not needed when $\nu = \infty$. Due to the complexity of the operations, this lengthy algorithm is replaced by a much faster one when $k \leq 15$ for both finite and infinite degrees of freedom ν. For $k \geq 16$, the lengthy computation is carried out. It is extremely expensive and very slow due to the complexity of the algorithm.

Comparisons

The MEANS statement in the GLM Procedure of SAS/STAT Software computes the following tests:

- Dunnett's one-sided test
- Dunnett's two-sided test
- Studentized Range

Examples

Example 1: Computing Probabilities by Using PROBMC This example shows how to compute probabilities.

```
data probs;
   array par{5};
      par{1}=.5;
      par{2}=.51;
      par{3}=.55;
      par{4}=.45;
      par{5}=.2;
   df=40;
   q=1;
   do test="dunnett1","dunnett2", "maxmod";
      prob=probmc(test, q, ., df, 5, of par1--par5);
      put test $10. df q e18.13 prob e18.13;
   end;
run;
```

SAS writes the following results to the log:

Output 4.65 Probabilities from PROBMC

```
DUNNETT1  40  1.00000000000E+00 4.82992196083E-01
DUNNETT2  40  1.00000000000E+00 1.64023105316E-01
MAXMOD    40  1.00000000000E+00 8.02784203408E-01
```

Example 2: Computing the Analysis of Means

```
data _null_;
   q1=probmc('anom',.,0.9,.,20);                        put q1=;
   q2=probmc('anom',.,0.9,20,5,0.1,0.1,0.1,0.1,0.1);  put q2=;
   q3=probmc('anom',.,0.9,20,5,0.5,0.5,0.5,0.5,0.5);  put q3=;
   q4=probmc('anom',.,0.9,20,5,0.1,0.2,0.3,0.4,0.5);  put q4=;
run;
```

SAS writes the following output to the log:

```
q1=2.7892895753
q2=2.4549773558
q3=2.4549773558
q4=2.4532130238
```

Example 3: Comparing Means This example shows how to compare group means to find where the significant differences lie. The data for this example is taken from a paper by Duncan (1955) (See "References" on page 1255) and can also be found in Hochberg and Tamhane (1987) (See "References" on page 1255).

The following values are the group means:

49.6
71.2
67.6
61.5
71.3
58.1
61.0

For this data, the mean square error is $s^2 = 79.64$ ($s = 8.924$) with $\nu = 30$.

```
data duncan;
   array tr{7}$;
   array mu{7};
   n=7;
   do i=1 to n;
      input tr{i} $1. mu{i};
   end;
   input df s alpha;
   prob= 1-alpha;
      /* compute the interval */
   x = probmc("RANGE", ., prob, df, 7);
   w = x * s / sqrt(6);
      /* compare the means */
   do i = 1 to n;
      do j = i + 1 to n;
         dmean = abs(mu{i} - mu{j});
         if dmean >= w then do;
            put tr{i} tr{j} dmean;
         end;
      end;
   end;
   datalines;
A 49.6
B 71.2
C 67.6
D 61.5
E 71.3
F 58.1
G 61.0
 30 8.924 .05
;
run;
```

SAS writes the following output to the log:

Output 4.66 Group Differences

```
A B 21.6
A C 18
A E 21.7
```

Example 4: Computing the Partitioned Range

```
data _null_;
   q1=probmc('partrange',.,0.9,.,4,3,4,5,6);  put q1=;
   q2=probmc('partrange',.,0.9,12,4,3,4,5,6); put q2=;
run;
```

SAS writes the following output to the log:

```
q1=4.1022395729
q2=4.788862411
```

Example 5: Computing Confidence Intervals This example shows how to compute 95% one-sided and two-sided confidence intervals of Dunnett's test. This example and the data come from Dunnett (1955) (See "References" on page 1255) and can also be found in Hochberg and Tamhane (1987) (See "References" on page 1255). The data are blood count measurements on three groups of animals. As shown in the following table, the third group serves as the control, while the first two groups were treated with different drugs. The numbers of animals in these three groups are unequal.

Treatment Group:	Drug A	Drug B	Control
	9.76	12.80	7.40
	8.80	9.68	8.50
	7.68	12.16	7.20
	9.36	9.20	8.24
		10.55	9.84
			8.32
Group Mean	8.90	10.88	8.25
n	4	5	6

The mean square error $s^2 = 1.3805$ ($s = 1.175$) with $\nu = 12$.

```
data a;
   array drug{3}$;
   array count{3};
   array mu{3};
   array lambda{2};
   array delta{2};
   array left{2};
   array right{2};

      /* input the table */
```

```
   do i = 1 to 3;
      input drug{i} count{i} mu{i};
   end;

      /* input the alpha level,    */
      /* the degrees of freedom,   */
      /* and the mean square error */
   input alpha df s;

      /* from the sample size, */
      /* compute the lambdas   */
   do i = 1 to 2;
      lambda{i} = sqrt(count{i}/
        (count{i} + count{3}));
   end;

      /* run the one-sided Dunnett's test */
   test="dunnett1";
      x = probmc(test, ., 1 - alpha, df,
                 2, of lambda1-lambda2);
      do i = 1 to 2;
         delta{i} = x * s *
            sqrt(1/count{i} + 1/count{3});
         left{i} = mu{i} - mu{3} - delta{i};
      end;
   put test $10. x left{1} left{2};

      /* run the two-sided Dunnett's test */
   test="dunnett2";
      x = probmc(test, ., 1 - alpha, df,
                 2, of lambda1-lambda2);
      do i=1 to 2;
         delta{i} = x * s *
            sqrt(1/count{i} + 1/count{3});
         left{i} = mu{i} - mu{3} - delta{i};
         right{i} = mu{i} - mu{3} + delta{i};
      end;
   put test $10. left{1} right{1};
   put test $10. left{2} right{2};
   datalines;
A 4 8.90
B 5 10.88
C 6 8.25
0.05 12 1.175
;
run;
```

SAS writes the following output to the log:

Output 4.67 Confidence Intervals

```
DUNNETT1  2.1210786586 -0.958751705 1.1208571303
DUNNETT2  -1.256411895 2.5564118953
DUNNETT2  0.8416271203 4.4183728797
```

Example 6: Computing Williams' Test In the following example, a substance has been tested at seven levels in a randomized block design of eight blocks. The observed treatment means are as follows:

Treatment	Mean
X_0	10.4
X_1	9.9
X_2	10.0
X_3	10.6
X_4	11.4
X_5	11.9
X_6	11.7

The mean square, with $(7 - 1)(8 - 1) = 42$ degrees of freedom, is $s^2 = 1.16$.
Determine the maximum likelihood estimates M_i through the averaging process.

- Because $X_0 > X_1$, form $X_{0,1} = (X_0 + X_1)/2 = 10.15$.
- Because $X_{0,1} > X_2$, form $X_{0,1,2} = (X_0 + X_1 + X_2)/3 = (2X_{0,1} + X_2)/3 = 10.1$.
- $X_{0,1,2} < X_3 < X_4 < X_5$
- Because $X_5 > X_6$, form $X_{5,6} = (X_5 + X_6)/2 = 11.8$.

Now the order restriction is satisfied.
The maximum likelihood estimates under the alternative hypothesis are:
$M_0 = M_1 = M_2 = X_{0,1,2} = 10.1$
$M_3 = X_3 = 10.6$
$M_4 = X_4 = 11.4$
$M_5 = M_6 = X_{5,6} = 11.8$

Now compute $t = (11.8 - 10.4) / \sqrt{2s^2/8} = 2.60$, and the probability that corresponds to $k = 6$, $\nu = 42$, and $t = 2.60$ is .9924467341, which shows strong evidence that there is a response to the substance. You can also compute the quantiles for the upper 5% and 1% tails, as shown in the following table.

SAS Statements	Results
`prob=probmc("williams",2.6,.,42,6);`	**0.99244673**
`quant5=probmc("williams",.,.95,42,6);`	**1.80654052**
`quant1=probmc("williams",.,.99,42,6);`	**2.49087829**

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

References

Guirguis, G. H. and R. D. Tobias. 2004. "On the computation of the distribution for the analysis of means." *Communications in Statistics: Simulation and Computation* 33: 861–887.

Nelson, P. R. 1981. "Numerical evaluation of an equicorrelated multivariate non-central *t* distribution." *Communications in Statistics: Part B - Simulation and Computation* 10: 41–50.

Nelson, P. R. 1982. "Exact critical points for the analysis of means." *Communications in Statistics: Part A - Theory and Methods* 11: 699–709.

Nelson, P. R. 1982a. "An Approximation for the Complex Normal Probability Integral." *BIT* 22(1): 94–100.

Nelson, P. R. 1988. "Application of the analysis of means." *Proceedings of the SAS Users Group International Conference* 13: 225–230.

Nelson, P. R. 1991. "Numerical evaluation of multivariate normal integrals with correlations $\rho_{lj=-\alpha_l\alpha_j}$." *The Frontiers of Statistical Scientific Theory and Industrial Applications* 2: 97–114.

Nelson, P. R. 1993. "Additional Uses for the Analysis of Means and Extended Tables of Critical Values." *Technometrics* 35: 61–71.

PROBNEGB Function

Returns the probability from a negative binomial distribution.

Category: Probability

See: "CDF Function" on page 558

Syntax

PROBNEGB(*p,n,m*)

Arguments

p
is a numeric probability of success parameter.

Range: $0 \leq p \leq 1$

n
is an integer number of successes parameter.

Range: $n \geq 1$

m
is a positive integer random variable, the number of failures.

Range: $m \geq 0$

Details

The PROBNEGB function returns the probability that an observation from a negative binomial distribution, with probability of success p and number of successes n, is less than or equal to m.

To compute the probability that an observation is equal to a given value m, compute the difference of two probabilities from the negative binomial distribution for m and m–1.

Examples

SAS Statements	Results
`x=probnegb(0.5,2,1);`	`0.5`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBNORM Function

Returns the probability from the standard normal distribution.

Category: Probability

See: "CDF Function" on page 558

Syntax

PROBNORM(*x*)

Arguments

x
is a numeric random variable.

Details

The PROBNORM function returns the probability that an observation from the standard normal distribution is less than or equal to *x*.

Note: PROBNORM is the inverse of the PROBIT function. △

Examples

SAS Statements	Results
`x=probnorm(1.96);`	`0.9750021049`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROBT Function

Returns the probability from a *t* distribution.

Category: Probability
See: "CDF Function" on page 558, "PDF Function" on page 986

Syntax

PROBT(*x*,*df*<,*nc*>)

Arguments

x
is a numeric random variable.

df
is a numeric degrees of freedom parameter.

Range: $df > 0$

nc
is an optional numeric noncentrality parameter.

Details

The PROBT function returns the probability that an observation from a Student's *t* distribution, with degrees of freedom *df* and noncentrality parameter *nc*, is less than or equal to *x*. This function accepts a noninteger degree of freedom parameter *df*. If the optional parameter, *nc*, is not specified or has the value 0, the value that is returned is from the central Student's *t* distribution.

The significance level of a two-tailed *t* test is given by

```
p=(1-probt(abs(x),df))*2;
```

Examples

SAS Statements	Results
`x=probt(0.9,5);`	`0.7953143998`

See Also

Functions:

"CDF Function" on page 558
"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120

PROPCASE Function

Converts all words in an argument to proper case.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

PROPCASE(*argument* <,*delimiters*>)

Arguments

argument
specifies a character constant, variable, or expression.

delimiter
specifies one or more delimiters that are enclosed in quotation marks. The default delimiters are blank, forward slash, hyphen, open parenthesis, period, and tab.

Tip: If you use this argument, then the default delimiters, including the blank, are no longer in effect.

Details

Length of Returned Variable In a DATA step, if the PROPCASE function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The PROPCASE function copies a character argument and converts all uppercase letters to lowercase letters. It then converts to uppercase the first character of a word that is preceded by a blank, forward slash, hyphen, open parenthesis, period, or tab. PROPCASE returns the value that is altered.

If you use the second argument, then the default delimiters are no longer in effect.

The results of the PROPCASE function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

Examples

Example 1: Changing the Case of Words The following example shows how PROPCASE handles the case of words:

```
data _null_;
   input place $ 1-40;
   name=propcase(place);
   put name;
   datalines;
INTRODUCTION TO THE SCIENCE OF ASTRONOMY
VIRGIN ISLANDS (U.S.)
SAINT KITTS/NEVIS
WINSTON-SALEM, N.C.
;

run;
```

SAS writes the following output to the log:

```
Introduction To The Science Of Astronomy
Virgin Islands (U.S.)
Saint Kitts/Nevis
Winston-Salem, N.C.
```

Example 2: Using PROPCASE with a Second Argument The following example uses a blank, a hyphen and a single quotation mark as the second argument so that names such as O'Keeffe and Burne-Jones are written correctly.

```
options pageno=1 nodate ls=80 ps=64;
data names;
   infile datalines dlm='#';
   input CommonName : $20. CapsName : $20.;
   PropcaseName=propcase(capsname, " -'");
   datalines;
Delacroix, Eugene# EUGENE DELACROIX
O'Keeffe, Georgia# GEORGIA O'KEEFFE
Rockwell, Norman# NORMAN ROCKWELL
Burne-Jones, Edward# EDWARD BURNE-JONES
;

proc print data=names noobs;
   title 'Names of Artists';
run;
```

Output 4.68 Output Showing the Results of Using PROPCASE with a Second Argument

```
                          Names of Artists                                1

       CommonName             CapsName            PropcaseName

    Delacroix, Eugene      EUGENE DELACROIX      Eugene Delacroix
    O'Keeffe, Georgia      GEORGIA O'KEEFFE      Georgia O'Keeffe
    Rockwell, Norman       NORMAN ROCKWELL       Norman Rockwell
    Burne-Jones, Edward    EDWARD BURNE-JONES    Edward Burne-Jones
```

See Also

Functions:

"UPCASE Function" on page 1177
"LOWCASE Function" on page 912

PRXCHANGE Function

Performs a pattern-matching replacement.

Category: Character String Matching

Syntax

PRXCHANGE(*perl-regular-expression* | *regular-expression-id*, *times*, *source*)

Arguments

perl-regular-expression
specifies a character constant, variable, or expression with a value that is a Perl regular expression.

regular-expression-id
specifies a numeric variable with a value that is a pattern identifier that is returned from the PRXPARSE function.

Restriction: If you use this argument, you must also use the PRXPARSE function.

times
is a numeric constant, variable, or expression that specifies the number of times to search for a match and replace a matching pattern.

Tip: If the value of *times* is –1, then matching patterns continue to be replaced until the end of *source* is reached.

source
specifies a character constant, variable, or expression that you want to search.

Details

The Basics If you use *regular-expression-id*, the PRXCHANGE function searches the variable *source* with the *regular-expression-id* that is returned by PRXPARSE. It returns the value in *source* with the changes that were specified by the regular expression. If there is no match, PRXCHANGE returns the unchanged value in *source*.

If you use *perl-regular-expression*, PRXCHANGE searches the variable *source* with the *perl-regular-expression*, and you do not need to call PRXPARSE. You can use PRXCHANGE with a *perl-regular-expression* in a WHERE clause and in PROC SQL.

For more information about pattern matching, see “Pattern Matching Using Perl Regular Expressions (PRX)” on page 333.

Compiling a Perl Regular Expression If *perl-regular-expression* is a constant or if it uses the /o option, then the Perl regular expression is compiled once and each use of PRXCHANGE reuses the compiled expression. If *perl-regular-expression* is not a constant and if it does not use the /o option, then the Perl regular expression is recompiled for each call to PRXCHANGE.

Note: The compile-once behavior occurs when you use PRXCHANGE in a DATA step, in a WHERE clause, or in PROC SQL. For all other uses, the *perl-regular-expression* is recompiled for each call to PRXCHANGE. △

Performing a Match Perl regular expressions consist of characters and special characters that are called metacharacters. When performing a match, SAS searches a source string for a substring that matches the Perl regular expression that you specify.

To view a short list of Perl regular expression metacharacters that you can use when you build your code, see the table “Tables of Perl Regular Expression (PRX) Metacharacters” on page 2205.

Comparisons

The PRXCHANGE function is similar to the CALL PRXCHANGE routine except that the function returns the value of the pattern-matching replacement as a return argument instead of as one of its parameters.

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in "Functions and CALL Routines by Category" on page 345.

Examples

Example 1: Changing the Order of First and Last Names

Changing the Order of First and Last Names by Using the DATA Step The following example uses the DATA step to change the order of first and last names.

```
   /* Create a data set that contains a list of names. */
data ReversedNames;
   input name & $32.;
   datalines;
Jones, Fred
Kavich, Kate
Turley, Ron
Dulix, Yolanda
;

   /* Reverse last and first names with a DATA step. */
options pageno=1 nodate ls=80 ps=64;
data names;
   set ReversedNames;
   name = prxchange('s/(\w+), (\w+)/$2 $1/', -1, name);
run;

proc print data=names;
run;
```

Output 4.69 Output from the DATA Step

```
                              The SAS System                              1

                          Obs        name

                           1     Fred Jones
                           2     Kate Kavich
                           3     Ron Turley
                           4     Yolanda Dulix
```

Changing the Order of First and Last Names by Using PROC SQL The following example uses PROC SQL to change the order of first and last names.

```
data ReversedNames;
   input name & $32.;
   datalines;
Jones, Fred
Kavich, Kate
Turley, Ron
Dulix, Yolanda
;

proc sql;
   create table names as
   select prxchange('s/(\w+), (\w+)/$2 $1/', -1, name) as name
   from ReversedNames;
quit;

options pageno=1 nodate ls=80 ps=64;
proc print data=names;
run;
```

Output 4.70 Output from PROC SQL

```
                              The SAS System                              1

                            Obs        name

                             1     Fred Jones
                             2     Kate Kavich
                             3     Ron Turley
                             4     Yolanda Dulix
```

Example 2: Matching Rows That Have the Same Name The following example compares the names in two data sets, and writes those names that are common to both data sets.

```
data names;
   input name & $32.;
   datalines;
Ron Turley
Judy Donnelly
Kate Kavich
Tully Sanchez
;

data ReversedNames;
   input name & $32.;
   datalines;
Jones, Fred
Kavich, Kate
Turley, Ron
Dulix, Yolanda
;
```

```
options pageno=1 nodate ls=80 ps=64;
proc sql;
   create table NewNames as
   select a.name from names as a, ReversedNames as b
   where a.name = prxchange('s/(\w+), (\w+)/$2 $1/', -1, b.name);
quit;

proc print data=NewNames;
run;
```

Output 4.71 Output from Matching Rows That Have the Same Names

```
                              The SAS System                              1

                            Obs        name

                             1     Ron Turley
                             2     Kate Kavich
```

Example 3: Changing Lowercase Text to Uppercase The following example uses the \U, \L and \E metacharacters to change the case of a string of text. Case modifications do not nest. In this example, note that "bear" does not convert to uppercase letters because the \E metacharacter ends all case modifications.

```
data _null_;
   length txt $32;
   txt = prxchange ('s/(big)(black)(bear)/\U$1\L$2\E$3/', 1, 'bigblackbear');
   put txt=;
run;
```

SAS returns the following output to the log:

```
txt=BIGblackbear
```

Example 4: Changing a Matched Pattern to a Fixed Value

This example locates a pattern in a variable and replaces the variable with a predefined value. The example uses a DATA step to find phone numbers and replace them with an informational message.

```
options nodate nostimer ls=78 ps=60;

   /* Create data set that contains confidential information. */
data a;
   input text $80.;
   datalines;
The phone number for Ed is (801)443-9876 but not until tonight.
He can be reached at (910)998-8762 tomorrow for testing purposes.
;
run;

   /* Locate confidential phone numbers and replace them with message */
   /* indicating that they have been removed.                         */
data b;
   set a;
   text = prxchange('s/\([2-9]\d\d\) ?[2-9]\d\d-\d\d\d\d/*PHONE NUMBER
      REMOVED*/', -1, text);
   put text=;
run;

proc print data = b;
run;
```

Output 4.72 Output from Changing a Matched Pattern to a Fixed Value

```
                                The SAS System                                1

Obs                                     text

 1  The phone number for Ed is *PHONE NUMBER REMOVED* but not until tonight.
 2  He can be reached at *PHONE NUMBER REMOVED* tomorrow for testing purposes.
```

See Also

Functions and CALL routines:

"CALL PRXCHANGE Routine" on page 479
"CALL PRXDEBUG Routine" on page 482
"CALL PRXFREE Routine" on page 484
"CALL PRXNEXT Routine" on page 485
"CALL PRXPOSN Routine" on page 487
"CALL PRXSUBSTR Routine" on page 490
"PRXMATCH Function" on page 1045
"PRXPAREN Function" on page 1049

"PRXPARSE Function" on page 1051
"PRXPOSN Function" on page 1053

PRXMATCH Function

Searches for a pattern match and returns the position at which the pattern is found.

Category: Character String Matching

Syntax

PRXMATCH (*regular-expression-id* | *perl-regular-expression*, *source*)

Arguments

regular-expression-id
specifies a numeric variable with a value that is a pattern identifier that is returned from the PRXPARSE function.

Restriction: If you use this argument, you must also use the PRXPARSE function.

perl-regular-expression
specifies a character constant, variable, or expression with a value that is a Perl regular expression.

source
specifies a character constant, variable, or expression that you want to search.

Details

The Basics If you use *regular-expression-id*, then the PRXMATCH function searches *source* with the *regular-expression-id* that is returned by PRXPARSE, and returns the position at which the string begins. If there is no match, PRXMATCH returns a zero.

If you use *perl-regular-expression*, PRXMATCH searches *source* with the *perl-regular-expression*, and you do not need to call PRXPARSE.

You can use PRXMATCH with a Perl regular expression in a WHERE clause and in PROC SQL. For more information about pattern matching, see "Pattern Matching Using Perl Regular Expressions (PRX)" on page 333.

Compiling a Perl Regular Expression If *perl-regular-expression* is a constant or if it uses the /o option, then the Perl regular expression is compiled once and each use of PRXMATCH reuses the compiled expression. If *perl-regular-expression* is not a constant and if it does not use the /o option, then the Perl regular expression is recompiled for each call to PRXMATCH.

Note: The compile-once behavior occurs when you use PRXMATCH in a DATA step, in a WHERE clause, or in PROC SQL. For all other uses, the *perl-regular-expression* is recompiled for each call to PRXMATCH. △

Comparisons

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these

functions and CALL routines, see the Character String Matching category in “Functions and CALL Routines by Category” on page 345.

Examples

Example 1: Finding the Position of a Substring in a String

Finding the Position of a Substring by Using PRXPARSE The following example searches a string for a substring, and returns its position in the string.

```
  data _null_;
      /* Use PRXPARSE to compile the Perl regular expression.    */
   patternID = prxparse('/world/');
      /* Use PRXMATCH to find the position of the pattern match. */
   position=prxmatch(patternID, 'Hello world!');
   put position=;
run;
```

SAS writes the following line to the log:

```
position=7
```

Finding the Position of a Substring by Using a Perl Regular Expression The following example uses a Perl regular expression to search a string (Hello world) for a substring (world) and to return the position of the substring in the string.

```
data _null_;
      /* Use PRXMATCH to find the position of the pattern match. */
   position=prxmatch('/world/', 'Hello world!');
   put position=;
run;
```

SAS writes the following line to the log:

```
position=7
```

Example 2: Finding the Position of a Substring in a String: A Complex Example The following example uses several Perl regular expression functions and a CALL routine to find the position of a substring in a string.

```
data _null_;
   if _N_ = 1 then
   do;
      retain PerlExpression;
      pattern = "/(\d+):(\d\d)(?:\.(\d+))?/";
      PerlExpression = prxparse(pattern);
   end;

   array match[3] $ 8;
   input minsec $80.;
   position = prxmatch(PerlExpression, minsec);
   if position ^= 0 then
   do;
      do i = 1 to prxparen(PerlExpression);
         call prxposn(PerlExpression, i, start, length);
         if start ^= 0 then
```

```
            match[i] = substr(minsec, start, length);
      end;
      put match[1] "minutes, " match[2] "seconds" @;
      if ^missing(match[3]) then
         put ", " match[3] "milliseconds";
   end;
   datalines;
14:56.456
45:32
;

run;
```

The following lines are written to the SAS log:

```
14 minutes, 56 seconds, 456 milliseconds
45 minutes, 32 seconds
```

Example 3: Extracting a ZIP Code from a Data Set

Extracting a ZIP Code by Using the DATA Step The following example uses a DATA step to search each observation in a data set for a nine-digit ZIP code, and writes those observations to the data set ZipPlus4.

```
data ZipCodes;
   input name: $16. zip:$10.;
   datalines;
Johnathan 32523-2343
Seth 85030
Kim 39204
Samuel 93849-3843
;

   /* Extract ZIP+4 ZIP codes with the DATA step. */
data ZipPlus4;
   set ZipCodes;
   where prxmatch('/\d{5}-\d{4}/', zip);
run;

options nodate pageno=1 ls=80 ps=64;
proc print data=ZipPlus4;
run;
```

Output 4.73 ZIP Code Output from the DATA Step

```
                              The SAS System                              1

                        Obs    name            zip

                         1     Johnathan    32523-2343
                         2     Samuel       93849-3843
```

Extracting a ZIP Code by Using PROC SQL The following example searches each observation in a data set for a nine-digit ZIP code, and writes those observations to the data set ZipPlus4.

```
data ZipCodes;
   input name: $16. zip:$10.;
   datalines;
Johnathan 32523-2343
Seth 85030
Kim 39204
Samuel 93849-3843
;

   /* Extract ZIP+4 ZIP codes with PROC SQL. */
proc sql;
   create table ZipPlus4 as
   select * from ZipCodes
   where prxmatch('/\d{5}-\d{4}/', zip);
run;

options nodate pageno=1 ls=80 ps=64;
proc print data=ZipPlus4;
run;
```

Output 4.74 ZIP Code Output from PROC SQL

```
                              The SAS System                              1

                        Obs    name            zip

                         1     Johnathan    32523-2343
                         2     Samuel       93849-3843
```

See Also

Functions and CALL routines:

"CALL PRXCHANGE Routine" on page 479
"CALL PRXDEBUG Routine" on page 482
"CALL PRXFREE Routine" on page 484
"CALL PRXNEXT Routine" on page 485
"CALL PRXPOSN Routine" on page 487
"CALL PRXSUBSTR Routine" on page 490
"CALL PRXCHANGE Routine" on page 479
"PRXCHANGE Function" on page 1039
"PRXPAREN Function" on page 1049
"PRXPARSE Function" on page 1051
"PRXPOSN Function" on page 1053

PRXPAREN Function

Returns the last bracket match for which there is a match in a pattern.

Category: Character String Matching

Restriction: Use with the PRXPARSE function.

Syntax

PRXPAREN (*regular-expression-id*)

Arguments

regular-expression-id
specifies a numeric variable with a value that is an identification number that is returned by the PRXPARSE function.

Details

The PRXPAREN function is useful in finding the largest capture-buffer number that can be passed to the CALL PRXPOSN routine, or in identifying which part of a pattern matched.

For more information about pattern matching, see "Pattern Matching Using Perl Regular Expressions (PRX)" on page 333.

Comparisons

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in “Functions and CALL Routines by Category” on page 345.

Examples

The following example uses Perl regular expressions and writes the results to the SAS log.

```
data _null_;
   ExpressionID = prxparse('/(magazine)|(book)|(newspaper)/');
   position = prxmatch(ExpressionID, 'find book here');
   if position then paren = prxparen(ExpressionID);
   put 'Matched paren ' paren;

   position = prxmatch(ExpressionID, 'find magazine here');
   if position then paren = prxparen(ExpressionID);
   put 'Matched paren ' paren;

   position = prxmatch(ExpressionID, 'find newspaper here');
   if position then paren = prxparen(ExpressionID);
   put 'Matched paren ' paren;
run;
```

The following lines are written to the SAS log:

```
Matched paren 2
Matched paren 1
Matched paren 3
```

See Also

Functions and CALL routines:

“CALL PRXCHANGE Routine” on page 479
“CALL PRXDEBUG Routine” on page 482
“CALL PRXFREE Routine” on page 484
“CALL PRXNEXT Routine” on page 485
“CALL PRXPOSN Routine” on page 487
“CALL PRXSUBSTR Routine” on page 490
“CALL PRXCHANGE Routine” on page 479
“PRXCHANGE Function” on page 1039
“PRXMATCH Function” on page 1045
“PRXPARSE Function” on page 1051
“PRXPOSN Function” on page 1053

PRXPARSE Function

Compiles a Perl regular expression (PRX) that can be used for pattern matching of a character value.

Category: Character String Matching

Restriction: Use with other Perl regular expressions.

Syntax

regular-expression-id=**PRXPARSE** (*perl-regular-expression*)

Arguments

regular-expression-id
is a numeric pattern identifier that is returned by the PRXPARSE function.

perl-regular-expression
specifies a character value that is a Perl regular expression.

Details

The Basics The PRXPARSE function returns a pattern identifier number that is used by other Perl functions and CALL routines to match patterns. If an error occurs in parsing the regular expression, SAS returns a missing value.

PRXPARSE uses metacharacters in constructing a Perl regular expression. To view a table of common metacharacters, see "Tables of Perl Regular Expression (PRX) Metacharacters" on page 2205.

For more information about pattern matching, see "Pattern Matching Using Perl Regular Expressions (PRX)" on page 333.

Compiling a Perl Regular Expression If *perl-regular-expression* is a constant or if it uses the /o option, the Perl regular expression is compiled only once. Successive calls to PRXPARSE will not cause a recompile, but will return the *regular-expression-id* for the regular expression that was already compiled. This behavior simplifies the code because you do not need to use an initialization block (IF _N_ =1) to initialize Perl regular expressions.

Note: If you have a Perl regular expression that is a constant, or if the regular expression uses the /o option, then calling PRXFREE to free the memory allocation results in the need to recompile the regular expression the next time that it is called by PRXPARSE.

The compile-once behavior occurs when you use PRXPARSE in a DATA step. For all other uses, the *perl-regular-expression* is recompiled for each call to PRXPARSE. △

Comparisons

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in "Functions and CALL Routines by Category" on page 345.

Examples

The following example uses metacharacters and regular characters to construct a Perl regular expression. The example parses addresses and writes formatted results to the SAS log.

```
data _null_;
   if _N_ = 1 then
   do;
      retain patternID;

         /* The i option specifies a case insensitive search. */
      pattern = "/ave|avenue|dr|drive|rd|road/i";
      patternID = prxparse(pattern);
   end;

   input street $80.;
   call prxsubstr(patternID, street, position, length);
   if position ^= 0 then
   do;
      match = substr(street, position, length);
      put match:$QUOTE. "found in " street:$QUOTE.;
   end;
   datalines;
153 First Street
6789 64th Ave
4 Moritz Road
7493 Wilkes Place
;
```

The following lines are written to the SAS log:

```
"Ave" found in "6789 64th Ave"
"Road" found in "4 Moritz Road"
```

See Also

Functions and CALL routines:

“CALL PRXCHANGE Routine” on page 479

“CALL PRXDEBUG Routine” on page 482

“CALL PRXFREE Routine” on page 484

“CALL PRXNEXT Routine” on page 485

“CALL PRXPOSN Routine” on page 487

“CALL PRXSUBSTR Routine” on page 490

“CALL PRXCHANGE Routine” on page 479

“PRXCHANGE Function” on page 1039

“PRXPAREN Function” on page 1049

“PRXMATCH Function” on page 1045

“PRXPOSN Function” on page 1053

PRXPOSN Function

Returns a character string that contains the value for a capture buffer.

Category: Character String Matching

Restriction: Use with the PRXPARSE function.

Syntax

PRXPOSN(*regular-expression-id*, *capture-buffer*, *source*)

Arguments

regular-expression-id
specifies a numeric variable with a value that is a pattern identifier that is returned by the PRXPARSE function.

capture-buffer
is a numeric constant, variable, or expression that identifies the capture buffer for which to retrieve a value:

- If the value of *capture-buffer* is zero, PRXPOSN returns the entire match.
- If the value of *capture-buffer* is between 1 and the number of open parentheses in the regular expression, then PRXPOSN returns the value for that capture buffer.
- If the value of *capture-buffer* is greater than the number of open parentheses, then PRXPOSN returns a missing value.

source
specifies the text from which to extract capture buffers.

Details

The PRXPOSN function uses the results of PRXMATCH, PRXSUBSTR, PRXCHANGE, or PRXNEXT to return a capture buffer. A match must be found by one of these functions for PRXPOSN to return meaningful information.

A capture buffer is part of a match, enclosed in parentheses, that is specified in a regular expression. This function simplifies using capture buffers by returning the text for the capture buffer directly, and by not requiring a call to SUBSTR as in the case of CALL PRXPOSN.

For more information about pattern matching, see “Pattern Matching Using Perl Regular Expressions (PRX)” on page 333.

Comparisons

The PRXPOSN function is similar to the CALL PRXPOSN routine, except that it returns the capture buffer itself rather than the position and length of the capture buffer.

The Perl regular expression (PRX) functions and CALL routines work together to manipulate strings that match patterns. To see a list and short description of these functions and CALL routines, see the Character String Matching category in “Functions and CALL Routines by Category” on page 345.

Examples

Example 1: Extracting First and Last Names The following example uses PRXPOSN to extract first and last names from a data set.

```
data ReversedNames;
   input name & $32.;
   datalines;
Jones, Fred
Kavich, Kate
Turley, Ron
Dulix, Yolanda
;

data FirstLastNames;
   length first last $ 16;
   keep first last;
   retain re;
   if _N_ = 1 then
      re = prxparse('/(\w+), (\w+)/');
   set ReversedNames;
   if prxmatch(re, name) then
      do;
         last = prxposn(re, 1, name);
         first = prxposn(re, 2, name);
      end;
run;

options pageno=1 nodate ls=80 ps=64;
proc print data = FirstLastNames;
run;
```

Output 4.75 Output from PRXPOSN: First and Last Names

```
                              The SAS System                              1

                        Obs    first       last

                         1     Fred        Jones
                         2     Kate        Kavich
                         3     Ron         Turley
                         4     Yolanda     Dulix
```

Example 2: Extracting Names When Some Names Are Invalid The following example creates a data set that contains a list of names. Observations that have only a first name or only a last name are invalid. PRXPOSN extracts the valid names from the data set, and writes the names to the data set NEW.

```
data old;
   input name $60.;
   datalines;
Judith S Reaveley
Ralph F. Morgan
Jess Ennis
Carol Echols
```

```
Kelly Hansen Huff
Judith
Nick
Jones
;

data new;
   length first middle last $ 40;
   keep first middle last;
   re = prxparse('/(\S+)\s+([^\s]+\s+)?(\S+)/o');
   set old;
   if prxmatch(re, name) then
      do;
         first = prxposn(re, 1, name);
         middle = prxposn(re, 2, name);
         last = prxposn(re, 3, name);
         output;
      end;
run;

options pageno=1 nodate ls=80 ps=64;
proc print data = new;
run;
```

Output 4.76 Output of Valid Names

```
                              The SAS System                                1

                    Obs    first     middle    last

                     1     Judith    S         Reaveley
                     2     Ralph     F.        Morgan
                     3     Jess                Ennis
                     4     Carol               Echols
                     5     Kelly     Hansen    Huff
```

See Also

Functions:

"CALL PRXCHANGE Routine" on page 479
"CALL PRXDEBUG Routine" on page 482
"CALL PRXFREE Routine" on page 484
"CALL PRXNEXT Routine" on page 485
"CALL PRXPOSN Routine" on page 487
"CALL PRXSUBSTR Routine" on page 490
"CALL PRXCHANGE Routine" on page 479
"PRXCHANGE Function" on page 1039
"PRXMATCH Function" on page 1045
"PRXPAREN Function" on page 1049
"PRXPARSE Function" on page 1051

PTRLONGADD Function

Returns the pointer address as a character variable on 32-bit and 64-bit platforms.

Category: Special

Syntax

PTRLONGADD(*pointer*<,*amount*>)

Arguments

pointer
is a character constant, variable, or expression that specifies the pointer address.

amount
is a numeric constant, variable, or expression that specifies the amount to add to the address.

Tip: *amount* can be a negative number.

Details

The PTRLONGADD function performs pointer arithmetic and returns a pointer address as a character string.

Examples

The following example returns the pointer address for the variable Z.

```
data _null_;
   x='ABCDE';
   y=ptrlongadd(addrlong(x),2);
   z=peekclong(y,1);
   put z=;
run;
```

The output from the SAS log is: **z=C**

PUT Function

Returns a value using a specified format.

Category: Special

Syntax

PUT(*source*, *format*.)

Arguments

source
identifies the constant, variable, or expression whose value you want to reformat. The *source* argument can be character or numeric.

format.
contains the SAS format that you want applied to the value that is specified in the source. This argument must be the name of a format with a period and optional width and decimal specifications, not a character constant, variable, or expression. By default, if the source is numeric, the resulting string is right aligned, and if the source is character, the result is left aligned. To override the default alignment, you can add an alignment specification to a format:

-L	left aligns the value.
-C	centers the value.
-R	right aligns the value.

Restriction: The *format*. must be of the same type as the source, either character or numeric. That is, if the source is character, the format name must begin with a dollar sign, but if the source is numeric, the format name must not begin with a dollar sign.

Details

If the PUT function returns a value to a variable that has not yet been assigned a length, by default the variable length is determined by the width of the format.

Use PUT to convert a numeric value to a character value. The PUT function has no effect on which formats are used in PUT statements or which formats are assigned to variables in data sets. You cannot use the PUT function to change the type of a variable in a data set from numeric to character.

Comparisons

The PUT statement and the PUT function are similar. The PUT function returns a value using a specified format. You must use an assignment statement to store the value in a variable. The PUT statement writes a value to an external destination (either the SAS log or a destination you specify).

Examples

Example 1: Converting Numeric Values to Character Value In this example, the first statement converts the values of CC, a numeric variable, into the four-character hexadecimal format, and the second writes the same value that the PUT function returns.

```
cchex=put(cc,hex4.);
put cc hex4.;
```

Example 2: Using PUT and INPUT Functions In this example, the PUT function returns a numeric value as a character string. The value 122591 is assigned to the CHARDATE variable. The INPUT function returns the value of the character string as a SAS date value using a SAS date informat. The value 11681 is stored in the SASDATE variable.

```
numdate=122591;
chardate=put(numdate,z6.);
sasdate=input(chardate,mmddyy6.);
```

See Also

Functions:

"INPUT Function" on page 823
"INPUTC Function" on page 826
"INPUTN Function" on page 827
"PUTC Function" on page 1058,
"PUTN Function" on page 1060

Statement:

"PUT Statement" on page 1708

PUTC Function

Enables you to specify a character format at run time.

Category: Special

Syntax

PUTC(*source*, *format*.<,*w*>)

Arguments

source
specifies a character constant, variable, or expression to which you want to apply the format.

***format*.**
is a character constant, variable, or expression with a value that is the character format you want to apply to *source*.

w
is a numeric constant, variable, or expression that specifies a width to apply to the format.

Interaction: If you specify a width here, it overrides any width specification in the format.

Details

If the PUTC function returns a value to a variable that has not yet been assigned a length, by default the variable length is determined by the length of the first argument.

Comparisons

The PUTN function enables you to specify a numeric format at run time.

The PUT function is faster than PUTC because PUT lets you specify a format at compile time rather than at run time.

Examples

The PROC FORMAT step in this example creates a format, TYPEFMT., that formats the variable values 1, 2, and 3 with the name of one of the three other formats that this step creates. These three formats output responses of "positive," "negative," and "neutral" as different words, depending on the type of question. After PROC FORMAT creates the formats, the DATA step creates a SAS data set from raw data consisting of a number identifying the type of question and a response. After reading a record, the DATA step uses the value of TYPE to create a variable, RESPFMT, that contains the value of the appropriate format for the current type of question. The DATA step also creates another variable, WORD, whose value is the appropriate word for a response. The PUTC function assigns the value of WORD based on the type of question and the appropriate format.

```
proc format;
   value typefmt 1='$groupx'
                 2='$groupy'
                 3='$groupz';
   value $groupx 'positive'='agree'
                 'negative'='disagree'
                 'neutral'='notsure ';
   value $groupy 'positive'='accept'
                 'negative'='reject'
                 'neutral'='possible';

   value $groupz 'positive'='pass    '
                 'negative'='fail'
                 'neutral'='retest';
run;

data answers;
   length word $ 8;
   input type response $;
   respfmt = put(type, typefmt.);
   word = putc(response, respfmt);
   datalines;
1 positive
1 negative
1 neutral
2 positive
2 negative
2 neutral
3 positive
3 negative
3 neutral
;
```

The value of the variable WORD is **agree** for the first observation. The value of the variable WORD is **retest** for the last observation.

See Also

Functions:

"INPUT Function" on page 823
"INPUTC Function" on page 826
"INPUTN Function" on page 827
"PUT Function" on page 1056,
"PUTN Function" on page 1060

PUTN Function

Enables you to specify a numeric format at run time.

Category: Special

Syntax

PUTN(*source*, *format*.<,*w*<,*d*>>)

Arguments

source
specifies a numeric constant, variable, or expression to which you want to apply the format.

***format*.**
is a character constant, variable, or expression with a value that is the numeric format you want to apply to *source*.

w
is a numeric constant, variable, or expression that specifies a width to apply to the format.

Interaction: If you specify a width here, it overrides any width specification in the format.

d
is a numeric constant, variable, or expression that specifies the number of decimal places to use.

Interaction: If you specify a number here, it overrides any decimal-place specification in the format.

Details

If the PUTN function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

Comparisons

The PUTC function enables you to specify a character format at run time.

The PUT function is faster than PUTN because PUT lets you specify a format at compile time rather than at run time.

Examples

The PROC FORMAT step in this example creates a format, WRITFMT., that formats the variable values 1 and 2 with the name of a SAS date format. The DATA step creates a SAS data set from raw data consisting of a number and a key. After reading a record, the DATA step uses the value of KEY to create a variable, DATEFMT, that contains the value of the appropriate date format. The DATA step also creates a new variable, DATE, whose value is the formatted value of the date. PUTN assigns the value of DATE based on the value of NUMBER and the appropriate format.

```
proc format;
   value writfmt 1='date9.'
                 2='mmddyy10.';
run;
data dates;
   input number key;
   datefmt=put(key,writfmt.);
   date=putn(number,datefmt);
   datalines;
15756 1
14552 2
;
```

See Also

Functions:

"INPUT Function" on page 823

"INPUTC Function" on page 826

"INPUTN Function" on page 827

"PUT Function" on page 1056

"PUTC Function" on page 1058

PVP Function

Returns the present value for a periodic cash flow stream (such as a bond), with repayment of principal at maturity.

Category: Financial

Syntax

PVP(*A,c,n,K,k*$_0$*,y*)

Arguments

A
specifies the par value.
Range: $A > 0$

c
specifies the nominal per-year coupon rate, expressed as a fraction.
Range: $0 \leq c < 1$

n
specifies the number of coupons per year.
Range: $n > 0$ and is an integer

K
specifies the number of remaining coupons.
Range: $K > 0$ and is an integer

k_0
specifies the time from the present date to the first coupon date, expressed in terms of the number of years.
Range: $0 < k_0 \leq \frac{1}{n}$

y
specifies the nominal per-year yield-to-maturity, expressed as a fraction.
Range: $y > 0$

Details

The PVP function is based on the relationship

$$P = \sum_{k=1}^{K} \frac{c(k)}{\left(1 + \frac{y}{n}\right)^{t_k}}$$

where

$$t_k = nk_0 + k - 1$$
$$c(k) = \frac{c}{n}A \quad for \;\; k = 1, \ldots, K-1$$
$$c(K) = \left(1 + \frac{c}{n}\right)A$$

Examples

```
data _null_;
p=pvp(1000,.01,4,14,.33/2,.10);
put p;
run;
```

The value returned is 743.168.

QTR Function

Returns the quarter of the year from a SAS date value.

Category: Date and Time

Syntax

QTR(*date*)

Arguments

date
specifies a numeric constant, variable, or expression that represents a SAS date value.

Details

The QTR function returns a value of 1, 2, 3, or 4 from a SAS date value to indicate the quarter of the year in which a date value falls.

Examples

The following SAS statements produce these results.

SAS Statements	Results

`x='20jan94'd;`
`y=qtr(x);`
`put y=;`
	`y=1`

See Also

Function:

"YYQ Function" on page 1237

QUANTILE Function

Returns the quantile from a distribution that you specify.

Category: Quantile
See: "CDF Function" on page 558

Syntax

QUANTILE(*dist, probability, parm-1,...,parm-k*)

Arguments

dist
is a character constant, variable, or expression that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	**BERNOULLI**
Beta	**BETA**
Binomial	**BINOMIAL**
Cauchy	**CAUCHY**
Chi-Square	**CHISQUARE**
Exponential	**EXPONENTIAL**
F	**F**
Gamma	**GAMMA**
Geometric	**GEOMETRIC**
Hypergeometric	**HYPERGEOMETRIC**
Laplace	**LAPLACE**
Logistic	**LOGISTIC**
Lognormal	**LOGNORMAL**
Negative binomial	**NEGBINOMIAL**
Normal	**NORMAL|GAUSS**
Normal mixture	**NORMALMIX**
Pareto	**PARETO**
Poisson	**POISSON**
T	**T**
Uniform	**UNIFORM**
Wald (inverse Gaussian)	**WALD|IGAUSS**
Weibull	**WEIBULL**

Note: Except for T, F, and NORMALMIX, you can minimally identify any distribution by its first four characters. △

probability
is a numeric constant, variable, or expression that specifies the value of a random variable.

parm-1,...,parm-k
are optional *shape*, *location*, or *scale* parameters appropriate for the specific distribution.

The QUANTILE function computes the probability from various continuous and discrete distributions. For more information, see the on page 559.

Examples

SAS Statements	Results
`y=quantile('BERN',.75,.25);`	`0`
`y=quantile('BETA',0.1,3,4);`	`0.2009088789`
`y=quantile('BINOM',.4,.5,10);`	`5`
`y=quantile('CAUCHY',.85);`	`1.9626105055`
`y=quantile('CHISQ',.6,11);`	`11.529833841`
`y=quantile('EXPO',.6);`	`0.9162907319`
`y=quantile('F',.8,2,3);`	`2.8860266073`
`y=quantile('GAMMA',.4,3);`	`2.285076904`
`y=quantile('HYPER',.5,200,50,10);`	`2`
`y=quantile('LAPLACE',.8);`	`0.9162907319`
`y=quantile('LOGISTIC',.7);`	`0.8472978604`
`y=quantile('LOGNORMAL',.5);`	`1`
`y=quantile('NEGB',.5,.5,2);`	`1`
`y=quantile('NORMAL',.975);`	`1.9599639845`
`y=quantile('PARETO',.01,1);`	`1.0101010101`
`y=quantile('POISSON',.9,1);`	`2`
`y=quantile('T',.8,5);`	`0.9195437802`
`y=quantile('UNIFORM',0.25);`	`0.25`
`y=quantile('WALD',.6,2);`	`0.9526209927`
`y=quantile('WEIBULL',.6,2);`	`0.9572307621`

See Also

Functions:

"LOGCDF Function" on page 907

"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"SDF Function" on page 1120
"CDF Function" on page 558

QUOTE Function

Adds double quotation marks to a character value.

Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

QUOTE(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

Length of Returned Variable In a DATA step, if the QUOTE function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The QUOTE function adds double quotation marks, the default character, to a character value. If double quotation marks are found within the argument, they are doubled in the output.

The length of the receiving variable must be long enough to contain the argument (including trailing blanks), leading and trailing quotation marks, and any embedded quotation marks that are doubled. For example, if the argument is ABC followed by three trailing blanks, then the receiving variable must have a length of at least eight to hold "ABC###". (The character # represents a blank space.) If the receiving field is not long enough, the QUOTE function returns a blank string, and writes an invalid argument note to the log.

Examples

SAS Statements	Results

`x='A"B';`
`y=quote(x);`
`put y;`
	`"A""B"`

`x='A''B';`
`y=quote(x);`
`put y;`
	`"A'B"`

`x='Paul''s';`
`y=quote(x);`
`put y;`
	`"Paul's"`

`x='Catering Service Center      ';`
`y=quote(x);`
`put y;`
	`"Catering Service Center      "`

`x='Paul''s Catering Service          ';`
`y=quote(trim(x));`
`put y;`
	`"Paul's Catering Service"`

RANBIN Function

Returns a random variate from a binomial distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANBIN routine instead of the RANBIN function.

Syntax

RANBIN(*seed,n,p*)

Arguments

seed
is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: $seed < 2^{31}-1$

See: "Seed Values" on page 314 for more information about seed values

n
is a numeric constant, variable, or expression with an integer value that specifies the number of independent Bernoulli trials parameter.

Range: $n > 0$

p
is a numeric constant, variable, or expression that specifies the probability of success.

Range: $0 < p < 1$

Details

The RANBIN function returns a variate that is generated from a binomial distribution with mean np and variance $np(1–p)$. If $n \leq 50$, $np \leq 5$, or $n(1–p) \leq 5$, an inverse transform method applied to a RANUNI uniform variate is used. If $n > 50$, $np > 5$, and $n(1–p) > 5$, the normal approximation to the binomial distribution is used. In that case, the Box-Muller transformation of RANUNI uniform variates is used.

For a discussion about seeds and streams of data, as well as examples of using the random-number functions, see "Generating Multiple Variables from One Seed in Random-Number Functions" on page 324.

Comparisons

The CALL RANBIN routine, an alternative to the RANBIN function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069

"CALL RANBIN Routine" on page 492

RANCAU Function

Returns a random variate from a Cauchy distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANCAU routine instead of the RANCAU function.

Syntax

RANCAU(*seed*)

Arguments

seed
is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: *seed* < 2^{31} –1

See: "Seed Values" on page 314 for more information about seed values

Details

The RANCAU function returns a variate that is generated from a Cauchy distribution with location parameter 0 and scale parameter 1. An acceptance-rejection procedure applied to RANUNI uniform variates is used. If *u* and *v* are independent uniform (–1/2, 1/2) variables and $u^2 + v^2 \leq 1/4$, then *u*/*v* is a Cauchy variate. A Cauchy variate X with location parameter ALPHA and scale parameter BETA can be generated:

```
x=alpha+beta*rancau(seed);
```

For a discussion about seeds and streams of data, as well as examples of using the random-number functions, see "Generating Multiple Variables from One Seed in Random-Number Functions" on page 324.

Comparisons

The CALL RANCAU routine, an alternative to the RANCAU function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069

"CALL RANCAU Routine" on page 494

RAND Function

Generates random numbers from a distribution that you specify.

Category: Random Number

Syntax

RAND (*dist*, *parm-1*,...,*parm-k*)

Arguments

dist
is a character constant, variable, or expression that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	**BERNOULLI**
Beta	**BETA**
Binomial	**BINOMIAL**
Cauchy	**CAUCHY**
Chi-Square	**CHISQUARE**
Erlang	**ERLANG**
Exponential	**EXPONENTIAL**
F	**F**
Gamma	**GAMMA**
Geometric	**GEOMETRIC**
Hypergeometric	**HYPERGEOMETRIC**
Lognormal	**LOGNORMAL**
Negative binomial	**NEGBINOMIAL**
Normal	**NORMAL|GAUSSIAN**
Poisson	**POISSON**
T	**T**
Tabled	**TABLE**
Triangular	**TRIANGLE**
Uniform	**UNIFORM**
Weibull	**WEIBULL**

Note: Except for T and F, you can minimally identify any distribution by its first four characters. △

parm-1,...,parm-k
are *shape*, *location*, or *scale* parameters appropriate for the specific distribution.

See: "Details" on page 1071 for complete information about these parameters

Details

Generating Random Numbers The RAND function generates random numbers from various continuous and discrete distributions. Wherever possible, the simplest form of the distribution is used.

The RAND function uses the Mersenne-Twister random number generator (RNG) that was developed by Matsumoto and Nishimura (1998). The random number generator has a very long period ($2^{19937} - 1$) and very good statistical properties. The period is a Mersenne prime, which contributes to the naming of the RNG. The algorithm is a twisted generalized feedback shift register (TGFSR) that explains the latter part of the name. The TGFSR gives the RNG a very high order of equidistribution (623-dimensional with 32-bit accuracy), which means that there is a very small correlation between successive vectors of 623 pseudo-random numbers.

The RAND function is started with a single seed. However, the state of the process cannot be captured by a single seed. You cannot stop and restart the generator from its stopping point.

Reproducing a Random Number Stream

If you want to create reproducible streams of random numbers, then use the CALL STREAMINIT routine to specify a seed value for random number generation. Use the CALL STREAMINIT routine once per DATA step before any invocation of the RAND function. If you omit the call to the CALL STREAMINIT routine (or if you specify a non-positive seed value in the CALL STREAMINIT routine), then RAND uses a call to the system clock to seed itself. For more information, see CALL STREAMINIT Example 1 on page 535.

Bernoulli Distribution

x = **RAND**('BERNOULLI',*p*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \begin{cases} 1 & p = 0, x = 0 \\ p^x (1-p)^{1-x} & 0 < p < 1, x = 0, 1 \\ 1 & p = 1, x = 1 \end{cases}$$

Range: *x* = 0, 1

p
is a numeric probability of success.

Range: $0 \leq p \leq 1$

Beta Distribution

x = **RAND**('BETA',*a*,*b*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} x^{a-1}(1-x)^{b-1}$$

Range: $0 < x < 1$

a
is a numeric shape parameter.
Range: $a > 0$

b
is a numeric shape parameter.
Range: $b > 0$

Binomial Distribution

x = **RAND**('BINOMIAL',*p*,*n*)

where

x
is an integer observation from the distribution with the following probability density function:

$$f(x) = \begin{cases} 1 & p = 0, x = 0 \\ \binom{n}{x} p^x (1-p)^{n-x} & 0 < p < 1, x = 0, \ldots, n \\ 1 & p = 1, x = n \end{cases}$$

Range: $x = 0, 1, \ldots, n$

p
is a numeric probability of success.
Range: $0 \le p \le 1$

n
is an integer parameter that counts the number of independent Bernoulli trials.
Range: $n = 1, 2, \ldots$

Cauchy Distribution

x = **RAND**('CAUCHY')

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

Range: $-\infty < x < \infty$

Chi-Square Distribution

x = **RAND**('CHISQUARE',*df*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{2^{\frac{-df}{2}}}{\Gamma\left(\frac{df}{2}\right)} x^{\frac{df}{2}-1} e^{\frac{-x}{2}}$$

Range: $x > 0$

df
is a numeric degrees of freedom parameter.
Range: $df > 0$

Erlang Distribution

x = **RAND**('ERLANG',*a*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{1}{\Gamma(a)} x^{a-1} e^{-x}$$

Range: $x > 0$

a
is an integer numeric shape parameter.
Range: $a = 1, 2, \ldots$

Exponential Distribution

x = **RAND**('EXPONENTIAL')

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = e^{-x}$$

Range: $x > 0$

F Distribution

x = **RAND**('F',*ndf*, *ddf*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{\Gamma\left(\frac{ndf+ddf}{2}\right)}{\Gamma\left(\frac{ndf}{2}\right)\Gamma\left(\frac{ddf}{2}\right)} \frac{ndf^{ndf/2}ddf^{ddf/2}x^{\frac{ndf}{2}-1}}{(ddf + ndfx)^{(ndf+ddf)/2}}$$

Range: $x > 0$

ndf
is a numeric numerator degrees of freedom parameter.
Range: $ndf > 0$

ddf
is a numeric denominator degrees of freedom parameter.
Range: $ddf > 0$

Gamma Distribution

x = **RAND**('GAMMA',*a*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{1}{\Gamma(a)} x^{a-1} e^{-x}$$

Range: $x > 0$

a
is a numeric shape parameter.
Range: $a > 0$

Geometric Distribution

x = **RAND**('GEOMETRIC',*p*)

where

x
is an integer count that denotes the number of trials that are needed to obtain one success. *X* is an integer observation from the distribution with the following probability density function:

$$f(x) = \begin{cases} (1-p)^{x-1} p & 0 < p < 1, x = 1, 2, \ldots \\ 1 & p = 1, x = 1 \end{cases}$$

Range: *x* = 1, 2, ...

p
is a numeric probability of success.
Range: $0 < p \leq 1$

Hypergeometric Distribution

x = **RAND**('HYPER',*N,R,n*)

where

x
is an integer observation from the distribution with the following probability density function:

$$f(x) = \frac{\binom{R}{x}\binom{N-R}{n-x}}{\binom{N}{n}}$$

Range: $x = \max(0, (n - (N - R))), ..., \min(n, R)$

N
is an integer population size parameter.
Range: $N = 1, 2, ...$

R
is an integer number of items in the category of interest.
Range: $R = 0, 1, ..., N$

n
is an integer sample size parameter.
Range: $n = 1, 2, ..., N$

The hypergeometric distribution is a mathematical formalization of an experiment in which you draw n balls from an urn that contains N balls, R of which are red. The hypergeometric distribution is the distribution of the number of red balls in the sample of n.

Lognormal Distribution

x = **RAND**('LOGNORMAL')

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{e^{-\ln^2(x)/2}}{x\sqrt{2\pi}}$$

Range: $x > 0$

Negative Binomial Distribution

x = **RAND**('NEGBINOMIAL',*p*,*k*)

where

x
is an integer observation from the distribution with the following probability density function:

$$f(x) = \begin{cases} \binom{x+k-1}{k-1} (1-p)^x p^k & 0 < p < 1, x = 0, 1, \ldots \\ 1 & p = 1, x = 0 \end{cases}$$

Range: *x* = 0, 1, ...

k
is an integer parameter that is the number of successes. However, non-integer *k* values are allowed as well.

Range: *k* = 1, 2, ...

p
is a numeric probability of success.

Range: $0 < p \leq 1$

The negative binomial distribution is the distribution of the number of failures before *k* successes occur in sequential independent trials, all with the same probability of success, *p*.

Normal Distribution

x = **RAND**('NORMAL',<,θ,λ>)

where

x
is an observation from the normal distribution with a mean of θ and a standard deviation of λ, that has the following probability density function:

$$f(x) = \frac{1}{\lambda\sqrt{2\pi}} \; exp\left(-\frac{(x-\theta)^2}{2\lambda^2}\right)$$

Range: $-\infty < x < \infty$

θ
is the mean parameter.

Default: 0

λ
is the standard deviation parameter.

Default: 1

Range: $\lambda > 0$

Poisson Distribution

x = **RAND**('POISSON',*m*)

where

x
is an integer observation from the distribution with the following probability density function:

$$f(x) = \frac{m^x e^{-m}}{x!}$$

Range: $x = 0, 1, \ldots$

m
is a numeric mean parameter.
Range: $m > 0$

T Distribution

x = **RAND**('T',*df*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\sqrt{df\,\pi}\,\Gamma\left(\frac{df}{2}\right)}\left(1 + \frac{x^2}{df}\right)^{-\frac{df+1}{2}}$$

Range: $-\infty < x < \infty$

df
is a numeric degrees of freedom parameter.
Range: $df > 0$

Tabled Distribution

x = **RAND**('TABLE',*p1*,*p2*, ...)

where

x
is an integer observation from one of the following distributions:
If $\sum_{i=1}^{n} p_i \leq 1$, then *x* is an observation from this probability density function:

$$f(i) = p_i, \quad i = 1, 2, \ldots, n$$

and

$$f(n+1) = 1 - \sum_{i=1}^{n} p_i$$

If for some index $\sum_{i=1}^{n} p_{i \geq 1}$, then *x* is an observation from this probability density function:

$$f(i) = p_i, \quad i = 1, 2, \ldots, j-1$$

and

$$f(j) = 1 - \sum_{i=1}^{j-1} p_i$$

p1, p2, ...
are numeric probability values.

Range: $0 \leq p1, p2, \ldots \leq 1$

Restriction: The maximum number of probability parameters depends on your operating environment, but the maximum number of parameters is at least 32,767.

The tabled distribution takes on the values 1, 2, ..., *n* with specified probabilities.

Note: By using the FORMAT statement, you can map the set {1, 2, ..., *n*} to any set of *n* or fewer elements. △

Triangular Distribution

x = **RAND**('TRIANGLE',*h*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \begin{cases} \frac{2x}{h} & 0 \leq x \leq h \\ \frac{2(1-x)}{1-h} & h < x \leq 1 \end{cases}$$

where $0 \leq h \leq 1$.

Range: $0 \leq x \leq 1$

Note: The distribution can be easily shifted and scaled. △

h
is the horizontal location of the peak of the triangle.

Range: $0 \leq h \leq 1$

Uniform Distribution

x = **RAND**('UNIFORM')

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = 1$$

Range: $0 < x < 1$

The uniform random number generator that the RAND function uses is the Mersenne-Twister (Matsumoto and Nishimura 1998). This generator has a period of $2^{19937} - 1$ and 623-dimensional equidistribution up to 32-bit accuracy. This algorithm underlies the generators for the other available distributions in the RAND function.

Weibull Distribution

x = **RAND**('WEIBULL',*a*,*b*)

where

x
is an observation from the distribution with the following probability density function:

$$f(x) = \frac{a}{b^a} x^{a-1} e^{-\left(\frac{x}{b}\right)^a}$$

Range: $x \geq 0$

a
is a numeric shape parameter.
Range: $a > 0$

b
is a numeric scale parameter.
Range: $b > 0$

Examples

SAS Statements	Results
`x=rand('BERN',.75);`	`0`
`x=rand('BETA',3,0.1);`	`.99920`
`x=rand('BINOM',10,0.75);`	`10`
`x=rand('CAUCHY');`	`-1.41525`
`x=rand('CHISQ',22);`	`25.8526`
`x=rand('ERLANG', 7);`	`7.67039`
`x=rand('EXPO');`	`1.48847`
`x=rand('F',12,322);`	`1.99647`
`x=rand('GAMMA',7.25);`	`6.59588`
`x=rand('GEOM',0.02);`	`43`
`x=rand('HYPER',10,3,5);`	`1`
`x=rand('LOGN');`	`0.66522`
`x=rand('NEGB',0.8,5);`	`33`
`x=rand('NORMAL');`	`1.03507`
`x=rand('POISSON',6.1);`	`6`
`x=rand('T',4);`	`2.44646`
`x=rand('TABLE',.2,.5);`	`2`
`x=rand('TRIANGLE',0.7);`	`.63811`
`x=rand('UNIFORM');`	`.96234`
`x=rand('WEIB',0.25,2.1);`	`6.55778`

See Also

CALL Routine:

"CALL STREAMINIT Routine" on page 535

References

Fishman, G. S. 1996. *Monte Carlo: Concepts, Algorithms, and Applications*. New York: Springer-Verlag.

Fushimi, M., and S. Tezuka. 1983. "The k-Distribution of Generalized Feedback Shift Register Pseudorandom Numbers." *Communications of the ACM* 26: 516–523.

Gentle, J. E. 1998. *Random Number Generation and Monte Carlo Methods*. New York: Springer-Verlag.

Lewis, T. G., and W. H. Payne. 1973. "Generalized Feedback Shift Register Pseudorandom Number Algorithm." *Journal of the ACM* 20: 456–468.

Matsumoto, M., and Y. Kurita. 1992. "Twisted GFSR Generators." *ACM Transactions on Modeling and Computer Simulation* 2: 179–194.

Matsumoto, M., and Y. Kurita. 1994. "Twisted GFSR Generators II." *ACM Transactions on Modeling and Computer Simulation* 4: 254–266.

Matsumoto, M., and T. Nishimura. 1998. "Mersenne Twister: A 623–Dimensionally Equidistributed Uniform Pseudo-Random Number Generator." *ACM Transactions on Modeling and Computer Simulation* 8: 3–30.

Ripley, B. D. 1987. *Stochastic Simulation*. New York: Wiley.

Robert, C. P., and G. Casella. 1999. *Monte Carlo Statistical Methods*. New York: Springer-Verlag.

Ross, S. M. 1997. *Simulation*. San Diego: Academic Press.

RANEXP Function

Returns a random variate from an exponential distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANEXP routine instead of the RANEXP function.

Syntax

RANEXP(*seed*)

Arguments

seed
is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: $seed < 2^{31}-1$

See: "Seed Values" on page 314 for more information about seed values

Details

The RANEXP function returns a variate that is generated from an exponential distribution with parameter 1. An inverse transform method applied to a RANUNI uniform variate is used.

An exponential variate X with parameter LAMBDA can be generated:

```
x=ranexp(seed)/lambda;
```

An extreme value variate X with location parameter ALPHA and scale parameter BETA can be generated:

```
x=alpha-beta*log(ranexp(seed));
```

A geometric variate X with parameter P can be generated as follows:

```
x=floor(-ranexp(seed)/log(1-p));
```

For a discussion about seeds and streams of data, as well as examples of using the random-number functions, see "Generating Multiple Variables from One Seed in Random-Number Functions" on page 324.

Comparisons

The CALL RANEXP routine, an alternative to the RANEXP function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069

"CALL RANEXP Routine" on page 497

RANGAM Function

Returns a random variate from a gamma distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANGAM routine instead of the RANGAM function.

Syntax

RANGAM(*seed,a*)

Arguments

seed

is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: $seed < 2^{31}-1$

See: "Seed Values" on page 314 for more information about seed values

a

is a numeric constant, variable, or expression that specifies the shape parameter.

Range: $a > 0$

Details

The RANGAM function returns a variate that is generated from a gamma distribution with parameter a. For $a > 1$, an acceptance-rejection method due to Cheng (1977) (See "References" on page 1255) is used. For $a \leq 1$, an acceptance-rejection method due to Fishman is used (1978, Algorithm G2) (See "References" on page 1255).

A gamma variate X with shape parameter ALPHA and scale BETA can be generated:

```
x=beta*rangam(seed,alpha);
```

If 2*ALPHA is an integer, a chi-square variate X with 2*ALPHA degrees of freedom can be generated:

```
x=2*rangam(seed,alpha);
```

If N is a positive integer, an Erlang variate X can be generated:

```
x=beta*rangam(seed,N);
```

It has the distribution of the sum of N independent exponential variates whose means are BETA.

And finally, a beta variate X with parameters ALPHA and BETA can be generated:

```
y1=rangam(seed,alpha);
y2=rangam(seed,beta);
x=y1/(y1+y2);
```

For a discussion about seeds and streams of data, as well as examples of using the random-number functions, see "Generating Multiple Variables from One Seed in Random-Number Functions" on page 324.

Comparisons

The CALL RANGAM routine, an alternative to the RANGAM function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069

"CALL RANGAM Routine" on page 499

RANGE Function

Returns the range of the nonmissing values.

Category: Descriptive Statistics

Syntax

RANGE(*argument-1*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least one nonmissing argument is required. Otherwise, the function returns a missing value. The argument list can consist of a variable list, which is preceded by OF.

Details

The RANGE function returns the difference between the largest and the smallest of the nonmissing arguments.

Examples

SAS Statements	Results
`x0=range(.,.);`	`.`
`x1=range(-2,6,3);`	`8`
`x2=range(2,6,3,.);`	`4`
`x3=range(1,6,3,1);`	`5`
`x4=range(of x1-x3);`	`4`

RANK Function

Returns the position of a character in the ASCII or EBCDIC collating sequence.

Category: Character

Restriction: "I18N Level 0" on page 313

See: RANK Function in the documentation for your operating environment.

Syntax

RANK(*x*)

Arguments

x
specifies a character constant, variable, or expression.

Details

The RANK function returns an integer that represents the position of the first character in the character expression. The result depends on your operating environment.

Examples

SAS Statements	Results	
	ASCII	EBCDIC
`n=rank('A');` `put n;`	`65`	`193`

See Also

Functions:

"BYTE Function" on page 431
"COLLATE Function" on page 589

RANNOR Function

Returns a random variate from a normal distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANNOR routine instead of the RANNOR function.

Syntax

RANNOR(*seed*)

Arguments

seed
is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: $seed < 2^{31}-1$

See: "Seed Values" on page 314 for more information about seed values

Details

The RANNOR function returns a variate that is generated from a normal distribution with mean 0 and variance 1. The Box-Muller transformation of RANUNI uniform variates is used.

A normal variate X with mean MU and variance S2 can be generated with this code:

```
x=MU+sqrt(S2)*rannor(seed);
```

A lognormal variate X with mean *exp*(MU + S2/2) and variance *exp*(2*MU + 2*S2) –*exp*(2*MU + S2) can be generated with this code:

```
x=exp(MU+sqrt(S2)*rannor(seed));
```

For a discussion about seeds and streams of data, as well as examples of using the random-number functions, see "Generating Multiple Variables from One Seed in Random-Number Functions" on page 324.

Comparisons

The CALL RANNOR routine, an alternative to the RANNOR function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069

"CALL RANNOR Routine" on page 501

RANPOI Function

Returns a random variate from a Poisson distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANPOI routine instead of the RANPOI function.

Syntax

RANPOI(*seed,m*)

Arguments

seed

is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: $seed < 2^{31}-1$

See: "Seed Values" on page 314 for more information about seed values

m

is a numeric constant, variable, or expression that specifies the mean of the distribution.

Range: $m \geq 0$

Details

The RANPOI function returns a variate that is generated from a Poisson distribution with mean *m*. For $m < 85$, an inverse transform method applied to a RANUNI uniform variate is used (Fishman 1976) (See "References" on page 1255). For $m \geq 85$, the normal

approximation of a Poisson random variable is used. To expedite execution, internal variables are calculated only on initial calls (that is, with each new *m*).

For a discussion about seeds and streams of data, as well as examples of using the random-number functions, see "Generating Multiple Variables from One Seed in Random-Number Functions" on page 324.

Comparisons

The CALL RANPOI routine, an alternative to the RANPOI function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069

"CALL RANPOI Routine" on page 508

RANTBL Function

Returns a random variate from a tabled probability distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANTBL routine instead of the RANTBL function.

Syntax

RANTBL(*seed,p*1 ,... p_i... ,p_n)

Arguments

seed

is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: *seed* < 2^{31}–1

See: "Seed Values" on page 314 for more information about seed values

p_i

is a numeric constant, variable, or expression.

Range: $0 \le p_i \le 1$ for $0 < i \le n$

Details

The RANTBL function returns a variate that is generated from the probability mass function defined by p_1 through p_n . An inverse transform method applied to a RANUNI uniform variate is used. RANTBL returns

$$
\begin{array}{ll}
1 & with\ probability\ p_1 \\
2 & with\ probability\ p_2 \\
. & \\
. & \\
. & \\
n & with\ probability\ p_n \\
n+1 & with\ probability\ 1-\sum_{i=1}^{n} p_i \quad if\ \sum_{i=1}^{n} p_i \leq 1
\end{array}
$$

If, for some index $j<n$, $\sum_{i=1}^{j} p_i \geq 1$, RANTBL returns only the indices 1 through j with the probability of occurrence of the index j equal to $1-\sum_{i=1}^{j-1} p_i$.

Let n=3 and P1, P2, and P3 be three probabilities with P1+P2+P3=1, and M1, M2, and M3 be three variables. The variable X in these statements

```
array m{3} m1-m3;
x=m{rantbl(seed,of p1-p3)};
```

will be assigned one of the values of M1, M2, or M3 with probabilities of occurrence P1, P2, and P3, respectively.

For a discussion and example of an effective use of the random number CALL routines, see "Starting, Stopping, and Restarting a Stream" on page 326.

Comparisons

The CALL RANTBL routine, an alternative to the RANTBL function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069
"CALL RANTBL Routine" on page 510

RANTRI Function

Returns a random variate from a triangular distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANTRI routine instead of the RANTRI function.

Syntax

RANTRI(*seed,h*)

Arguments

seed
is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: $seed < 2^{31}-1$

See: "Seed Values" on page 314 for more information about seed values

h
is a numeric constant, variable, or expression that specifies the mode of the distribution.

range: $0 < h < 1$

Details

The RANTRI function returns a variate that is generated from the triangular distribution on the interval (0,1) with parameter *h*, which is the modal value of the distribution. An inverse transform method applied to a RANUNI uniform variate is used.

A triangular distribution X on the interval (A,B) with mode C, where A ≤ C ≤ B, can be generated:

```
x=(b-a)*rantri(seed,(c-a)/(b-a))+a;
```

For a discussion about seeds and streams of data, as well as examples of using the random-number functions, see "Generating Multiple Variables from One Seed in Random-Number Functions" on page 324.

Comparisons

The CALL RANTRI routine, an alternative to the RANTRI function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069
"CALL RANTRI Routine" on page 513

RANUNI Function

Returns a random variate from a uniform distribution.

Category: Random Number

Tip: If you want to change the seed value during execution, you must use the CALL RANUNI routine instead of the RANUNI function.

Syntax

RANUNI(*seed*)

Arguments

seed

is a numeric constant, variable, or expression with an integer value. If *seed* ≤ 0, the time of day is used to initialize the seed stream.

Range: $seed < 2^{31}-1$

See: "Seed Values" on page 314 for more information about seed values

Details

The RANUNI function returns a number that is generated from the uniform distribution on the interval (0,1) using a prime modulus multiplicative generator with modulus $2^{31}-$ and multiplier 397204094 (Fishman and Moore 1982) (See "References" on page 1255).

You can use a multiplier to change the length of the interval and an added constant to move the interval. For example,

```
random_variate=a*ranuni(seed)+b;
```

returns a number that is generated from the uniform distribution on the interval (b,a+b).

Comparisons

The CALL RANUNI routine, an alternative to the RANUNI function, gives greater control of the seed and random number streams.

See Also

Functions and CALL routines:

"RAND Function" on page 1069

"CALL RANUNI Routine" on page 515

RENAME Function

Renames a member of a SAS library, an entry in a SAS catalog, an external file, or a directory.

Category: External Files

Category: SAS File I/O

See: The RENAME Function in the documentation for your operating environment.

Syntax

RENAME(*old-name*, *new-name* <, *type*<, *description* <, *password* <, *generation*>>>>)

Arguments

old-name
specifies a character constant, variable, or expression that is the current name of a member of a SAS library, an entry in a SAS catalog, an external file, or an external directory.

For a data set, *old-name* can be a one-level or two-level name. For a catalog entry, *old-name* can be a one-level, two-level, or four-level name. For an external file or directory, *old-name* must be the full pathname of the file or the directory. If the value for *old-name* is not specified, then SAS uses the current directory.

new-name
specifies a character constant, variable, or expression that is the new one-level name for the library member, catalog entry, external file, or directory.

type
is a character constant, variable, or expression that specifies the type of element to rename. *Type* can be a null argument, or one of the following values:

ACCESS	specifies a SAS/ACCESS descriptor that was created using SAS/ACCESS software.
CATALOG	specifies a SAS catalog or catalog entry.
DATA	specifies a SAS data set.
VIEW	specifies a SAS data set view.
FILE	specifies an external file or directory.

Default: 'DATA'

description
specifies a character constant, variable, or expression that is the description of a catalog entry. You can specify *description* only when the value of *type* is CATALOG. *Description* can be a null argument.

password
is a character constant, variable, or expression that specifies the password for the data set that is being renamed. *Password* can be a null argument.

generation
is a numeric constant, variable, or expression that specifies the generation number of the data set that is being renamed. *Generation* can be a null argument.

Details

You can use the RENAME function to rename members of a SAS library or entries in a SAS catalog. SAS returns 0 if the operation was successful, and a value other than 0 if the operation was not successful.

To rename an entry in a catalog, specify the four-level name for *old-name* and a one-level name for *new-name*. You must specify CATALOG for *type* when you rename an entry in a catalog.

Operating Environment Information: Use RENAME in directory-based operating environments only. If you use RENAME in a mainframe operating environment, SAS generates an error. △

Examples

Example 1: Renaming Data Sets and Catalog Entries The following examples rename a SAS data set from DATA1 to DATA2, and also rename a catalog entry from A.SCL to B.SCL.

```
rc1=rename('mylib.data1', 'data2');
rc2=rename('mylib.mycat.a.scl', 'b', 'catalog');
```

Example 2: Renaming an External File The following examples rename external files.

```
   /* Rename a file that is located in another directory. */
rc=rename('/local/u/testdir/first',
          '/local/u/second', 'file');
   /* Rename a PC file. */
rc=rename('d:\temp', 'd:\testfile', 'file');
```

Example 3: Renaming a Directory The following example renames a directory in the UNIX operating environment.

```
rc=rename('/local/u/testdir/', '/local/u/oldtestdir', 'file');
```

Example 4: Renaming a Generation Data Set The following example renames the generation data set WORK.ONE to WORK.TWO, where the password for WORK.ONE#003 is *my-password*.

```
rc=rename('work.one','two',,,3,'my-password');
```

See Also

Functions:

"FDELETE Function" on page 682
"FILEEXIST Function" on page 689
"EXIST Function" on page 675

REPEAT Function

Returns a character value that consists of the first argument repeated n+1 times.

Category: Character

Restriction: "I18N Level 1" on page 314

Syntax

REPEAT(*argument,n*)

Arguments

argument
specifies a character constant, variable, or expression.

n
specifies the number of times to repeat *argument*.

Restriction: *n* must be greater than or equal to 0.

Details

In a DATA step, if the REPEAT function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The REPEAT function returns a character value consisting of the first argument repeated *n* times. Thus, the first argument appears *n*+1 times in the result.

Examples

SAS Statements	Results
`x=repeat('ONE',2);` `put x;`	`ONEONEONE`

RESOLVE Function

Returns the resolved value of the argument after it has been processed by the macro facility.

Category: Macro

Syntax

RESOLVE(*argument*)

Arguments

argument
is a character constant, variable, or expression with a value that is a macro expression.

Details

If the RESOLVE function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

RESOLVE is fully documented in *SAS Macro Language: Reference*.

See Also

Function:

"SYMGET Function" on page 1151

REVERSE Function

Reverses a character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KREVERSE in *SAS National Language Support (NLS): Reference Guide*.

Syntax

REVERSE(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

In a DATA step, if the REVERSE function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the first argument.

The last character in the argument becomes the first character in the result, the next-to-last character in the argument becomes the second character in the result, and so on.

Note: Trailing blanks in the argument become leading blanks in the result. △

Examples

SAS Statements	Results
	`----+----1`
`backward=reverse('xyz  ');` `put backward $5.;`	`  zyx`

REWIND Function

Positions the data set pointer at the beginning of a SAS data set.

Category: SAS File I/O

Syntax

REWIND(*data-set-id*)

Arguments

data-set-id
is a numeric variable that specifies the data set identifier that the OPEN function returns.

Restriction: The data set cannot be opened in IS mode.

Details

REWIND returns 0 if the operation was successful, ≠0 if it was not successful. After a call to REWIND, a call to FETCH reads the first observation in the data set.

If there is an active WHERE clause, REWIND moves the data set pointer to the first observation that satisfies the WHERE condition.

Examples

This example calls FETCHOBS to fetch the tenth observation in the data set MYDATA. Next, the example calls REWIND to return to the first observation and fetch the first observation.

```
%let dsid=%sysfunc(open(mydata,i));
%let rc=%sysfunc(fetchobs(&dsid,10));
%let rc=%sysfunc(rewind(&dsid));
%let rc=%sysfunc(fetch(&dsid));
```

See Also

Functions:

"FETCH Function" on page 684
"FETCHOBS Function" on page 685
"FREWIND Function" on page 773
"NOTE Function" on page 956
"OPEN Function" on page 980
"POINT Function" on page 1010

RIGHT Function

Right aligns a character expression.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KRIGHT in *SAS National Language Support (NLS): Reference Guide*.

Syntax

RIGHT(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

In a DATA step, if the RIGHT function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the first argument.

The RIGHT function returns an argument with trailing blanks moved to the start of the value. The length of the result is the same as the length of the argument.

Examples

SAS Statements	Results
	`----+----1----+`

`a='Due Date  ';`
`b=right(a);`
`put a $10.;`
`put b $10.;`
	`Due Date`
	`  Due Date`

See Also

Functions:

"COMPRESS Function" on page 604

"LEFT Function" on page 881

"TRIM Function" on page 1173

RMS Function

Returns the root mean square of the nonmissing arguments.

Category: Descriptive Statistics

Syntax

RMS(*argument*<,*argument*,...>)

Arguments

argument
is a numeric constant, variable, or expression.

Tip: The argument list can consist of a variable list, which is preceded by OF.

Details

The root mean square is the square root of the arithmetic mean of the squares of the values. If all the arguments are missing values, then the result is a missing value. Otherwise, the result is the root mean square of the non-missing values.

Let n be the number of arguments with non-missing values, and let $x_1, x_2, \ldots, x_n$ be the values of those arguments. The root mean square is

$$\sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}}$$

Examples

SAS Statements	Results
`x1=rms(1,7);`	`5`
`x2=rms(.,1,5,11);`	`7`
`x3=rms(of x1-x2);`	`6.0827625303`

ROUND Function

Rounds the first argument to the nearest multiple of the second argument, or to the nearest integer when the second argument is omitted.

Category: Truncation

Syntax

ROUND (*argument* <,*rounding-unit*>)

Arguments

argument
is a numeric constant, variable, or expression to be rounded.

rounding-unit
is a positive, numeric constant, variable, or expression that specifies the rounding unit.

Details

Basic Concepts The ROUND function rounds the first argument to a value that is very close to a multiple of the second argument. The result might not be an exact multiple of the second argument.

Differences between Binary and Decimal Arithmetic Computers use binary arithmetic with finite precision. If you work with numbers that do not have an exact binary representation, computers often produce results that differ slightly from the results that are produced with decimal arithmetic.

For example, the decimal values 0.1 and 0.3 do not have exact binary representations. In decimal arithmetic, 3*0.1 is exactly equal to 0.3, but this equality is not true in binary arithmetic. As the following example shows, if you write these two values in SAS, they appear the same. If you compute the difference, however, you can see that the values are different.

```
data _null_;
   point_three=0.3;
   three_times_point_one=3*0.1;
   difference=point_three - three_times_point_one;
   put point_three= ;
   put three_times_point_one= ;
   put difference= ;
run;
```

The following lines are written to the SAS log:

```
point_three= 0.3
three_times_point_one= 0.3
difference= -5.55112E-17
```

Operating Environment Information: The example above was executed in a z/OS environment. If you use other operating environments, the results will be slightly different. △

The Effects of Rounding Rounding by definition finds an exact multiple of the rounding unit that is closest to the value to be rounded. For example, 0.33 rounded to the nearest tenth equals 3*0.1 or 0.3 in decimal arithmetic. In binary arithmetic, 0.33 rounded to the nearest tenth equals 3*0.1, and not 0.3, because 0.3 is not an exact multiple of one tenth in binary arithmetic.

The ROUND function returns the value that is based on decimal arithmetic, even though this value is sometimes not the exact, mathematically correct result. In the example **ROUND(0.33,0.1)**, ROUND returns 0.3 and not 3*0.1.

Expressing Binary Values If the characters "0.3" appear as a constant in a SAS program, the value is computed by the standard informat as 3/10. To be consistent with the standard informat, **ROUND(0.33,0.1)** computes the result as 3/10, and the following statement produces the results that you would expect.

```
if round(x,0.1) = 0.3 then
   ... more SAS statements ...
```

However, if you use the variable Y instead of the constant 0.3, as the following statement shows, the results might be unexpected depending on how the variable Y is computed.

```
if round(x,0.1) = y then
   ... more SAS statements ...
```

If SAS reads Y as the characters "0.3" using the standard informat, the result is the same as if a constant 0.3 appeared in the IF statement. If SAS reads Y with a different informat, or if a program other than SAS reads Y, then there is no guarantee that the characters "0.3" would produce a value of exactly 3/10. Imprecision can also be caused by computation involving numbers that do not have exact binary representations, or by porting data sets from one operating environment to another that has a different floating-point representation.

If you know that Y is a decimal number with one decimal place, but are not certain that Y has exactly the same value as would be produced by the standard informat, it is better to use the following statement:

```
if round(x,0.1) = round(y,0.1) then
   ... more SAS statements ...
```

Testing for Approximate Equality You should not use the ROUND function as a general method to test for approximate equality. Two numbers that differ only in the least significant bit can round to different values if one number rounds down and the other number rounds up. Testing for approximate equality depends on how the numbers have been computed. If both numbers are computed to high relative precision, you could test for approximate equality by using the ABS and the MAX functions, as the following example shows.

```
if abs(x-y) <= 1e-12 * max( abs(x), abs(y) ) then
   ... more SAS statements ...
```

Producing Expected Results In general, **ROUND(argument, rounding-unit)** produces the result that you expect from decimal arithmetic if the result has no more than nine significant digits and any of the following conditions are true:

- The rounding unit is an integer.
- The rounding unit is a power of 10 greater than or equal to 1e-15. *
- The result that you expect from decimal arithmetic has no more than four decimal places.

For example:

```
data rounding;
   d1 = round(1234.56789,100)     - 1200;
   d2 = round(1234.56789,10)      - 1230;
   d3 = round(1234.56789,1)       - 1235;
   d4 = round(1234.56789,.1)      - 1234.6;
   d5 = round(1234.56789,.01)     - 1234.57;
   d6 = round(1234.56789,.001)    - 1234.568;
   d7 = round(1234.56789,.0001)   - 1234.5679;
   d8 = round(1234.56789,.00001)  - 1234.56789;
   d9 = round(1234.56789,.1111)   - 1234.5432;
      /* d10 has too many decimal places in the value for */
      /* rounding-unit.                                   */
   d10 = round(1234.56789,.11111) - 1234.54321;
run;

proc print data=rounding noobs;
run;
```

* If the rounding unit is less than one, ROUND treats it as a power of 10 if the reciprocal of the rounding unit differs from a power of 10 in at most the three or four least significant bits.

The following output shows the results.

Output 4.77 Results of Rounding Based on the Value of the Rounding Unit

```
                                The SAS System                                 1

         d1     d2     d3     d4     d5     d6     d7     d8     d9       d10

          0      0      0      0      0      0      0      0      0    -2.2737E-13
```

Operating Environment Information: The example above was executed in a z/OS environment. If you use other operating environments, the results will be slightly different. △

When the Rounding Unit Is the Reciprocal of an Integer When the rounding unit is the reciprocal of an integer *, the ROUND function computes the result by dividing by the integer. Therefore, you can safely compare the result from ROUND with the ratio of two integers, but not with a multiple of the rounding unit. For example:

```
data rounding2;
   drop pi unit;
   pi = arcos(-1);
   unit=1/7;
   d1=round(pi,unit) - 22/7;
   d2=round(pi, unit) - 22*unit;
run;

proc print data=rounding2 noobs;
run;
```

The following output shows the results.

Output 4.78 Results of Rounding by the Reciprocal of an Integer

```
                                The SAS System                                 1

                               d1         d2

                                0    2.2204E-16
```

* ROUND treats the rounding unit as a reciprocal of an integer if the reciprocal of the rounding unit differs from an integer in at most the three or four least significant bits.

Operating Environment Information: The example above was executed in an z/OS environment. If you use other operating environments, the results will be slightly different. △

Computing Results in Special Cases The ROUND function computes the result by multiplying an integer by the rounding unit when all of the following conditions are true:

- The rounding unit is not an integer.
- The rounding unit is not a power of 10.
- The rounding unit is not the reciprocal of an integer.
- The result that you expect from decimal arithmetic has no more than four decimal places.

For example:

```
data _null_;
   difference=round(1234.56789,.11111) - 11111*.11111;
   put difference=;
run;
```

The following line is written to the SAS log:

```
   difference=0
```

Operating Environment Information: The example above was executed in a z/OS environment. If you use other operating environments, the results might be slightly different. △

Computing Results When the Value Is Halfway between Multiples of the Rounding Unit When the value to be rounded is approximately halfway between two multiples of the rounding unit, the ROUND function rounds up the absolute value and restores the original sign. For example:

```
options pageno=1 nodate ls=80 ps=64;

data test;
   do i=8 to 17;
      value=0.5 - 10**(-i);
      round=round(value);
      output;
   end;
   do i=8 to 17;
      value=-0.5 + 10**(-i);
      round=round(value);
      output;
   end;
run;

proc print data=test noobs;
   format value 19.16;
run;
```

The following output shows the results.

Output 4.79 Results of Rounding When Values Are Halfway between Multiples of the Rounding Unit

The SAS System 1

i	value	round
8	0.4999999900000000	0
9	0.4999999990000000	0
10	0.4999999999000000	0
11	0.4999999999900000	0
12	0.4999999999990000	0
13	0.4999999999999000	1
14	0.4999999999999900	1
15	0.4999999999999900	1
16	0.5000000000000000	1
17	0.5000000000000000	1
8	-0.4999999900000000	0
9	-0.4999999990000000	0
10	-0.4999999999000000	0
11	-0.4999999999900000	0
12	-0.4999999999990000	0
13	-0.4999999999999000	-1
14	-0.4999999999999900	-1
15	-0.4999999999999900	-1
16	-0.5000000000000000	-1
17	-0.5000000000000000	-1

Operating Environment Information: The example above was executed in a z/OS environment. If you use other operating environments, the results might be slightly different. △

The approximation is relative to the size of the value to be rounded, and is computed in a manner that is shown in the following DATA step. This DATA step code will not always produce results exactly equivalent to the ROUND function.

```
data testfile;
   do i = 1 to 17;
      value = 0.5 - 10**(-i);
      epsilon = min(1e-6, value * 1e-12);
      temp = value + .5 + epsilon;
      fraction = modz(temp, 1);
      round = temp - fraction;
      output;
   end;
run;
```

Comparisons

The ROUND function is the same as the ROUNDE function except when the first argument is halfway between the two nearest multiples of the second argument, ROUNDE returns an even multiple. ROUND returns the multiple with the larger absolute value.

The ROUNDZ function returns a multiple of the rounding unit without trying to make the result match the result that is computed with decimal arithmetic.

Examples

The following example compares the results that are returned by the ROUND function with the results that are returned by the ROUNDE function. The output was generated from the UNIX operating environment.

```
options pageno=1 nodate linesize=80 pagesize=60;

data results;
   do x=0 to 4 by .25;
      Rounde=rounde(x);
      Round=round(x);
      output;
   end;
run;

proc print data=results noobs;
run;
```

The following output shows the results.

Output 4.80 Results That Are Returned by the ROUND and ROUNDE Functions

```
                          The SAS System                          1

        x      Rounde     Round

     0.00        0         0
     0.25        0         0
     0.50        0         1
     0.75        1         1
     1.00        1         1
     1.25        1         1
     1.50        2         2
     1.75        2         2
     2.00        2         2
     2.25        2         2
     2.50        2         3
     2.75        3         3
     3.00        3         3
     3.25        3         3
     3.50        4         4
     3.75        4         4
     4.00        4         4
```

See Also

Functions:

"CEIL Function" on page 573
"CEILZ Function" on page 575
"FLOOR Function" on page 757
"FLOORZ Function" on page 758
"INT Function" on page 829
"INTZ Function" on page 861
"ROUNDE Function" on page 1106
"ROUNDZ Function" on page 1108

ROUNDE Function

Rounds the first argument to the nearest multiple of the second argument, and returns an even multiple when the first argument is halfway between the two nearest multiples.

Category: Truncation

Syntax

ROUNDE (*argument* <,*rounding-unit*>)

Arguments

argument
is a numeric constant, variable, or expression to be rounded.

rounding-unit
is a positive, numeric constant, variable, or expression that specifies the rounding unit.

Details

The ROUNDE function rounds the first argument to the nearest multiple of the second argument. If you omit the second argument, ROUNDE uses a default value of 1 for *rounding-unit*.

Comparisons

The ROUNDE function is the same as the ROUND function except when the first argument is halfway between the two nearest multiples of the second argument, ROUNDE returns an even multiple. ROUND returns the multiple with the larger absolute value.

Examples

The following example compares the results that are returned by the ROUNDE function with the results that are returned by the ROUND function.

```
options pageno=1 nodate linesize=80 pagesize=60;

data results;
   do x=0 to 4 by .25;
      Rounde=rounde(x);
      Round=round(x);
      output;
   end;
run;

proc print data=results noobs;
run;
```

The following output shows the results.

Output 4.81 Results That are Returned by the ROUNDE and ROUND Functions

```
                          The SAS System                          1

        x      Rounde     Round

     0.00        0         0
     0.25        0         0
     0.50        0         1
     0.75        1         1
     1.00        1         1
     1.25        1         1
     1.50        2         2
     1.75        2         2
     2.00        2         2
     2.25        2         2
     2.50        2         3
     2.75        3         3
     3.00        3         3
     3.25        3         3
     3.50        4         4
     3.75        4         4
     4.00        4         4
```

See Also

Function:

"CEIL Function" on page 573
"CEILZ Function" on page 575
"FLOOR Function" on page 757
"FLOORZ Function" on page 758
"INT Function" on page 829
"INTZ Function" on page 861
"ROUND Function" on page 1099
"ROUNDZ Function" on page 1108

ROUNDZ Function

Rounds the first argument to the nearest multiple of the second argument, using zero fuzzing.

Category: Truncation

Syntax

ROUNDZ (*argument* <,*rounding-unit*>)

Arguments

argument
is a numeric constant, variable, or expression to be rounded.

rounding-unit
is a positive, numeric constant, variable, or expression that specifies the rounding unit.

Details

The ROUNDZ function rounds the first argument to the nearest multiple of the second argument. If you omit the second argument, ROUNDZ uses a default value of 1 for *rounding-unit*.

Comparisons

The ROUNDZ function is the same as the ROUND function except that:

- ROUNDZ returns an even multiple when the first argument is exactly halfway between the two nearest multiples of the second argument. ROUND returns the multiple with the larger absolute value when the first argument is approximately halfway between the two nearest multiples.
- When the rounding unit is less than one and not the reciprocal of an integer, the result that is returned by ROUNDZ might not agree exactly with the result from decimal arithmetic. ROUNDZ does not fuzz the result. ROUND performs extra computations, called fuzzing, to try to make the result agree with decimal arithmetic.

Examples

Example 1: Comparing Results from the ROUNDZ and ROUND Functions The following example compares the results that are returned by the ROUNDZ and the ROUND function.

```
options pageno=1 nodate linesize=60 pagesize=60;

data test;
   do i=10 to 17;
      Value=2.5 - 10**(-i);
      Roundz=roundz(value);
      Round=round(value);
      output;
   end;
   do i=16 to 12 by -1;
      value=2.5 + 10**(-i);
      roundz=roundz(value);
      round=round(value);
      output;
   end;
run;

proc print data=test noobs;
   format value 19.16;
run;
```

The following output shows the results.

Output 4.82 Results That Are Returned by the ROUNDZ and ROUND Functions

```
                          The SAS System                          1

 i                  Value     Roundz     Round

10     2.4999999999000000        2         2
11     2.4999999999900000        2         2
12     2.4999999999990000        2         3
13     2.4999999999999000        2         3
14     2.4999999999999900        2         3
15     2.4999999999999900        2         3
16     2.5000000000000000        2         3
17     2.5000000000000000        2         3
16     2.5000000000000000        2         3
15     2.5000000000000000        3         3
14     2.5000000000000100        3         3
13     2.5000000000001000        3         3
12     2.5000000000010000        3         3
```

Example 2: Sample Output from the ROUNDZ Function

These examples show the results that are returned by ROUNDZ.

SAS Statement	Results

`var1=223.456;`
`x=roundz(var1,1);`
`put x 9.5;`
	`223.00000`

`var2=223.456;`
`x=roundz(var2,.01);`
`put x 9.5;`
	`223.46000`

`x=roundz(223.456,100);` `put x 9.5;`	`200.00000`
`x=roundz(223.456);` `put x 9.5;`	`223.00000`
`x=roundz(223.456,.3);` `put x 9.5;`	`223.50000`

See Also

Functions:

"ROUND Function" on page 1099
"ROUNDE Function" on page 1106

SAVING Function

Returns the future value of a periodic saving.

Category: Financial

Syntax

SAVING(*f,p,r,n*)

Arguments

f
is numeric, the future amount (at the end of n periods).

Range: $f \geq 0$

p
is numeric, the fixed periodic payment.

Range: $p \geq 0$

r
is numeric, the periodic interest rate expressed as a decimal.

Range: $r \geq 0$

n
is an integer, the number of compounding periods.

Range: $n \geq 0$

Details

The SAVING function returns the missing argument in the list of four arguments from a periodic saving. The arguments are related by

$$f = \frac{p\,(1+r)\left((1+r)^{n}-1\right)}{r}$$

One missing argument must be provided. It is then calculated from the remaining three. No adjustment is made to convert the results to round numbers.

Examples

A savings account pays a 5 percent nominal annual interest rate, compounded monthly. For a monthly deposit of $100, the number of payments that are needed to accumulate at least $12,000, can be expressed as

```
number=saving(12000,100,.05/12,.);
```

The value returned is 97.18 months. The fourth argument is set to missing, which indicates that the number of payments is to be calculated. The 5 percent nominal annual rate is converted to a monthly rate of 0.05/12. The rate is the fractional (not the percentage) interest rate per compounding period.

SCAN Function

Returns the *n*th word from a character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KSCAN in *SAS National Language Support (NLS): Reference Guide*.

Syntax

SCAN(*string*, *count*<,*charlist* <,*modifiers*>>)

Arguments

string
specifies a character constant, variable, or expression.

count
is a nonzero numeric constant, variable, or expression that has an integer value that specifies the number of the word in the character string that you want SCAN to select. For example, a value of 1 indicates the first word, a value of 2 indicates the second word, and so on. The following rules apply:

- If *count* is positive, SCAN counts words from left to right in the character string.
- If *count* is negative, SCAN counts words from right to left in the character string.

charlist
specifies an optional character expression that initializes a list of characters. This list determines which characters are used as the delimiters that separate words. The following rules apply:

- By default, all characters in *charlist* are used as delimiters.
- If you specify the K modifier in the *modifier* argument, then all characters that are *not* in *charlist* are used as delimiters.

Tip: You can add more characters to *charlist* by using other modifiers.

modifier
specifies a character constant, a variable, or an expression in which each non-blank character modifies the action of the SCAN function. Blanks are ignored. You can use the following characters as modifiers:

a or A	adds alphabetic characters to the list of characters.
b or B	scans backward from right to left instead of from left to right, regardless of the sign of the *count* argument.
c or C	adds control characters to the list of characters.
d or D	adds digits to the list of characters.
f or F	adds an underscore and English letters (that is, valid first characters in a SAS variable name using VALIDVARNAME=V7) to the list of characters.
g or G	adds graphic characters to the list of characters. Graphic characters are characters that, when printed, produce an image on paper.
h or H	adds a horizontal tab to the list of characters.

i or I — ignores the case of the characters.

k or K — causes all characters that are not in the list of characters to be treated as delimiters. That is, if K is specified, then characters that are in the list of characters are kept in the returned value rather than being omitted because they are delimiters. If K is not specified, then all characters that are in the list of characters are treated as delimiters.

l or L — adds lowercase letters to the list of characters.

m or M — specifies that multiple consecutive delimiters, and delimiters at the beginning or end of the *string* argument, refer to words that have a length of zero. If the M modifier is not specified, then multiple consecutive delimiters are treated as one delimiter, and delimiters at the beginning or end of the *string* argument are ignored.

n or N — adds digits, an underscore, and English letters (that is, the characters that can appear in a SAS variable name using VALIDVARNAME=V7) to the list of characters.

o or O — processes the *charlist* and *modifier* arguments only once, rather than every time the SCAN function is called.

Tip: Using the O modifier in the DATA step (excluding WHERE clauses), or in the SQL procedure can make SCAN run faster when you call it in a loop where the *charlist* and *modifier* arguments do not change. The O modifier applies separately to each instance of the SCAN function in your SAS code, and does *not* cause all instances of the SCAN function to use the same delimiters and modifiers.

p or P — adds punctuation marks to the list of characters.

q or Q — ignores delimiters that are inside of substrings that are enclosed in quotation marks. If the value of the *string* argument contains unmatched quotation marks, then scanning from left to right will produce different words than scanning from right to left.

r or R — removes leading and trailing blanks from the word that SCAN returns.

Tip: If you specify both the Q and R modifiers, then the SCAN function first removes leading and trailing blanks from the word. Then, if the word begins with a quotation mark, SCAN also removes one layer of quotation marks from the word.

s or S — adds space characters to the list of characters (blank, horizontal tab, vertical tab, carriage return, line feed, and form feed).

t or T — trims trailing blanks from the *string* and *charlist* arguments.

Tip: If you want to remove trailing blanks from only one character argument instead of both character arguments, then use the TRIM function instead of the SCAN function with the T modifier.

u or U — adds uppercase letters to the list of characters.

w or W — adds printable (writable) characters to the list of characters.

x or X — adds hexadecimal characters to the list of characters.

Tip: If the *modifier* argument is a character constant, then enclose it in quotation marks. Specify multiple modifiers in a single set of quotation marks. A *modifier* argument can also be expressed as a character variable or expression.

Details

Definition of "Delimiter" and "Word" A delimiter is any of several characters that are used to separate words. You can specify the delimiters in the *charlist* and *modifier* arguments.

If you specify the Q modifier, then delimiters inside of substrings that are enclosed in quotation marks are ignored.

In the SCAN function, "word" refers to a substring that has all of the following characteristics:

- is bounded on the left by a delimiter or the beginning of the string
- is bounded on the right by a delimiter or the end of the string
- contains no delimiters

A word can have a length of zero if there are delimiters at the beginning or end of the string, or if the string contains two or more consecutive delimiters. However, the SCAN function ignores words that have a length of zero unless you specify the M modifier.

Note: The definition of "word" is the same in both the SCAN and COUNTW functions. △

Using Default Delimiters in ASCII and EBCDIC Environments If you use the SCAN function with only two arguments, then the default delimiters depend on whether your computer uses ASCII or EBCDIC characters.

- If your computer uses ASCII characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ^ ¦

  In ASCII environments that do not contain the ^ character, the SCAN function uses the ~ character instead.
- If your computer uses EBCDIC characters, then the default delimiters are as follows:

  blank ! $ % & ( )* + , - . / ; < ¬ | ¢ ¦

If you use the *modifier* argument without specifying any characters as delimiters, then the only delimiters that will be used are delimiters that are defined by the *modifier* argument. In this case, the lists of default delimiters for ASCII and EBCDIC environments are not used. In other words, modifiers add to the list of delimiters that are explicitly specified by the *charlist* argument. Modifiers do not add to the list of default modifiers.

The Length of the Result In a DATA step, most variables have a fixed length. If the word returned by the SCAN function is assigned to a variable that has a fixed length greater than the length of the returned word, then the value of that variable will be padded with blanks. Macro variables have varying lengths and are not padded with blanks.

The maximum length of the word that is returned by the SCAN function depends on the environment from which it is called:

- In a DATA step, if the SCAN function returns a value to a variable that has not yet been given a length, then that variable is given a length of 200 characters. If you need the SCAN function to assign to a variable a word that is longer than 200 characters, then you should explicitly specify the length of that variable.

If you use the SCAN function in an expression that contains operators or other functions, a word that is returned by the SCAN function can have a length of up to 32,767 characters, except in a WHERE clause. In that case, the maximum length is 200 characters.

- In the SQL procedure, or in a WHERE clause in any procedure, the maximum length of a word that is returned by the SCAN function is 200 characters.
- In the macro processor, the maximum length of a word that is returned by the SCAN function is 65,534 characters.

The minimum length of the word that is returned by the SCAN function depends on whether the M modifier is specified, as described in "Using the SCAN Function with the M Modifier" on page 1115, and "Using the SCAN Function without the M Modifier" on page 1115.

Using the SCAN Function with the M Modifier If you specify the M modifier, then the number of words in a string is defined as one plus the number of delimiters in the string. However, if you specify the Q modifier, delimiters that are inside quotation marks are ignored.

If you specify the M modifier, then the SCAN function returns a word with a length of zero if one of the following conditions is true:

- The string begins with a delimiter and you request the first word.
- The string ends with a delimiter and you request the last word.
- The string contains two consecutive delimiters and you request the word that is between the two delimiters.

Using the SCAN Function without the M Modifier If you do not specify the M modifier, then the number of words in a string is defined as the number of maximal substrings of consecutive non-delimiters. However, if you specify the Q modifier, delimiters that are inside quotation marks are ignored.

If you do not specify the M modifier, then the SCAN function does the following:

- ignores delimiters at the beginning or end of the string
- treats two or more consecutive delimiters as if they were a single delimiter

If the string contains no characters other than delimiters, or if you specify a count that is greater in absolute value than the number of words in the string, then the SCAN function returns one of the following:

- a single blank when you call the SCAN function from a DATA step
- a string with a length of zero when you call the SCAN function from the macro processor

Using Null Arguments The SCAN function allows character arguments to be null. Null arguments are treated as character strings with a length of zero. Numeric arguments cannot be null.

Examples

Example 1: Finding the First and Last Words in a String The following example scans a string for the first and last words. Note the following:

- A negative count instructs the SCAN function to scan from right to left.
- Leading and trailing delimiters are ignored because the M modifier is not used.

□ In the last observation, all characters in the string are delimiters.

```
options pageno=1 nodate ls=80 ps=64;

data firstlast;
   input String $60.;
   First_Word = scan(string, 1);
   Last_Word = scan(string, -1);
   datalines4;
Jack and Jill
& Bob & Carol & Ted & Alice &
Leonardo
! $ % & ( ) * + , - . / ;
;;;;

proc print data=firstlast;
run;
```

Output 4.83 Results of Finding the First and Last Words in a String

```
                              The SAS System                                1

                                                 First_      Last_
         Obs    String                           Word        Word

          1     Jack and Jill                    Jack        Jill
          2     & Bob & Carol & Ted & ALice &    Bob         Alice
          3     Leonardo                         Leonardo    Leonardo
          4     ! $ % & ( ) * + , - . / ;
```

Example 2: Finding All Words in a String without Using the M Modifier The following example scans a string from left to right until the word that is returned is blank. Because the M modifier is not used, the SCAN function does not return any words that have a length of zero. Because blanks are included among the default delimiters, the SCAN function returns a blank word only when the count exceeds the number of words in the string. Therefore, the loop can be stopped when SCAN returns a blank word.

```
options pageno=1 nodate ls=80 ps=64;

data all;
   length word $20;
   drop string;
   string = ' The quick brown fox jumps over the lazy dog.   ';
   do until(word=' ');
      count+1;
      word = scan(string, count);
      output;
   end;
run;

proc print data=all noobs;
run;
```

Output 4.84 Results of Finding All Words without Using the M Modifier

```
                                The SAS System                                1

                                word     count

                                The          1
                                quick        2
                                brown        3
                                fox          4
                                jumps        5
                                over         6
                                the          7
                                lazy         8
                                dog          9
                                            10
```

Example 3: Finding All Words in a String by Using the M and O Modifiers The following example shows the results of using the M modifier with a comma as a delimiter. With the M modifier, leading, trailing, and multiple consecutive delimiters cause the SCAN function to return words that have a length of zero. Therefore, you should not end the loop by testing for a blank word. Instead, you can use the COUNTW function with the same modifiers and delimiters to count the words in the string.

The O modifier is used for efficiency because the delimiters and modifiers are the same in every call to the SCAN and COUNTW functions.

```
options pageno=1 nodate ls=80 ps=64;

data comma;
   keep count word;
   length word $30;
   string = ',leading, trailing,and multiple,,delimiters,,';
   delim = ',';
   modif = 'mo';
   nwords = countw(string, delim, modif);
   do count = 1 to nwords;
      word = scan(string, count, delim, modif);
      output;
   end;
run;

proc print data=comma noobs;
run;
```

Output 4.85 Results of Finding All Words by Using the M and O Modifiers

```
                              The SAS System                                1

                          word            count

                                            1
                          leading           2
                              trailing      3
                          and multiple      4
                                            5
                          delimiters        6
                                            7
                                            8
```

Example 4: Using Comma-Separated Values, Substrings in Quotation Marks, and the O and R Modifiers The following example uses the SCAN function with the O modifier and a comma as a delimiter, both with and without the R modifier.

The O modifier is used for efficiency because in each call of the SCAN or COUNTW function, the delimiters and modifiers do not change. The O modifier applies separately to each of the two instances of the SCAN function:

- The first instance of the SCAN function uses the same delimiters and modifiers every time SCAN is called. Consequently, you can use the O modifier for this instance.
- The second instance of the SCAN function uses the same delimiters and modifiers every time SCAN is called. Consequently, you can use the O modifier for this instance.
- The first instance of the SCAN function does not use the same modifiers as the second instance, but this fact has no bearing on the use of the O modifier.

```
options pageno=1 nodate ls=80 ps=64;

data test;
   keep count word word_r;
   length word word_r $30;
   string = 'He said, "She said, ""No!""", not "Yes!"';
   delim = ',';
   modif = 'oq';
   nwords = countw(string, delim, modif);
   do count = 1 to nwords;
      word   = scan(string, count, delim, modif);
      word_r = scan(string, count, delim, modif||'r');
      output;
   end;
run;

proc print data=test noobs;
run;
```

Output 4.86 Results of Comma-Separated Values and Substrings in Quotation Marks

```
                          The SAS System                                1

word                    word_r             count

He said                 He said              1
 "She said, ""No!"""    She said, "No!"      2
 not "Yes!"             not "Yes!"           3
```

Example 5: Finding Substrings of Digits by Using the D and K Modifiers The following example finds substrings of digits. The *charlist* argument is null. Consequently, the list of characters is initially empty. The D modifier adds digits to the list of characters. The K modifier treats all characters that are not in the list as delimiters. Therefore, all characters except digits are delimiters.

```
options pageno=1 nodate ls=80 ps=64;

data digits;
   keep count digits;
   length digits $20;
   string = 'Call (800) 555--1234 now!';
   do until(digits = ' ');
      count+1;
      digits = scan(string, count, , 'dko');
      output;
   end;
run;

proc print data=digits noobs;
run;
```

Output 4.87 Results of Finding Substrings of Digits by Using the D and K Modifiers

```
                           The SAS System                               1

                         digits    count

                          800        1
                          555        2
                          1234       3
                                     4
```

See Also

Functions and CALL routines:

"CALL SCAN Routine" on page 516

"COUNTW Function" on page 621

"FINDW Function" on page 743

SDF Function

Returns a survival function.

Category: Probability
See: "CDF Function" on page 558

Syntax

SDF(*dist, quantile, parm-1,...,parm-k*)

Arguments

dist
is a character string that identifies the distribution. Valid distributions are as follows:

Distribution	Argument
Bernoulli	**BERNOULLI**
Beta	**BETA**
Binomial	**BINOMIAL**
Cauchy	**CAUCHY**
Chi-Square	**CHISQUARE**
Exponential	**EXPONENTIAL**
F	**F**
Gamma	**GAMMA**
Geometric	**GEOMETRIC**
Hypergeometric	**HYPERGEOMETRIC**
Laplace	**LAPLACE**
Logistic	**LOGISTIC**
Lognormal	**LOGNORMAL**
Negative binomial	**NEGBINOMIAL**
Normal	**NORMAL|GAUSS**
Normal mixture	**NORMALMIX**
Pareto	**PARETO**
Poisson	**POISSON**
T	**T**
Uniform	**UNIFORM**
Wald (inverse Gaussian)	**WALD|IGAUSS**
Weibull	**WEIBULL**

Note: Except for T, F, and NORMALMIX, you can minimally identify any distribution by its first four characters. △

quantile
is a numeric constant, variable or expression that specifies the value of a random variable.

parm-1,...,parm-k
are optional *shape*, *location*, or *scale* parameters appropriate for the specific distribution.

The SDF function computes the survival function (upper tail) from various continuous and discrete distributions. For more information, see the on page 559.

Examples

SAS Statements	Results
`y=sdf('BERN',0,.25);`	0.25
`y=sdf('BETA',0.2,3,4);`	0.09011
`y=sdf('BINOM',4,.5,10);`	0.62305
`y=sdf('CAUCHY',2);`	0.14758
`y=sdf('CHISQ',11.264,11);`	0.42142
`y=sdf('EXPO',1);`	0.36788
`y=sdf('F',3.32,2,3);`	0.17361
`y=sdf('GAMMA',1,3);`	0.91970
`y=sdf('HYPER',2,200,50,10);`	0.47633
`y=sdf('LAPLACE',1);`	0.18394
`y=sdf('LOGISTIC',1);`	0.26894
`y=sdf('LOGNORMAL',1);`	0.5
`y=sdf('NEGB',1,.5,2);`	0.5
`y=sdf('NORMAL',1.96);`	0.025
`y=pdf('NORMALMIX',2.3,3,.33,.33,.34, .5,1.5,2.5,.79,1.6,4.3);`	0.2819
`y=sdf('PARETO',1,1);`	1
`y=sdf('POISSON',2,1);`	0.08030
`y=sdf('T',.9,5);`	0.20469
`y=sdf('UNIFORM',0.25);`	0.75
`y=sdf('WALD',1,2);`	0.37230
`y=sdf('WEIBULL',1,2);`	0.36788

See Also

Functions:

"LOGCDF Function" on page 907
"LOGPDF Function" on page 909
"LOGSDF Function" on page 910
"PDF Function" on page 986
"CDF Function" on page 558
"QUANTILE Function" on page 1064

SECOND Function

Returns the second from a SAS time or datetime value.

Category: Date and Time

Syntax

SECOND(*time* | *datetime*)

Arguments

time
is a numeric constant, variable, or expression with a value that represents a SAS time value.

datetime
is a numeric constant, variable, or expression with a value that represents a SAS datetime value.

Details

The SECOND function produces a numeric value that represents a specific second of the minute. The result can be any number that is >= 0 and < 60.

Examples

SAS Statements	Results

`time='3:19:24't;`
`s=second(time);`
`put s;`
	`24`

`time='6:25:65't;`
`s=second(time);`
`put s;`
	`5`

`time='3:19:60't;`
`s=second(time);`
`put s;`
	`0`

See Also

Functions:

"HOUR Function" on page 807
"MINUTE Function" on page 928

SIGN Function

Returns the sign of a value.

Category: Mathematical

Syntax

SIGN(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The SIGN function returns the following values:

-1	if *argument* < 0
0	if *argument* = 0
1	if *argument* > 0.

Examples

SAS Statements	Results
`x=sign(-5);`	`-1`
`x=sign(5);`	`1`
`x=sign(0);`	`0`

SIN Function

Returns the sine.

Category: Trigonometric

Syntax

SIN(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression and is expressed in radians. If the magnitude of *argument* is so great that **`mod(argument,pi)`** is accurate to less than about three decimal places, SIN returns a missing value.

Examples

SAS Statements	Results
`x=sin(0.5);`	`0.4794255386`
`x=sin(0);`	`0`
`x=sin(3.14159/4);`	`.7071063121`

SINH Function

Returns the hyperbolic sine.

Category: Hyperbolic

Syntax

SINH(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The SINH function returns the hyperbolic sine of the argument, which is given by

$$\left(e^{argument} - e^{-argument}\right)/2$$

Examples

SAS Statements	Results
`x=sinh(0);`	`0`
`x=sinh(1);`	`1.1752011936`
`x=sinh(-1.0);`	`-1.175201194`

SKEWNESS Function

Returns the skewness of the nonmissing arguments.

Category: Descriptive Statistics

Syntax

SKEWNESS(*argument-1,argument-2,argument-3*<,*...argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

At least three non-missing arguments are required. Otherwise, the function returns a missing value. If all non-missing arguments have equal values, the skewness is mathematically undefined. The SKEWNESS function returns a missing value and sets _ERROR_ equal to 1.

The argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=skewness(0,1,1);`	`-1.732050808`
`x2=skewness(2,4,6,3,1);`	`0.5901286564`
`x3=skewness(2,0,0);`	`1.7320508076`
`x4=skewness(of x1-x3);`	`-0.953097714`

SLEEP Function

For a specified period of time, suspends the execution of a program that invokes this function.

Category: Special

See: SLEEP Function in the documentation for your operating environment.

Syntax

SLEEP(*n*<, *unit*>)

Arguments

n
is a numeric constant, variable, or expression that specifies the number of units of time for which you want to suspend execution of a program.

Range: $n \geq 0$

unit
is a numeric constant, variable, or expression that specifies the unit of time, as a power of 10, which is applied to *n*. For example, 1 corresponds to a second, and .001 to a millisecond.

Default: 1 in a Windows PC environment, .001 in other environments

Details

The SLEEP function suspends the execution of a program that invokes this function for a period of time that you specify. The program can be a DATA step, macro, IML, SCL, or anything that can invoke a function. The maximum sleep period for the SLEEP function is 46 days.

Examples

Example 1: Suspending Execution for a Specified Period of Time The following example tells SAS to delay the execution of the DATA step PAYROLL for 20 seconds:

```
data payroll;
   time_slept=sleep(20,1);
   ...more SAS statements...
run;
```

Example 2: Suspending Execution Based on a Calculation of Sleep Time The following example tells SAS to suspend the execution of the DATA step BUDGET until March 1, 2006, at 3:00 AM. SAS calculates the length of the suspension based on the target date and the date and time that the DATA step begins to execute.

```
data budget;
   sleeptime='01mar2006:03:00'dt-datetime();
   time_calc=sleep(sleeptime,1);
   ...more SAS statements...;
run;
```

See Also

CALL routine
"CALL SLEEP Routine" on page 526

SMALLEST Function

Returns the *k*th smallest nonmissing value.

Category: Descriptive Statistics

Syntax

SMALLEST (*k*, *value-1*<, *value-2* ...>)

Arguments

k
is a numeric constant, variable, or expression that specifies which value to return.

value
specifies a numeric constant, variable, or expression.

Details

If *k* is missing, less than zero, or greater than the number of values, the result is a missing value and _ERROR_ is set to 1. Otherwise, if *k* is greater than the number of non-missing values, the result is a missing value but _ERROR_ is not set to 1.

Comparisons

The SMALLEST function differs from the ORDINAL function in that the SMALLEST function ignores missing values, but the ORDINAL function counts missing values.

Examples

This example compares the values that are returned by the SMALLEST function with values that are returned by the ORDINAL function.

```
options pageno=1 nodate linesize=80 pagesize=60;

data comparison;
   label smallest_num='SMALLEST Function' ordinal_num='ORDINAL Function';
   do k = 1 to 4;
      smallest_num = smallest(k, 456, 789, .Q, 123);
      ordinal_num  = ordinal (k, 456, 789, .Q, 123);
      output;
   end;
run;
```

```
proc print data=comparison label noobs;
   var k smallest_num ordinal_num;
   title 'Results From the SMALLEST and the ORDINAL Functions';
run;
```

Output 4.88 Comparison of Values: The SMALLEST and the ORDINAL Functions

```
            Results From the SMALLEST and the ORDINAL Functions              1

                          SMALLEST     ORDINAL
                       k  Function    Function

                       1      123           Q
                       2      456         123
                       3      789         456
                       4        .         789
```

See Also

Functions:

"LARGEST Function" on page 877
"ORDINAL Function" on page 982
"PCTL Function" on page 985

SOUNDEX Function

Encodes a string to facilitate searching.

Category: Character
Restriction: SOUNDEX algorithm is English-biased.
Restriction: "I18N Level 0" on page 313

Syntax

SOUNDEX(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

Length of Returned Variable In a DATA step, if the SOUNDEX function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The SOUNDEX function encodes a character string according to an algorithm that was originally developed by Margaret K. Odell and Robert C. Russel (US Patents 1261167 (1918) and 1435663 (1922)). The algorithm is described in Knuth, *The Art of Computer Programming, Volume 3*. (See "References" on page 1255.) Note that the SOUNDEX algorithm is English-biased and is less useful for languages other than English.

The SOUNDEX function returns a copy of the *argument* that is encoded by using the following steps:

1 Retain the first letter in the *argument* and discard the following letters:

A E H I O U W Y

2 Assign the following numbers to these classes of letters:

1: B F P V

2: C G J K Q S X Z

3: D T

4: L

5: M N

6: R

3 If two or more adjacent letters have the same classification from Step 2, then discard all but the first. (Adjacent refers to the position in the word before discarding letters.)

The algorithm that is described in Knuth adds trailing zeros and truncates the result to the length of 4. You can perform these operations with other SAS functions.

Examples

SAS Statements	Results
`x=soundex('Paul');` `put x;`	`P4`

`word='amnesty';`
`x=soundex(word);`
`put x;`
	`A523`

SPEDIS Function

Determines the likelihood of two words matching, expressed as the asymmetric spelling distance between the two words.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

SPEDIS(*query*,*keyword*)

Arguments

query
identifies the word to query for the likelihood of a match. SPEDIS removes trailing blanks before comparing the value.

keyword
specifies a target word for the query. SPEDIS removes trailing blanks before comparing the value.

Details

Length of Returned Variable In a DATA step, if the SPEDIS function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics SPEDIS returns the distance between the query and a keyword, a nonnegative value that is usually less than 100 but never greater than 200 with the default costs.

SPEDIS computes an asymmetric spelling distance between two words as the normalized cost for converting the keyword to the query word by using a sequence of operations. SPEDIS(*QUERY*, *KEYWORD*) is *not* the same as SPEDIS(*KEYWORD*, *QUERY*).

Costs for each operation that is required to convert the keyword to the query are listed in the following table:

Operation	Cost	Explanation
match	0	no change
singlet	25	delete one of a double letter
doublet	50	double a letter
swap	50	reverse the order of two consecutive letters
truncate	50	delete a letter from the end
append	35	add a letter to the end
delete	50	delete a letter from the middle
insert	100	insert a letter in the middle
replace	100	replace a letter in the middle
firstdel	100	delete the first letter
firstins	200	insert a letter at the beginning
firstrep	200	replace the first letter

The distance is the sum of the costs divided by the length of the query. If this ratio is greater than one, the result is rounded down to the nearest whole number.

Comparisons

The SPEDIS function is similar to the COMPLEV and COMPGED functions, but COMPLEV and COMPGED are much faster, especially for long strings.

Examples

```
options nodate pageno=1 linesize=64;

data words;
   input Operation $ Query $ Keyword $;
   Distance = spedis(query,keyword);
   Cost = distance * length(query);
   datalines;
match       fuzzy        fuzzy
singlet     fuzy         fuzzy
doublet     fuuzzy       fuzzy
swap        fzuzy        fuzzy
truncate    fuzz         fuzzy
append      fuzzys       fuzzy
delete      fzzy         fuzzy
insert      fluzzy       fuzzy
replace     fizzy        fuzzy
firstdel    uzzy         fuzzy
firstins    pfuzzy       fuzzy
firstrep    wuzzy        fuzzy
several     floozy       fuzzy
;

proc print data = words;
run;
```

The output from the DATA step is as follows.

Output 4.89 Costs for SPEDIS Operations

```
                        The SAS System                          1

  Obs    Operation    Query     Keyword    Distance    Cost

    1    match        fuzzy      fuzzy         0          0
    2    singlet      fuzy       fuzzy         6         24
    3    doublet      fuuzzy     fuzzy         8         48
    4    swap         fzuzy      fuzzy        10         50
    5    truncate     fuzz       fuzzy        12         48
    6    append       fuzzys     fuzzy         5         30
    7    delete       fzzy       fuzzy        12         48
    8    insert       fluzzy     fuzzy        16         96
    9    replace      fizzy      fuzzy        20        100
   10    firstdel     uzzy       fuzzy        25        100
   11    firstins     pfuzzy     fuzzy        33        198
   12    firstrep     wuzzy      fuzzy        40        200
   13    several      floozy     fuzzy        50        300
```

See Also

Functions:

"COMPLEV Function" on page 601

"COMPGED Function" on page 596

SQRT Function

Returns the square root of a value.

Category: Mathematical

Syntax

SQRT(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression. *Argument* must be nonnegative.

Examples

SAS Statements	Results
`x=sqrt(36);`	`6`
`x=sqrt(25);`	`5`
`x=sqrt(4.4);`	`2.0976176963`

STD Function

Returns the standard deviation of the nonmissing arguments.

Category: Descriptive Statistics

Syntax

STD(*argument-1,argument-2*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least two nonmissing arguments are required. Otherwise, the function returns a missing value. The argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=std(2,6);`	`2.8284271247`
`x2=std(2,6,.);`	`2.8284271427`
`x3=std(2,4,6,3,1);`	`1.9235384062`
`x4=std(of x1-x3);`	`0.5224377453`

STDERR Function

Returns the standard error of the mean of the nonmissing arguments.

Category: Descriptive Statistics

Syntax

STDERR(*argument-1,argument-2*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least two nonmissing arguments are required. Otherwise, the function returns a missing value. The argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=stderr(2,6);`	`2`
`x2=stderr(2,6,.);`	`2`
`x3=stderr(2,4,6,3,1);`	`0.8602325267`
`x4=stderr(of x1-x3);`	`0.3799224911`

STFIPS Function

Converts state postal codes to FIPS state codes.

Category: State and ZIP Code

Syntax

STFIPS(*postal-code*)

Arguments

postal-code
specifies a character expression that contains the two-character standard state postal code. Characters can be mixed case. The function ignores trailing blanks, but generates an error if the expression contains leading blanks.

Details

The STFIPS function converts a two-character state postal code (or world-wide GSA geographic code for U.S. territories) to the corresponding numeric U.S. Federal Information Processing Standards (FIPS) code.

Comparisons

The STFIPS, STNAME, and STNAMEL functions take the same argument but return different values. STFIPS returns a numeric U.S. Federal Information Processing Standards (FIPS) code. STNAME returns an uppercase state name. STNAMEL returns a mixed case state name.

Examples

The examples show the differences when using STFIPS, STNAME, and STNAMEL.

SAS Statements	Results
`fips=stfips ('NC');` `put fips;`	`37`
`state=stname('NC');` `put state;`	`NORTH CAROLINA`
`state=stnamel('NC');` `put state;`	`North Carolina`

See Also

Functions:

"FIPNAME Function" on page 752
"FIPNAMEL Function" on page 753
"FIPSTATE Function" on page 754
"STNAME Function" on page 1136,
"STNAMEL Function" on page 1137

STNAME Function

Converts state postal codes to uppercase state names.

Category: State and ZIP Code

Syntax

STNAME(*postal-code*)

Arguments

postal-code
specifies a character expression that contains the two-character standard state postal code. Characters can be mixed case. The function ignores trailing blanks, but generates an error if the expression contains leading blanks.

Details

The STNAME function converts a two-character state postal code (or world-wide GSA geographic code for U.S. territories) to the corresponding state name in uppercase.

Note: For Version 6, the maximum length of the value that is returned is 200 characters. For Version 7 and beyond, the maximum length is 20 characters. △

Comparisons

The STFIPS, STNAME, and STNAMEL functions take the same argument but return different values. STFIPS returns a numeric U.S. Federal Information Processing Standards (FIPS) code. STNAME returns an uppercase state name. STNAMEL returns a mixed case state name.

Examples

SAS Statements	Results
`fips=stfips ('NC');` `put fips;`	`37`
`state=stname('NC');` `put state;`	`NORTH CAROLINA`
`state=stnamel('NC');` `put state;`	`North Carolina`

See Also

Functions:

"FIPNAME Function" on page 752

"FIPNAMEL Function" on page 753
"FIPSTATE Function" on page 754
"STFIPS Function" on page 1134
"STNAMEL Function" on page 1137

STNAMEL Function

Converts state postal codes to mixed case state names.

Category: State and ZIP Code

Syntax

STNAMEL(*postal-code*)

Arguments

postal-code
specifies a character expression that contains the two-character standard state postal code. Characters can be mixed case. The function ignores trailing blanks, but generates an error if the expression contains leading blanks.

Details

If the STNAMEL function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

The STNAMEL function converts a two-character state postal code (or world-wide GSA geographic code for U.S. territories) to the corresponding state name in mixed case.

Note: For Version 6, the maximum length of the value that is returned is 200 characters. For Version 7 and beyond, the maximum length is 20 characters. △

Comparisons

The STFIPS, STNAME, and STNAMEL functions take the same argument but return different values. STFIPS returns a numeric U.S. Federal Information Processing Standards (FIPS) code. STNAME returns an uppercase state name. STNAMEL returns a mixed case state name.

Examples

The examples show the differences when using STFIPS, STNAME, and STNAMEL.

SAS Statements	Results
`fips=stfips ('NC');` `put fips;`	`37`
`state=stname('NC');` `put state;`	`NORTH CAROLINA`
`state=stnamel('NC');` `put state;`	`North Carolina`

See Also

Functions:

"FIPNAME Function" on page 752
"FIPNAMEL Function" on page 753
"FIPSTATE Function" on page 754
"STFIPS Function" on page 1134

STRIP Function

Returns a character string with all leading and trailing blanks removed.

Category: Character

Restriction: "I18N Level 0" on page 313

Syntax

STRIP(*string*)

Arguments

string
is a character constant, variable, or expression.

Details

Length of Returned Variable In a DATA step, if the STRIP function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the argument.

The Basics The STRIP function returns the argument with all leading and trailing blanks removed. If the argument is blank, STRIP returns a string with a length of zero.

Assigning the results of STRIP to a variable does not affect the length of the receiving variable. If the value that is trimmed is shorter than the length of the receiving variable, SAS pads the value with new trailing blanks.

Note: The STRIP function is useful for concatenation because the concatenation operator does not remove leading or trailing blanks. △

Comparisons

The following list compares the STRIP function with the TRIM and TRIMN functions:

- For strings that are blank, the STRIP and TRIMN functions return a string with a length of zero, whereas the TRIM function returns a single blank.
- For strings that lack leading blanks, the STRIP and TRIMN functions return the same value.
- For strings that lack leading blanks but have at least one non-blank character, the STRIP and TRIM functions return the same value.

Note: **STRIP(*string*)** returns the same result as **TRIMN(LEFT(*string*))**, but the STRIP function runs faster. △

Examples

The following example shows the results of using the STRIP function to delete leading and trailing blanks.

```
options pageno=1 nodate ls=80 ps=60;

data lengthn;
   input string $char8.;
   original = '*' || string || '*';
   stripped = '*' || strip(string) || '*';
   datalines;
abcd
  abcd
    abcd
abcdefgh
 x y z
;

proc print data=lengthn;
run;
```

Output 4.90 Results from the STRIP Function

```
                              The SAS System                                1

                Obs     string      original     stripped

                 1      abcd        *abcd    *    *abcd*
                 2        abcd      *  abcd  *    *abcd*
                 3          abcd    *    abcd*    *abcd*
                 4      abcdefgh    *abcdefgh*    *abcdefgh*
                 5       x y z      * x y z  *    *x y z*
```

See Also

Functions:

"CAT Function" on page 543
"CATS Function" on page 550
"CATT Function" on page 552
"CATX Function" on page 554
"LEFT Function" on page 881
"TRIM Function" on page 1173
"TRIMN Function" on page 1175

SUBPAD Function

Returns a substring that has a length you specify, using blank padding if necessary.

Category: Character
Restriction: "I18N Level 1" on page 314

Syntax

SUBPAD(*string*, *position* <, *length*>)

Arguments

string
specifies a character constant, variable, or expression.

position
is a positive integer that specifies the position of the first character in the substring.

length
is a non-negative integer that specifies the length of the substring. If you do not specify *length*, the SUBPAD function returns the substring that extends from the position that you specify to the end of the string.

Details

In a DATA step, if the SUBPAD function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

If the substring that you specify extends beyond the length of the string, the result is padded with blanks.

Comparisons

The SUBPAD function is similar to the SUBSTR function except for the following differences:

- If the value of *length* in SUBPAD is zero, SUBPAD returns a zero-length string. If the value of *length* in SUBSTR is zero, SUBSTR
  - writes a note to the log stating that the third argument is invalid
  - sets _ERROR_=1
  - returns the substring that extends from the position that you specified to the end of the string.
- If the substring that you specify extends past the end of the string, SUBPAD pads the result with blanks to yield the length that you requested. If the substring that you specify extends past the end of the string, SUBSTR
  - writes a note to the log stating that the third argument is invalid
  - sets _ERROR_=1
  - returns the substring that extends from the position that you specified to the end of the string.

See Also

Function:

"SUBSTRN Function" on page 1144

SUBSTR (left of =) Function

Replaces character value contents.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent functions are KSUBSTR and KSUBSTRB in *SAS National Language Support (NLS): Reference Guide*.

Syntax

SUBSTR(*variable, position*<,*length*>)=*characters-to-replace*

Arguments

variable
specifies a character variable.

position
specifies a numeric constant, variable, or expression that is the beginning character position.

length
specifies a numeric constant, variable, or expression that is the length of the substring that will be replaced.

Restriction: *length* cannot be larger than the length of the expression that remains in *variable* after *position*.

Tip: If you omit *length*, SAS uses all of the characters on the right side of the assignment statement to replace the values of *variable*.

characters-to-replace
specifies a character constant, variable, or expression that will replace the contents of *variable*.

Tip: Enclose a literal string of characters in quotation marks.

Details

If you use an undeclared variable, it will be assigned a default length of 8 when the SUBSTR function is compiled.

When you use the SUBSTR function on the left side of an assignment statement, SAS replaces the value of *variable* with the expression on the right side. SUBSTR replaces *length* characters starting at the character that you specify in *position*.

Examples

SAS Statements	Results

`a='KIDNAP';`
`substr(a,1,3)='CAT';`
`put a;`
	`CATNAP`

`b=a;`
`substr(b,4)='TY';`
`put b;`
	`CATTY`

See Also

Function:

"SUBSTR (right of =) Function" on page 1143

SUBSTR (right of =) Function

Extracts a substring from an argument.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent functions are KSUBSTR and KSUBSTRB in *SAS National Language Support (NLS): Reference Guide*.

Syntax

<*variable*=>**SUBSTR**(*string, position*<,*length*>)

Arguments

variable
specifies a valid SAS variable name.

string
specifies a character constant, variable, or expression.

position
specifies a numeric constant, variable, or expression that is the beginning character position.

length
specifies a numeric constant, variable, or expression that is the length of the substring to extract.

Interaction: If *length* is zero, a negative value, or larger than the length of the expression that remains in *string* after *position*, SAS extracts the remainder of the expression. SAS also sets _ERROR_ to 1 and prints a note to the log indicating that the *length* argument is invalid.

Tip: If you omit *length*, SAS extracts the remainder of the expression.

Details

In a DATA step, if the SUBSTR (right of =) function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the first argument.

The SUBSTR function returns a portion of an expression that you specify in *string*. The portion begins with the character that you specify by *position*, and is the number of characters that you specify in *length*.

Examples

SAS Statements	Results
	`----+----1----+----2`

`date='06MAY98';`
`month=substr(date,3,3);`
`year=substr(date,6,2);`
`put @1 month @5 year;`
	`MAY 98`

See Also

Functions:

"SUBPAD Function" on page 1140
"SUBSTR (left of =) Function" on page 1141
"SUBSTRN Function" on page 1144

SUBSTRN Function

Returns a substring, allowing a result with a length of zero.

Category: Character

Restriction: "I18N Level 1" on page 314

Tip: KSUBSTR in *SAS National Language Support (NLS): Reference Guide* has the same functionality.

Syntax

SUBSTRN(*string*, *position* <, *length*>)

Arguments

string
specifies a character or numeric constant, variable, or expression.

If *string* is numeric, then it is converted to a character value that uses the BEST32. format. Leading and trailing blanks are removed, and no message is sent to the SAS log.

position
is an integer that specifies the position of the first character in the substring.

length
is an integer that specifies the length of the substring. If you do not specify *length*, the SUBSTRN function returns the substring that extends from the position that you specify to the end of the string.

Details

Length of Returned Variable In a DATA step, if the SUBSTRN function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the first argument.

The Basics The following information applies to the SUBSTRN function:

- The SUBSTRN function returns a string with a length of zero if either *position* or *length* has a missing value.
- If the position that you specify is non-positive, the result is truncated at the beginning, so that the first character of the result is the first character of the string. The length of the result is reduced accordingly.
- If the length that you specify extends beyond the end of the string, the result is truncated at the end, so that the last character of the result is the last character of the string.

Using the SUBSTRN Function in a Macro If you call SUBSTRN by using the %SYSFUNC macro, then the macro processor resolves the first argument (*string*) to determine whether the argument is character or numeric. If you do not want the first argument to be evaluated as a macro expression, use one of the macro-quoting functions in the first argument.

Comparisons

The following table lists comparisons between the SUBSTRN and the SUBSTR functions:

Table 4.6 Comparisons between SUBSTRN and SUBSTR

Condition	Function	Result
the value of *position* is nonpositive	SUBSTRN	returns a result beginning at the first character of the string.
the value of *position* is nonpositive	SUBSTR	□ writes a note to the log stating that the second argument is invalid. □ sets _ERROR_ =1. □ returns the substring that extends from the position that you specified to the end of the string.
the value of *length* is nonpositive	SUBSTRN	returns a result with a length of zero.

Condition	Function	Result
the value of *length* is nonpositive	SUBSTR	□ writes a note to the log stating that the third argument is invalid. □ sets _ERROR_ =1. □ returns the substring that extends from the position that you specified to the end of the string.
the substring that you specify extends past the end of the string	SUBSTRN	truncates the result.
the substring that you specify extends past the end of the string	SUBSTR	□ writes a note to the log stating that the third argument is invalid. □ sets _ERROR_=1. □ returns the substring that extends from the position that you specified to the end of the string.

Examples

Example 1: Manipulating Strings with the SUBSTRN Function The following example shows how to manipulate strings with the SUBSTRN function.

```
options pageno=1 nodate ls=80 ps=60;

data test;
   retain string "abcd";
   drop string;
   do Position = -1 to 6;
      do Length = max(-1,-position) to 7-position;
         Result = substrn(string, position, length);
         output;
      end;
   end;
   datalines;
abcd
;

proc print noobs data=test;
run;
```

Output 4.91 Output from the SUBSTRN Function

The SAS System 1

Position	Length	Result
-1	1	
-1	2	
-1	3	a
-1	4	ab
-1	5	abc
-1	6	abcd
-1	7	abcd
-1	8	abcd
0	0	
0	1	
0	2	a
0	3	ab
0	4	abc
0	5	abcd
0	6	abcd
0	7	abcd
1	-1	
1	0	
1	1	a
1	2	ab
1	3	abc
1	4	abcd
1	5	abcd
1	6	abcd
2	-1	
2	0	
2	1	b
2	2	bc
2	3	bcd
2	4	bcd
2	5	bcd
3	-1	
3	0	
3	1	c
3	2	cd
3	3	cd
3	4	cd
4	-1	
4	0	
4	1	d
4	2	d
4	3	d
5	-1	
5	0	
5	1	
5	2	
6	-1	
6	0	
6	1	

Example 2: Comparison between the SUBSTR and SUBSTRN Functions The following example compares the results of using the SUBSTR function and the SUBSTRN function when the first argument is numeric.

```
data _null_;
     substr_result = "*" || substr(1234.5678,2,6) || "*";
     put substr_result=;
     substrn_result = "*" || substrn(1234.5678,2,6)  || "*";
     put substrn_result=;
run;
```

Output 4.92 Results from the SUBSTR and SUBSTRN Functions

```
substr_result=*  1234*
substrn_result=*234.56*
```

See Also

Functions:

"SUBPAD Function" on page 1140

"SUBSTR (left of =) Function" on page 1141

"SUBSTR (right of =) Function" on page 1143

SUM Function

Returns the sum of the nonmissing arguments.

Category: Descriptive Statistics

Syntax

SUM(*argument,argument, ...*)

Arguments

argument
specifies a numeric constant, variable, or expression. If all the arguments have missing values, then one of the following occurs:

- If you use only one argument, then the value of that argument is returned.
- If you use two or more arguments, then a standard missing value (.) is returned.

Otherwise, the result is the sum of the nonmissing values. The argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=sum(4,9,3,8);`	`24`
`x2=sum(4,9,3,8,.);`	`24`

`x1=9;`
`x2=39;`
`x3=sum(of x1-x2);`
	`48`

`x1=5; x2=6; x3=4; x4=9;`
`y1=34; y2=12; y3=74; y4=39;`
`result=sum(of x1-x4, of y1-y5);`
	`183`

`x1=55;`
`x2=35;`
`x3=6;`
`x4=sum(of x1-x3, 5);`
	`101`

`x1=7;`
`x2=7;`
`x5=sum(x1-x2);`
	`0`

`y1=20;`
`y2=30;`
`x6=sum(of y:);`
	`50`

SUMABS Function

Returns the sum of the absolute values of the non-missing arguments.

Category: Descriptive Statistics

Syntax

SUMABS(*value-1* <,*value-2* ...>)

Arguments

value
specifies a numeric expression.

Details

If all arguments have missing values, then the result is a missing value. Otherwise, the result is the sum of the absolute values of the non-missing values.

Examples

Example 1: Calculating the Sum of Absolute Values The following example returns the sum of the absolute values of the non-missing arguments.

```
data _null_;
   x=sumabs(1,.,-2,0,3,.q,-4);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=10
```

Example 2: Calculating the Sum of Absolute Values When You Use a Variable List The following example uses a variable list and returns the sum of the absolute value of the non-missing arguments.

```
data _null_;
   x1 = 1;
   x2 = 3;
   x3 = 4;
   x4 = 3;
   x5 = 1;
   x = sumabs(of x1-x5);
   put x=;
run;
```

SAS writes the following output to the log:

```
x=12
```

SYMEXIST Function

Returns an indication of the existence of a macro variable.

Category: Macro
See: SYMEXIST Function in *SAS Macro Language: Reference*

Syntax

SYMEXIST (*argument*)

Argument

argument
can be one of the following items:

- the name of a macro variable within double quotation marks but without an ampersand
- the name of a DATA step character variable, specified with no quotation marks, which contains a macro variable name
- a character expression that constructs a macro variable name

Details

The SYMEXIST function searches any enclosing local symbol tables and then the global symbol table for the indicated macro variable and returns **1** if the macro variable is found or **0** if the macro variable is not found.

For more information, see the "SYMEXIST Function" in *SAS Macro Language: Reference*.

SYMGET Function

Returns the value of a macro variable during DATA step execution.

Category: Macro

Syntax

SYMGET(*argument*)

Arguments

argument
can be one of the following items:

- the name of a macro variable within double quotation marks but without an ampersand
- the name of a DATA step character variable, specified with no quotation marks, which contains a macro variable name
- a character expression that constructs a macro variable name

Details

If the SYMGET function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

The SYMGET function returns the value of a macro variable during DATA step execution. For more information, see the "SYMGET Function" in *SAS Macro Language: Reference*.

See Also

CALL routine:

"CALL SYMPUT Routine" on page 536

SAS Macro Language: Reference

SYMGLOBL Function

Returns an indication of whether a macro variable is in global scope to the DATA step during DATA step execution.

Category: Macro

See: SYMGLOBL Function in *SAS Macro Language: Reference*

Syntax

SYMGLOBL (*argument*)

Argument

argument
can be one of the following items:

- the name of a macro variable within double quotation marks but without an ampersand.
- the name of a DATA step character variable, specified with no quotation marks, which contains a macro variable name.
- a character expression that constructs a macro variable name.

Details

The SYMGLOBL function searches only the global symbol table for the indicated macro variable and returns **1** if the macro variable is found or **0** if the macro variable is not found.

SYMGLOBL is fully documented in *SAS Macro Language: Reference*.

SYMLOCAL Function

Returns an indication of whether a macro variable is in local scope to the DATA step during DATA step execution.

Category: Macro

See: SYMLOCAL Function in *SAS Macro Language: Reference*

Syntax

SYMLOCAL (*argument*)

Argument

argument
can be one of the following items:

- the name of a macro variable within double quotation marks but without an ampersand.
- the name of a DATA step character variable, specified with no quotation marks, which contains a macro variable name.
- a character expression that constructs a macro variable name.

Details

The SYMLOCAL function searches the enclosing local symbol tables for the indicated macro variable and returns `1` if the macro variable is found or `0` if the macro variable is not found.

SYMLOCAL is fully documented in *SAS Macro Language: Reference*.

SYSGET Function

Returns the value of the specified operating environment variable.

Category: Special

See: SYSGET Function in the documentation for your operating environment.

Syntax

SYSGET(*operating-environment-variable*)

Arguments

operating-environment-variable
is a character constant, variable, or expression with a value that is the name of an operating environment variable. The case of *operating-environment-variable* must agree with the case that is stored in the operating environment. Trailing blanks in the argument of SYSGET are significant. Use the TRIM function to remove them.

Operating Environment Information: The term *operating-environment-variable* used in the description of this function refers to a name that represents a numeric, character, or logical value in the operating environment. Refer to the SAS documentation for your operating environment for details. △

Details

If the SYSGET function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

If the value of the operating environment variable is truncated or the variable is not defined in the operating environment, SYSGET displays a warning message in the SAS log.

Examples

This example obtains the value of two environment variables in the UNIX environment:

```
data _null_;
   length result $200;
   input env_var $;
   result=sysget(trim(env_var));
   put env_var= result=;
```

```
   datalines;
USER
PATH
;
```

Executing this DATA step for user ABCDEF displays these lines:

```
ENV_VAR=USER RESULT=abcdef
ENV_VAR=PATH RESULT=path-for-abcdef
```

See Also

- Functions:
  - "ENVLEN Function" on page 671

SYSMSG Function

Returns error or warning message text from processing the last data set or external file function.

Category: SAS File I/O
Category: External Files

Syntax

SYSMSG()

Details

SYSMSG returns the text of error messages or warning messages that are produced when a data set or external file access function encounters an error condition. If no error message is available, the returned value is blank. The internally stored error message is reset to blank after a call to SYSMSG, so subsequent calls to SYSMSG before another error condition occurs return blank values.

Examples

This example uses SYSMSG to write to the SAS log the error message generated if FETCH cannot copy the next observation into the Data Set Data Vector. The return code is 0 only when a record is fetched successfully:

```
%let rc=%sysfunc(fetch(&dsid));
%if &rc ne 0 %then
   %put %sysfunc(sysmsg());
```

See Also

Functions:

"FETCH Function" on page 684

"SYSRC Function" on page 1158

SYSPARM Function

Returns the system parameter string.

Category: Special

Syntax

SYSPARM()

Details

If the SYSPARM function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

SYSPARM allows you to access a character string specified with the SYSPARM= system option at SAS invocation or in an OPTIONS statement.

Note: If the SYSPARM= system option is not specified, the SYSPARM function returns a string with a length of zero. △

Examples

This example shows the SYSPARM= system option and the SYSPARM function.

```
options sysparm='yes';
data a;
   If sysparm()='yes' then
      do;
      ...SAS Statements...
      end;
run;
```

See Also

System option:

SYSPARM= System Option in *SAS Macro Language: Reference*

SYSPROCESSID Function

Returns the process ID of the current process.

Category: Special

Syntax

SYSPROCESSID()

Details

The SYSPROCESSID function returns the 32–character hexadecimal ID of the current process. This ID can be passed to the SYSPROCESSNAME function to obtain the name of the current process.

Examples

Example 1: Using a DATA Step The following DATA step writes the current process id to the SAS log:

```
data _null_;
   id=sysprocessid();
   put id;
run;
```

Example 2: Using SAS Macro Language The following SAS Macro Language code writes the current process id to the SAS log:

```
%let id=%sysfunc(sysprocessid());
%put &id;
```

See Also

Function:

"SYSPROCESSNAME Function" on page 1156

SYSPROCESSNAME Function

Returns the process name that is associated with a given process ID, or returns the name of the current process.

Category: Special

Syntax

SYSPROCESSNAME(<*process_id*>)

Arguments

process_id
specifies a 32–character hexadecimal process id.

Details

The SYSPROCESSNAME function returns the process name associated with the process id you supply as an argument. You can use the value returned from the SYSPROCESSID function as the argument to SYSPROCESSNAME. If you omit the argument, then SYSPROCESSNAME returns the name of the current process.

You can also use the values stored in the automatic macro variables SYSPROCESSID and SYSSTARTID as arguments to SYSPROCESSNAME.

Examples

Example 1: Using SYSPROCESSNAME Without an Argument in a DATA Step The following DATA step writes the current process name to the SAS log:

```
data _null_;
   name=sysprocessname();
   put name;
run;
```

Example 2: Using SYSPROCESSNAME With an Argument in SAS Macro Language The following SAS Macro Language code writes the process name associated with the given process id to the SAS log:

```
%let id=&sysprocessid;
%let name=%sysfunc(sysprocessname(&id));
%put &name;
```

See Also

Function:

"SYSPROCESSID Function" on page 1155

SYSPROD Function

Determines whether a product is licensed.

Category: Special

Syntax

SYSPROD(*product-name*)

Arguments

product-name

specifies a character constant, variable, or expression with a value that is the name of a SAS product.

Requirement: *Product-name* must be the correct official name of the product or solution.

Details

The SYSPROD function returns 1 if a specific SAS software product is licensed, 0 if it is a SAS software product but not licensed for your system, and -1 if the product name is not recognized. Use SYSPROD in the DATA step, in an IML step, or in an SCL program.

If SYSPROD indicates that a product is licensed, it means that the final license expiration date has not passed. To determine the final expiration date for the product, execute the following program:

```
proc setinit noalias;
run;
```

It is possible for a SAS software product to exist on your system even though the product is no longer licensed. In this case, SAS cannot access this product. Similarly, it is possible for a product to be licensed, but not installed.

You can enter the product name in uppercase, in lowercase, or in mixed case. You can prefix the product with 'SAS/'. You can prefix SAS/ACCESS product names with 'ACC-'. To view a list of products that are available on your system, execute the following program:

```
proc setinit noalias;
run;
```

Examples

These examples determine whether a specified product is licensed.

- **x=sysprod('graph');**

  If SAS/GRAPH software is currently licensed, then SYSPROD returns a value of 1. If SAS/GRAPH software is not currently licensed, then SYSPROD returns a value of 0.

- **x=sysprod('abc');**

  SYSPROD returns a value of –1 because ABC is not a valid product name.

- **x=sysprod('base');**

  or

  x=sysprod('base sas');

  SYSPROD always returns a value of 1 because the Base product must be licensed for the SYSPROD function to run successfully.

SYSRC Function

Returns a system error number.

Category: SAS File I/O

Category: External Files

Syntax

SYSRC()

Details

SYSRC returns the error number for the last system error encountered by a call to one of the data set functions or external file functions.

Examples

This example determines the error message if FILEREF does not exist:

```
%if %sysfunc(fileref(myfile)) ne 0 %then
   %put %sysfunc(sysrc()) - %sysfunc(sysmsg());
```

See Also

Functions:

"FILEREF Function" on page 692
"SYSMSG Function" on page 1154

SYSTEM Function

Issues an operating environment command during a SAS session, and returns the system return code.

Category: Special

See: SYSTEM Function in the documentation for your operating environment.

Syntax

SYSTEM(*command*)

Arguments

command

specifies any of the following: a system command that is enclosed in quotation marks (explicit character string), an expression whose value is a system command, or the name of a character variable whose value is a system command that is executed.

Operating Environment Information: See the SAS documentation for your operating environment for information about what you can specify. The system return code is dependent on your operating environment. △

Restriction: The length of the command cannot be greater than 1024 characters, including trailing blanks.

Comparisons

The SYSTEM function is similar to the X statement, the X command, and the CALL SYSTEM routine. In most cases, the X statement, X command, or %SYSEXEC macro statement are preferable because they require less overhead. However, the SYSTEM

function can be executed conditionally, and accepts expressions as arguments. The X statement is a global statement and executes as a DATA step is being compiled, regardless of whether SAS encounters a conditional statement.

Examples

Execute the host command TIMEDATA if the macro variable SYSDAY is **Friday**.

```
data _null_;
   if "&sysday"="Friday" then do;
      rc=system("timedata");
   end;
   else rc=system("errorck");
run;
```

See Also

CALL Routine:
"CALL SYSTEM Routine" on page 538
Statement:
"X Statement" on page 1808

TAN Function

Returns the tangent.

Category: Trigonometric

Syntax

TAN(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression and is expressed in radians. If the magnitude of *argument* is so great that **mod(argument,pi)** is accurate to less than about three decimal places, TAN returns a missing value.
Restriction: cannot be an odd multiple of π /2

Examples

SAS Statements	Results
`x=tan(0.5);`	`0.5463024898`
`x=tan(0);`	`0`
`x=tan(3.14159/3);`	`1.7320472695`

TANH Function

Returns the hyperbolic tangent.

Category: Hyperbolic

Syntax

TANH(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Details

The TANH function returns the hyperbolic tangent of the argument, which is given by

$$\frac{\left(e^{argument} - e^{-argument}\right)}{\left(e^{argument} + e^{-argument}\right)}$$

Examples

SAS Statements	Results
`x=tanh(0);`	`0`
`x=tanh(0.5);`	`0.4621171573`
`x=tanh(-0.5);`	`-0.462117157`

TIME Function

Returns the current time of day as a numeric SAS time value.

Category: Date and Time

Syntax

TIME()

Examples

SAS assigns CURRENT a SAS time value corresponding to 14:32:00 if the following statements are executed exactly at 2:32 PM:

```
current=time();
put current=time.;
```

TIMEPART Function

Extracts a time value from a SAS datetime value.

Category: Date and Time

Syntax

TIMEPART(*datetime*)

Arguments

datetime
is a numeric constant, variable, or expression that represents a SAS datetime value.

Examples

SAS assigns TIME a SAS value that corresponds to 10:40:17 if the following statements are executed exactly at 10:40:17 a.m. on any date:

```
datim=datetime();
time=timepart(datim);
```

TINV Function

Returns a quantile from the *t* distribution.

Category: Quantile

Syntax

TINV(*p*,*df*<,*nc*>)

Arguments

p
is a numeric probability.

Range: $0 < p < 1$

df
is a numeric degrees of freedom parameter.

Range: $df > 0$

nc
is an optional numeric noncentrality parameter.

Details

The TINV function returns the p^{th} quantile from the Student's t distribution with degrees of freedom df and a noncentrality parameter nc. The probability that an observation from a t distribution is less than or equal to the returned quantile is p.

TINV accepts a noninteger degree of freedom parameter df. If the optional parameter nc is not specified or is 0, the quantile from the central t distribution is returned.

CAUTION:
For large values of *nc*, the algorithm can fail. In that case, a missing value is returned. △

Note: TINV is the inverse of the PROBT function. △

Examples

SAS Statements	Results
`x=tinv(.95,2);`	`2.9199855804`
`x=tinv(.95,2.5,3);`	`11.033833625`

See Also

Functions:

"QUANTILE Function" on page 1064

TNONCT Function

Returns the value of the noncentrality parameter from the Student's *t* distribution.

Category: Mathematical

Syntax

TNONCT(*x*,*df*,*prob*)

Arguments

x
is a numeric random variable.

df
is a numeric degrees of freedom parameter.
Range: $df > 0$

prob
is a probability.
Range: $0 < prob < 1$

Details

The TNONCT function returns the nonnegative noncentrality parameter from a noncentral *t* distribution whose parameters are *x*, *df*, and *nc*. A Newton-type algorithm is used to find a root *nc* of the equation

$$P_t\left(x|df,nc\right) - prob = 0$$

where

$$P_t\left(x|df,nc\right) = \frac{1}{\Gamma\left(\frac{df}{2}\right)} \int_0^{\infty} v^{\frac{df}{2}-1} e^{-v} \int_{-\infty}^{x\sqrt{\frac{2v}{df}}} e^{-\frac{(u-nc)^2}{2}} du dv$$

If the algorithm fails to converge to a fixed point, a missing value is returned.

Examples

```
data work;
   x=2;
   df=4;
   do nc=1 to 3 by .5;
      prob=probt(x,df,nc);
      ncc=tnonct(x,df,prob);
```

```
      output;
   end;
run;
proc print;
run;
```

Output 4.93 Computations of the Noncentrality Parameter from the *t* Distribution

OBS	x	df	nc	prob	ncc
1	2	4	1.0	0.76457	1.0
2	2	4	1.5	0.61893	1.5
3	2	4	2.0	0.45567	2.0
4	2	4	2.5	0.30115	2.5
5	2	4	3.0	0.17702	3.0

TODAY Function

Returns the current date as a numeric SAS date value.

Category: Date and Time
Alias: DATE

Syntax

TODAY()

Details

The TODAY function produces the current date in the form of a SAS date value, which is the number of days since January 1, 1960.

Examples

These statements illustrate a practical use of the TODAY function:

```
data _null_;
   tday=today();
   if (tday-datedue)> 15 then
      do;
         put 'As of ' tday date9. ' Account #'
             account 'is more than 15 days overdue.';
      end;
run;
```

TRANSLATE Function

Replaces specific characters in a character string.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KTRANSLATE in *SAS National Language Support (NLS): Reference Guide*.

See: TRANSLATE Function in the documentation for your operating environment.

Syntax

TRANSLATE(*source,to-1,from-1*<,*…to-n,from-n*>)

Arguments

source
specifies a character constant, variable, or expression that contains the original character string.

to
specifies the characters that you want TRANSLATE to use as substitutes.

from
specifies the characters that you want TRANSLATE to replace.

Interaction: Values of *to* and *from* correspond on a character-by-character basis; TRANSLATE changes the first character of *from* to the first character of *to*, and so on. If *to* has fewer characters than *from*, TRANSLATE changes the extra *from* characters to blanks. If *to* has more characters than *from*, TRANSLATE ignores the extra *to* characters.

Operating Environment Information: You must have pairs of *to* and *from* arguments on some operating environments. On other operating environments, a segment of the collating sequence replaces null *from* arguments. See the SAS documentation for your operating environment for more information. △

Details

In a DATA step, if the TRANSLATE function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the first argument.

The maximum number of pairs of *to* and *from* arguments that TRANSLATE accepts depends on the operating environment you use to run SAS. There is no functional difference between using several pairs of short arguments, or fewer pairs of longer arguments.

Comparisons

The TRANWRD function differs from TRANSLATE in that it scans for words (or patterns of characters) and replaces those words with a second word (or pattern of characters).

Examples

SAS Statements	Results
`x=translate('XYZW','AB','VW');` `put x;`	`XYZB`

See Also

Function:

"TRANWRD Function" on page 1169

TRANSTRN Function

Replaces or removes all occurrences of a substring in a character string.

Category: Character

Syntax

TRANSTRN(*source,target,replacement*)

Arguments

source
specifies a character constant, variable, or expression that you want to translate.

target
specifies a character constant, variable, or expression that is searched for in *source*.

Requirement: The length for *target* must be greater than zero.

replacement
specifies a character constant, variable, or expression that replaces *target*.

Details

Length of Returned Variable In a DATA step, if the TRANSTRN function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes. You can use the LENGTH statement, before calling TRANSTRN, to change the length of the value.

The Basics The TRANSTRN function replaces or removes all occurrences of a given substring within a character string. The TRANSTRN function does not remove trailing blanks in the *target* string and the *replacement* string. To remove all occurrences of *target*, specify *replacement* as TRIMN("").

Comparisons

The TRANWRD function differs from the TRANSTRN function because TRANSTRN allows the replacement string to have a length of zero. TRANWRD uses a single blank instead when the replacement string has a length of zero.

The TRANSLATE function converts every occurrence of a user-supplied character to another character. TRANSLATE can scan for more than one character in a single call. In doing this scan, however, TRANSLATE searches for every occurrence of any of the individual characters within a string. That is, if any letter (or character) in the target string is found in the source string, it is replaced with the corresponding letter (or character) in the replacement string.

The TRANSTRN function differs from TRANSLATE in that TRANSTRN scans for substrings and replaces those substrings with a second substring.

Examples

Example 1: Replacing All Occurrences of a Word These statements and these values produce these results:

```
name=transtrn(name, "Mrs.", "Ms.");
name=transtrn(name, "Miss", "Ms.");
put name;
```

Values	Results
Mrs. Joan Smith	**Ms. Joan Smith**
Miss Alice Cooper	**Ms. Alice Cooper**

Example 2: Removing Blanks from the Search String In this example, the TRANSTRN function does not replace the source string because the target string contains blanks.

```
data list;
   input salelist $;
   length target $10 replacement $3;
   target='FISH';
   replacement='NIP';
   salelist=transtrn(salelist,target,replacement);
   put salelist;
   datalines;
CATFISH
;
```

The LENGTH statement pads *target* with blanks to the length of 10, which causes the TRANSTRN function to search for the character string **'FISH '** in SALELIST. Because the search fails, this line is written to the SAS log:

```
CATFISH
```

You can use the TRIM function to exclude trailing blanks from a target or replacement variable. Use the TRIM function with *target*:

```
salelist=transtrn(salelist,trim(target),replacement);
put salelist;
```

Now, this line is written to the SAS log:

```
CATNIP
```

Example 3: Zero Length in the Third Argument of the TRANSTRN Function The following example shows the results of the TRANSTRN function when the third argument, *replacement*, has a length of zero. In the DATA step, a character constant that consists of two quotation marks represents a single blank, and not a zero-length string. In the following example, the results for *string1* are different from the results for *string2*.

```
data _null_;
   string1='*' || transtrn('abcxabc', 'abc', trimn('')) || '*';
   put string1=;
   string2='*' || transtrn('abcxabc', 'abc',  '') || '*';
  put string2=;
run;
```

SAS writes the following output to the log:

Output 4.94 Output When the Third Argument of TRANSTRN Has a Length of Zero

```
string1=*x*
string2=* x *
```

See Also

Function:

"TRANSLATE Function" on page 1166

TRANWRD Function

Replaces all occurrences of a substring in a character string.

Category: Character

Restriction: "I18N Level 2" on page 314

Syntax

TRANWRD(*source,target,replacement*)

Arguments

source
specifies a character constant, variable, or expression that you want to translate.

target
specifies a character constant, variable, or expression that is searched for in *source*.

Requirement: The length for *target* must be greater than zero.

replacement
specifies a character constant, variable, or expression that replaces *target*. When the replacement string has a length of zero, TRANWRD uses a single blank instead.

Details

Length of Returned Variable

In a DATA step, if the TRANWRD function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes. You can use the LENGTH statement, before calling TRANWRD, to change the length of the value.

The Basics

The TRANWRD function replaces all occurrences of a given substring within a character string. The TRANWRD function does not remove trailing blanks in the *target* string and the *replacement* string.

Comparisons

The TRANWRD function differs from the TRANSTRN function because TRANSTRN allows the replacement string to have a length of zero. TRANWRD uses a single blank instead when the replacement string has a length of zero.

The TRANSLATE function converts every occurrence of a user-supplied character to another character. TRANSLATE can scan for more than one character in a single call. In doing this scan, however, TRANSLATE searches for every occurrence of any of the individual characters within a string. That is, if any letter (or character) in the target string is found in the source string, it is replaced with the corresponding letter (or character) in the replacement string.

The TRANWRD function differs from TRANSLATE in that TRANWRD scans for substrings and replaces those substrings with a second substring.

Examples

Example 1: Replacing All Occurrences of a Word

These statements and these values produce these results:

```
name=tranwrd(name, "Mrs.", "Ms.");
name=tranwrd(name, "Miss", "Ms.");
put name;
```

Values	Results
Mrs. Joan Smith	**Ms. Joan Smith**
Miss Alice Cooper	**Ms. Alice Cooper**

Example 2: Removing Blanks From the Search String In this example, the TRANWRD function does not replace the source string because the target string contains blanks.

```
data list;
   input salelist $;
   length target $10 replacement $3;
   target='FISH';
   replacement='NIP';
   salelist=tranwrd(salelist,target,replacement);
   put salelist;
   datalines;
CATFISH
;
```

The LENGTH statement pads *target* with blanks to the length of 10, which causes the TRANWRD function to search for the character string **'FISH '** in SALELIST. Because the search fails, this line is written to the SAS log:

```
CATFISH
```

You can use the TRIM function to exclude trailing blanks from a target or replacement variable. Use the TRIM function with *target*:

```
salelist=tranwrd(salelist,trim(target),replacement);
put salelist;
```

Now, this line is written to the SAS log:

```
CATNIP
```

Example 3: Zero Length in the Third Argument of the TRANWRD Function The following example shows the results of the TRANWRD function when the third argument, *replacement*, has a length of zero. In this case, TRANWRD uses a single blank. In the DATA step, a character constant that consists of two consecutive quotation marks represents a single blank, and not a zero-length string. In this example, the results for *string1* and *string2* are the same:

```
data _null_;
   string1='*' || tranwrd('abcxabc', 'abc', trimn('')) || '*';
   put string1=;
   string2='*' || tranwrd('abcxabc', 'abc', '') || '*';
   put string2=;
run;
```

SAS writes the following output to the log:

Output 4.95 Output When the Third Argument of TRANWRD Has a Length of Zero

```
string1=* x *
string2=* x *
```

Removing Repeated Commas

You can use the TRANWRD function to remove repeated commas in text, and replace the repeated commas with a single comma. In the following example, the TRANWRD function is used twice: to replace three commas with one comma, and to replace the ending two commas with a period:

```
data _null_;
   mytxt='If you exercise your power to vote,,,then your opinion will be heard,,';
   newtext=tranwrd(mytxt, ',,,', ',');
   newtext2=tranwrd(newtext, ',,' , '.');
   put // mytxt= / newtext= / newtext2=;
run;
```

SAS writes the following output to the log:

Output 4.96 Output from Removing Repeated Commas

```
mytxt=If you exercise your power to vote,,,then your opinion will be heard,,
newtext=If you exercise your power to vote,then your opinion will be heard,,
newtext2=If you exercise your power to vote,then your opinion will be heard.
```

See Also

Function:

"TRANSLATE Function" on page 1166

TRIGAMMA Function

Returns the value of the trigamma function.

Category: Mathematical

Syntax

TRIGAMMA(*argument*)

Arguments

argument
specifies a numeric constant, variable, or expression.

Restriction: Nonpositive integers are invalid.

Details

The TRIGAMMA function returns the derivative of the DIGAMMA function. For *argument* > 0, the TRIGAMMA function is the second derivative of the LGAMMA function.

Examples

SAS Statements	Results
`x=trigamma(3);`	`0.3949340668`

TRIM Function

Removes trailing blanks from a character string, and returns one blank if the string is missing.

Category: Character

Restriction: "I18N Level 0" on page 313

Tip: DBCS equivalent function is KTRIM in *SAS National Language Support (NLS): Reference Guide*.

Syntax

TRIM(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

Length of Returned Variable In a DATA step, if the TRIM function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the argument.

The Basics TRIM copies a character argument, removes trailing blanks, and returns the trimmed argument as a result. If the argument is blank, TRIM returns one blank. TRIM is useful for concatenating because concatenation does not remove trailing blanks.

Assigning the results of TRIM to a variable does not affect the length of the receiving variable. If the trimmed value is shorter than the length of the receiving variable, SAS pads the value with new blanks as it assigns it to the variable.

Comparisons

The TRIM and TRIMN functions are similar. TRIM returns one blank for a blank string. TRIMN returns a string with a length of zero for a blank string.

Examples

Example 1: Removing Trailing Blanks These statements and this data line produce these results:

```
data test;
   input part1 $ 1-10 part2 $ 11-20;
   hasblank=part1||part2;
   noblank=trim(part1)||part2;
   put hasblank;
   put noblank;
   datalines;
```

Data Line	Results
	`----+----1----+----2`
`apple     sauce`	`apple     sauce`
	`applesauce`

Example 2: Concatenating a Blank Character Expression

SAS Statements	Results
`x="A"||trim(" ")||"B"; put x;`	`A B`
`x="    "; y=">"||trim(x)||"<"; put y;`	`> <`

See Also

Functions:

"COMPRESS Function" on page 604

"LEFT Function" on page 881
"RIGHT Function" on page 1097

TRIMN Function

Removes trailing blanks from character expressions, and returns a string with a length of zero if the expression is missing.

Category: Character
Restriction: "I18N Level 0" on page 313

Syntax

TRIMN(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

Length of Returned Variable In a DATA step, if the TRIMN function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the argument.

Assigning the results of TRIMN to a variable does not affect the length of the receiving variable. If the trimmed value is shorter than the length of the receiving variable, SAS pads the value with new blanks as it assigns it to the variable.

The Basics TRIMN copies a character argument, removes all trailing blanks, and returns the trimmed argument as a result. If the argument is blank, TRIMN returns a string with a length of zero. TRIMN is useful for concatenating because concatenation does not remove trailing blanks.

Comparisons

The TRIMN and TRIM functions are similar. TRIMN returns a string with a length of zero for a blank string. TRIM returns one blank for a blank string.

Examples

SAS Statements	Results
`x="A"||trimn("")||"B";` `put x;`	**AB**

`x="   ";`
`z=">"||trimn(x)||"<";`
`put z;`
	><

See Also

Functions:

"COMPRESS Function" on page 604
"LEFT Function" on page 881
"RIGHT Function" on page 1097
"TRIM Function" on page 1173

TRUNC Function

Truncates a numeric value to a specified number of bytes.

Category: Truncation

Syntax

TRUNC(*number*,*length*)

Arguments

number
specifies a numeric constant, variable, or expression.

length
specifies an integer.

Details

The TRUNC function truncates a full-length *number* (stored as a double) to a smaller number of bytes, as specified in *length* and pads the truncated bytes with 0s. The truncation and subsequent expansion duplicate the effect of storing numbers in less than full length and then reading them.

Examples

```
data test;
  length x 3;
  x=1/5;
run;
data test2;
  set test;
  if x ne 1/5 then
     put 'x ne 1/5';
  if x eq trunc(1/5,3) then
     put 'x eq trunc(1/5,3)';
run;
```

The variable X is stored with a length of 3 and, therefore, each of the above comparisons is true.

UNIFORM Function

Returns a random variate from a uniform distribution.

Category: Random Number
Alias: RANUNI
See: "RANUNI Function" on page 1090

UPCASE Function

Converts all letters in an argument to uppercase.

Category: Character
Restriction: "I18N Level 2" on page 314

Syntax

UPCASE(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

In a DATA step, if the UPCASE function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the argument.

The UPCASE function copies a character argument, converts all lowercase letters to uppercase letters, and returns the altered value as a result.

The results of the UPCASE function depend directly on the translation table that is in effect (see "TRANTAB System Option") and indirectly on the "ENCODING System Option" and the "LOCALE System Option" in *SAS National Language Support (NLS): Reference Guide*.

Examples

SAS Statements	Results
`name=upcase('John B. Smith');` `put name;`	`JOHN B. SMITH`

See Also

Functions:

"LOWCASE Function" on page 912
"PROPCASE Function" on page 1037

URLDECODE Function

Returns a string that was decoded using the URL escape syntax.

Category: Web Tools
Restriction: "I18N Level 2" on page 314

Syntax

URLDECODE(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

Length of Returned Variable in a DATA Step If the URLDECODE function returns a value to a variable that has not previously been assigned a length, then that variable is given the length of the argument.

The Basics The URL escape syntax is used to hide characters that might otherwise be significant when used in a URL.

A URL escape sequence can be one of the following:

- a plus sign, which is replaced by a blank
- a sequence of three characters beginning with a percent sign and followed by two hexadecimal characters, which is replaced by a single character that has the specified hexadecimal value.

Operating Environment Information: In operating environments that use EBCDIC, SAS performs an extra translation step after it recognizes an escape sequence. The specified character is assumed to be an ASCII encoding. SAS uses the transport-to-local translation table to convert this character to an EBCDIC character in operating environments that use EBCDIC. For more information see TRANTAB= System Option in *SAS National Language Support (NLS): Reference Guide*. △

Examples

SAS Statements	Results
`x1=urldecode ('abc+def');` `put x1;`	`abc def`
`x2=urldecode ('why%3F');` `put x2;`	`why?`
`x3=urldecode ('%41%42%43%23%31');` `put x3;`	`ABC#1`

See Also

Function:

"URLENCODE Function" on page 1179

URLENCODE Function

Returns a string that was encoded using the URL escape syntax.

Category: Web Tools

Restriction: "I18N Level 2" on page 314

Syntax

URLENCODE(*argument*)

Arguments

argument
specifies a character constant, variable, or expression.

Details

Length of Returned Variable in a DATA Step If the URLDECODE function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The URLENCODE function encodes characters that might otherwise be significant when used in a URL. This function encodes all characters except for the following:

- all alphanumeric characters
- dollar sign ($)
- hyphen (-)

- underscore (_)
- at sign (@)
- period (.)
- exclamation point (!)
- asterisk (*)
- open parenthesis (( )and close parenthesis ( ))
- comma (,).

Note: The encoded string might be longer than the original string. Ensure that you consider the additional length when you use this function. △

Examples

SAS Statements	Results
`x1=urlencode ('abc def');` `put x1;`	`abc%20def`
`x2=urlencode ('why?');` `put x2;`	`why%3F`
`x3=urlencode ('ABC#1');` `put x3;`	`ABC%231`

See Also

Function:

"URLDECODE Function" on page 1178

USS Function

Returns the uncorrected sum of squares of the nonmissing arguments.

Category: Descriptive Statistics

Syntax

USS(*argument-1*<,...*argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least one nonmissing argument is required. Otherwise, the function returns a missing value. If you have more than one argument, the argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=uss(4,2,3.5,6);`	`68.25`
`x2=uss(4,2,3.5,6,.);`	`68.25`
`x3=uss(of x1-x2);`	`9316.125`

UUIDGEN Function

Returns the short or binary form of a Universal Unique Identifier (UUID).

Category: Special

Syntax

UUIDGEN(<*max-warnings*<,*binary-result*>>)

Arguments

max-warnings
specifies an integer value that represents the maximum number of warnings that this function writes to the log.

Default: 1

binary-result
specifies an integer value that indicates whether this function should return a binary result. Nonzero indicates a binary result should be returned. Zero indicates that a character result should be returned.

Default: 0

Details

Length of Returned Variable in a DATA Step If the UUIDGEN function returns a value to a variable that has not previously been assigned a length, then that variable is given a length of 200 bytes.

The Basics The UUIDGEN function returns a UUID (a unique value) for each cell. The default result is 36 characters long and it looks like:

```
5ab6fa40--426b-4375--bb22--2d0291f43319
```

A binary result is 16 bytes long.

See Also

"Universal Unique Identifiers" in *SAS Language Reference: Concepts*

VAR Function

Returns the variance of the nonmissing arguments.

Category: Descriptive Statistics

Syntax

VAR(*argument-1,argument-2*<,*...argument-n*>)

Arguments

argument
specifies a numeric constant, variable, or expression. At least two nonmissing arguments are required. Otherwise, the function returns a missing value. The argument list can consist of a variable list, which is preceded by OF.

Examples

SAS Statements	Results
`x1=var(4,2,3.5,6);`	`2.7291666667`
`x2=var(4,6,.);`	`2`
`x3=var(of x1-x2);`	`0.2658420139`

VARFMT Function

Returns the format that is assigned to a SAS data set variable.

Category: SAS File I/O

Syntax

VARFMT(*data-set-id*,*var-num*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

var-num
specifies the number of the variable's position in the SAS data set.

Tip: This number is next to the variable in the list that is produced by the CONTENTS procedure.

Tip: The VARNUM function returns this number.

Details

If no format has been assigned to the variable, a blank string is returned.

Examples

Example 1: Using VARFMT to Obtain the Format of the Variable NAME This example obtains the format of the variable NAME in the SAS data set MYDATA.

```
%let dsid=%sysfunc(open(mydata,i));
%if &dsid %then
   %do;
      %let fmt=%sysfunc(varfmt(&dsid,
                        %sysfunc(varnum
                        (&dsid,NAME))));
      %let rc=%sysfunc(close(&dsid));
   %end;
```

Example 2: Using VARFMT to Obtain the Format of all the Numeric Variables in a Data Set This example creates a data set that contains the name and formatted content of each numeric variable in the SAS data set MYDATA.

```
data vars;
   length name $ 8 content $ 12;
   drop dsid i num rc fmt;
   dsid=open("mydata","i");
   num=attrn(dsid,"nvars");
   do while (fetch(dsid)=0);
      do i=1 to num;
         name=varname(dsid,i);
         if (vartype(dsid,i)='N') then do;
            fmt=varfmt(dsid,i);
            if fmt='' then fmt="BEST12.";
            content=putc(putn(getvarn
                        (dsid,i),fmt),"$char12.");
            output;
            end;
      end;
   end;
   rc=close(dsid);
run;
```

See Also

Functions:

"VARINFMT Function" on page 1184

"VARNUM Function" on page 1188

VARINFMT Function

Returns the informat that is assigned to a SAS data set variable.

Category: SAS File I/O

Syntax

VARINFMT(*data-set-id*,*var-num*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

var-num
specifies the number of the variable's position in the SAS data set.

Tip: This number is next to the variable in the list that is produced by the CONTENTS procedure.

Tip: The VARNUM function returns this number.

Details

If no informat has been assigned to the variable, a blank string is returned.

Examples

Example 1: Using VARINFMT to Obtain the Informat of the Variable NAME This example obtains the informat of the variable NAME in the SAS data set MYDATA.

```
%let dsid=%sysfunc(open(mydata,i));
%if &dsid %then
   %do;
      %let fmt=%sysfunc(varinfmt(&dsid,
                        %sysfunc(varnum
                           (&dsid,NAME))));
      %let rc=%sysfunc(close(&dsid));
   %end;
```

Example 2: Using VARINFMT to Obtain the Informat of all the Variables in a Data Set This example creates a data set that contains the name and informat of the variables in MYDATA.

```
data vars;
   length name $ 8 informat $ 10 ;
   drop dsid i num rc;
   dsid=open("mydata","i");
   num=attrn(dsid,"nvars");
   do i=1 to num;
      name=varname(dsid,i);
      informat=varinfmt(dsid,i);
```

```
      output;
   end;
   rc=close(dsid);
run;
```

See Also

Functions:

"OPEN Function" on page 980
"VARFMT Function" on page 1182
"VARNUM Function" on page 1188

VARLABEL Function

Returns the label that is assigned to a SAS data set variable.

Category: SAS File I/O

Syntax

VARLABEL(*data-set-id*,*var-num*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

var-num
specifies the number of the variable's position in the SAS data set.

Tip: This number is next to the variable in the list that is produced by the CONTENTS procedure.

Tip: The VARNUM function returns this number.

Details

If no label has been assigned to the variable, a blank string is returned.

Comparisons

VLABEL returns the label that is associated with the given variable.

Examples

This example obtains the label of the variable NAME in the SAS data set MYDATA.

Example Code 4.1 Obtaining the Label of the Variable NAME

```
%let dsid=%sysfunc(open(mydata,i));
%if &dsid %then
```

```
%do;
   %let fmt=%sysfunc(varlabel(&dsid,
                   %sysfunc(varnum
                           (&dsid,NAME))));
   %let rc=%sysfunc(close(&dsid));
%end;
```

See Also

Functions:

"OPEN Function" on page 980
"VARNUM Function" on page 1188

VARLEN Function

Returns the length of a SAS data set variable.

Category: SAS File I/O

Syntax

VARLEN(*data-set-id*,*var-num*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

var-num
specifies the number of the variable's position in the SAS data set.

Tip: This number is next to the variable in the list that is produced by the CONTENTS procedure.

Tip: The VARNUM function returns this number.

Comparisons

VLENGTH returns the compile-time (allocated) size of the given variable.

Examples

□ This example obtains the length of the variable ADDRESS in the SAS data set MYDATA.

```
%let dsid=%sysfunc(open(mydata,i));
%if &dsid %then
   %do;
      %let len=%sysfunc(varlen(&dsid,
                       %sysfunc(varnum
```

```
                        (&dsid,ADDRESS))));
      %let rc=%sysfunc(close(&dsid));
   %end;
```

□ This example creates a data set that contains the name, type, and length of the variables in MYDATA.

```
data vars;
   length name $ 8 type $ 1;
   drop dsid i num rc;
   dsid=open("mydata","i");
   num=attrn(dsid,"nvars");
   do i=1 to num;
      name=varname(dsid,i);
      type=vartype(dsid,i);
      length=varlen(dsid,i);
      output;
   end;
   rc=close(dsid);
run;
```

See Also

Functions:

"OPEN Function" on page 980

"VARNUM Function" on page 1188

VARNAME Function

Returns the name of a SAS data set variable.

Category: SAS File I/O

Syntax

VARNAME(*data-set-id*,*var-num*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

var-num
specifies the number of the variable's position in the SAS data set.

Tip: This number is next to the variable in the list that is produced by the CONTENTS procedure.

Tip: The VARNUM function returns this number.

Examples

This example copies the names of the first five variables in the SAS data set CITY (or all of the variables if there are fewer than five) into a macro variable.

```
%macro names;
   %let dsid=%sysfunc(open(city,i));
   %let varlist=;
   %do i=1 %to
      %sysfunc(min(5,%sysfunc(attrn
                              (&dsid,nvars))));
      %let varlist=&varlist %sysfunc(varname
                                  (&dsid,&i));
   %end;
   %put varlist=&varlist;
%mend names;
%names
```

See Also

Functions:

"OPEN Function" on page 980
"VARNUM Function" on page 1188

VARNUM Function

Returns the number of a variable's position in a SAS data set.

Category: SAS File I/O

Syntax

VARNUM(*data-set-id*,*var-name*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

var-name
specifies the variable's name.

Details

VARNUM returns the number of a variable's position in a SAS data set, or 0 if the variable is not in the SAS data set. This is the same variable number that is next to the variable in the output from PROC CONTENTS.

Examples

- This example obtains the number of a variable's position in the SAS data set CITY, given the name of the variable.

```
%let dsid=%sysfunc(open(city,i));
%let citynum=%sysfunc(varnum(&dsid,CITYNAME));
%let rc=%sysfunc(fetch(&dsid));
%let cityname=%sysfunc(getvarc
                      (&dsid,&citynum));
```

- This example creates a data set that contains the name, type, format, informat, label, length, and position of the variables in SASUSER.HOUSES.

```
data vars;
   length name $ 8 type $ 1
          format informat $ 10 label $ 40;
   drop dsid i num rc;
   dsid=open("sasuser.houses","i");
   num=attrn(dsid,"nvars");
   do i=1 to num;
      name=varname(dsid,i);
      type=vartype(dsid,i);
      format=varfmt(dsid,i);
      informat=varinfmt(dsid,i);
      label=varlabel(dsid,i);
      length=varlen(dsid,i);
      position=varnum(dsid,name);
      output;
   end;
   rc=close(dsid);
run;
```

See Also

Functions:

"OPEN Function" on page 980

"VARNAME Function" on page 1187

VARRAY Function

Returns a value that indicates whether the specified name is an array.

Category: Variable Information

Restriction: Use only with the DATA step

Syntax

VARRAY (*name*)

Arguments

name
specifies a name that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

VARRAY returns 1 if the given name is an array; it returns 0 if the given name is not an array.

Comparisons

- VARRAY returns a value that indicates whether the specified name is an array. VARRAYX returns a value that indicates whether the value of the specified expression is an array.
- VARRAY does not accept an expression as an argument. VARRAYX accepts expressions, but the value of the specified variable cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results
`array x(3) x1-x3;`	
`a=varray(x);`	
`B=varray(x1);`	
`put a=;`	`a=1`
`put B=;`	`B=0`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VARRAYX Function

Returns a value that indicates whether the value of the specified argument is an array.

Category: Variable Information

Syntax

VARRAYX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression.

Restriction: The value of the specified expression cannot denote an array reference.

Details

VARRAYX returns 1 if the value of the given argument is the name of an array; it returns 0 if the value of the given argument is not the name of an array.

Comparisons

- VARRAY returns a value that indicates whether the specified name is the name of an array. VARRAYX returns a value that indicates whether the value of the specified expression is the name of an array.
- VARRAY does not accept an expression as an argument. VARRAYX accepts expressions, but the value of the specified variable cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`array vx(4) $6 vx1 vx2 vx3 vx4`
`  ('x' 'x1' 'x2' 'x3');`
`y=varrayx(vx(1));`
`z=varrayx(vx(2));`
`put y=;`
	`y=1`

`put z=;`	`z=0`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VARTYPE Function

Returns the data type of a SAS data set variable.

Category: SAS File I/O

Syntax

VARTYPE(*data-set-id*,*var-num*)

Arguments

data-set-id
specifies the data set identifier that the OPEN function returns.

var-num
specifies the number of the variable's position in the SAS data set.

Tip: This number is next to the variable in the list that is produced by the CONTENTS procedure.

Tip: The VARNUM function returns this number.

Details

VARTYPE returns C for a character variable or N for a numeric variable.

Examples

Example 1: Using VARTYPE to Determine which Variables are Numeric This example places the names of all the numeric variables of the SAS data set MYDATA into a macro variable.

```
%let dsid=%sysfunc(open(mydata,i));
%let varlist=;
%do i=1 %to %sysfunc(attrn(&dsid,nvars));
  %if (%sysfunc(vartype(&dsid,&i)) = N) %then
     %let varlist=&varlist %sysfunc(varname
                                  (&dsid,&i));
%end;
%let rc=%sysfunc(close(&dsid));
```

Example 2: Using VARTYPE to Determine which Variables are Character This example creates a data set that contains the name and formatted contents of each character variable in the SAS data set MYDATA.

```
data vars;
   length name $ 8 content $ 20;
   drop dsid i num fmt rc;
   dsid=open("mydata","i");
   num=attrn(dsid,"nvars");
   do while (fetch(dsid)=0);
      do i=1 to num;
         name=varname(dsid,i);
         fmt=varfmt(dsid,i);
         if (vartype(dsid,i)='C') then do;
            content=getvarc(dsid,i);
            if (fmt ne '' ) then
             content=left(putc(content,fmt));
            output;
            end;
      end;
```

```
    end;
    rc=close(dsid);
run;
```

See Also

Function:
"VARNUM Function" on page 1188

VERIFY Function

Returns the position of the first character in a string that is not in any of several other strings.

Category: Character
Restriction: "I18N Level 0" on page 313
Tip: DBCS equivalent function is KVERIFY in *SAS National Language Support (NLS): Reference Guide*.

Syntax

VERIFY(*source,excerpt-1*<, …, *excerpt-n*>)

Arguments

source
specifies a character constant, variable, or expression.

excerpt
specifies a character constant, variable, or expression. If you specify more than one excerpt, separate them with a comma.

Details

The VERIFY function returns the position of the first character in *source* that is not present in any *excerpt*. If VERIFY finds every character in *source* in at least one *excerpt*, it returns a 0.

Examples

SAS Statements	Results

`data scores;`
`  input Grade : $1. @@;`
`  check='abcdf';`
`  if verify(grade,check)>0 then`
`  put @1 'INVALID ' grade=;`
`  datalines;`
`a b c b c d f a a q a b d d b`
`;`
	`INVALID Grade=q`

See Also

Functions:
"FINDC Function" on page 737

VFORMAT Function

Returns the format that is associated with the specified variable.

Category: Variable Information
Restriction: Use only with the DATA step

Syntax

VFORMAT (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.
Restriction: You cannot use an expression as an argument.

Details

If the VFORMAT function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VFORMAT returns the complete format name, which includes the width and the period (for example, $CHAR20.).

Comparisons

- VFORMAT returns the format that is associated with the specified variable. VFORMATX, however, evaluates the argument to determine the variable name. The function then returns the format that is associated with that variable name.
- VFORMAT does not accept an expression as an argument. VFORMATX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 best6.;`
`y=vformat(x(1));`
`put y=;`
	`y=BEST6.`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VFORMATD Function

Returns the decimal value of the format that is associated with the specified variable.

Category: Variable Information

Syntax

VFORMATD (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Comparisons

- VFORMATD returns the format decimal value that is associated with the specified variable. VFORMATDX, however, evaluates the argument to determine the variable name. The function then returns the format decimal value that is associated with that variable name.
- VFORMATD does not accept an expression as an argument. VFORMATDX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, and length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 comma8.2;`
`y=vformatd(x(1));`
`put y=;`
	`y=2`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VFORMATDX Function

Returns the decimal value of the format that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VFORMATDX (*expression*)

Arguments

expression
specifies a SAS character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Comparisons

- VFORMATD returns the format decimal value that is associated with the specified variable. VFORMATDX, however, evaluates the argument to determine the variable name. The function then returns the format decimal value that is associated with that variable name.
- VFORMATD does not accept an expression as an argument. VFORMATDX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 comma8.2;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vformatdx(vx(1));`
`z=vformatdx('x'||'1');`
`put y=;`
`put z=;`
	
	`y=2`
	`z=2`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VFORMATN Function

Returns the format name that is associated with the specified variable.

Category: Variable Information

Syntax

VFORMATN (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

If the VFORMATN function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VFORMATN returns only the format name, which does not include the width or the period (for example, $CHAR).

Comparisons

- VFORMATN returns the format name that is associated with the specified variable. VFORMATNX, however, evaluates the argument to determine the variable name. The function then returns the format name that is associated with that variable name.
- VFORMATN does not accept an expression as an argument. VFORMATNX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, and length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 best6.;`
`y=vformatn(x(1));`
`put y=;`
	`y=BEST`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VFORMATNX Function

Returns the format name that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VFORMATNX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VFORMATNX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VFORMATNX returns only the format name, which does not include the length or the period (for example, $CHAR).

Comparisons

- VFORMATN returns the format name that is associated with the specified variable. VFORMATNX, however, evaluates the argument to determine the variable name. The function then returns the format name that is associated with that variable name.
- VFORMATN does not accept an expression as an argument. VFORMATNX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, and type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 best6.;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vformatnx(vx(1));`
`put y=;`
	`y=BEST`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VFORMATW Function

Returns the format width that is associated with the specified variable.

Category: Variable Information

Syntax

VFORMATW (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Comparisons

- VFORMATW returns the format width that is associated with the specified variable. VFORMATWX, however, evaluates the argument to determine the variable name. The function then returns the format width that is associated with that variable name.
- VFORMATW does not accept an expression as an argument. VFORMATWX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, and length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 best6.;`
`y=vformatw(x(1));`
`put y=;`
	`y=6`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VFORMATWX Function

Returns the format width that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VFORMATWX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Comparisons

- VFORMATW returns the format width that is associated with the specified variable. VFORMATWX, however, evaluates the argument to determine the variable name. The function then returns the format width that is associated with that variable name.
- VFORMATW does not accept an expression as an argument. VFORMATWX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, and type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 best6.;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vformatwx(vx(1));`
`put y=;`
	`y=6`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VFORMATX Function

Returns the format that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VFORMATX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VFORMATX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VFORMATX returns the complete format name which includes the width and the period (for example, $CHAR20.).

Comparisons

- VFORMAT returns the format that is associated with the specified variable. VFORMATX, however, evaluates the argument to determine the variable name. The function then returns the format that is associated with that variable name.
- VFORMAT does not accept an expression as an argument. VFORMATX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, and type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`format x1 best6.;`
`format x2 20.10;`
`array vx(3) $6 vx1 vx2 vx3`
`('x1' 'x2' 'x3');`
`y=vformatx(vx(1));`
`z=vformatx(vx(2));`	

`put y=;`	`y=BEST6.`
`put z=;`	`z=F20.10`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINARRAY Function

Returns a value that indicates whether the specified variable is a member of an array.

Category: Variable Information

Restriction: Use only with the DATA step

Syntax

VINARRAY (*var*)

Arguments

var

specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

VINARRAY returns 1 if the given variable is a member of an array; it returns 0 if the given variable is not a member of an array.

Comparisons

- VINARRAY returns a value that indicates whether the specified variable is a member of an array. VINARRAYX, however, evaluates the argument to determine the variable name. The function then returns a value that indicates whether the variable name is a member of an array.
- VINARRAY does not accept an expression as an argument. VINARRAYX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results
`array x(3) x1-x3;`	
`y=vinarray(x);`	
`Z=vinarray(x1);`	
`put y=;`	`y=0`
`put Z=;`	`z=1`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINARRAYX Function

Returns a value that indicates whether the value of the specified argument is a member of an array.

Category: Variable Information

Syntax

VINARRAYX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

VINARRAYX returns 1 if the value of the given argument is a member of an array; it returns 0 if the value of the given argument is not a member of an array.

Comparisons

- VINARRAY returns a value that indicates whether the specified variable is a member of an array. VINARRAYX, however, evaluates the argument to determine the variable name. The function then returns a value that indicates whether the variable name is a member of an array.
- VINARRAY does not accept an expression as an argument. VINARRAYX accepts expressions, but the value of the specified expression cannot denote an array reference.

- Related functions return the value of other variable attributes, such as the variable name, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`array vx(4) $6 vx1 vx2 vx3 vx4`
`  ('x' 'x1' 'x2' 'x3');`
`y=vinarrayx(vx(1));`
`z=vinarrayx(vx(2));`
`put y=;`
`put z=;`
	
	`y=0`
	`z=1`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMAT Function

Returns the informat that is associated with the specified variable.

Category: Variable Information

Restriction: Use only with the DATA step

Syntax

VINFORMAT (*var*)

Arguments

var

specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

If the VINFORMAT function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VINFORMAT returns the complete informat name, which includes the width and the period (for example, $CHAR20.).

Comparisons

- VINFORMAT returns the informat that is associated with the specified variable. VINFORMATX, however, evaluates the argument to determine the variable name. The function then returns the informat that is associated with that variable name.
- VINFORMAT does not accept an expression as an argument. VINFORMATX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, and length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x $char6.;`
`input x;`
`y=vinformat(x);`
`put y=;`
	`y=$CHAR6.`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMATD Function

Returns the decimal value of the informat that is associated with the specified variable.

Category: Variable Information

Syntax

VINFORMATD (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Comparisons

- VINFORMATD returns the informat decimal value that is associated with the specified variable. VINFORMATDX, however, evaluates the argument to determine the variable name. The function then returns the informat decimal value that is associated with that variable name.
- VINFORMATD does not accept an expression as an argument. VINFORMATDX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, and length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x comma8.2;`
`input x;`
`y=vinformatd(x);`
`put y=;`
	`y=2`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMATDX Function

Returns the decimal value of the informat that is associated with the value of the specified variable.

Category: Variable Information

Syntax

VINFORMATDX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified variable cannot denote an array reference.

Comparisons

- VINFORMATD returns the informat decimal value that is associated with the specified variable. VINFORMATDX, however, evaluates the argument to determine the variable name. The function then returns the informat decimal value that is associated with that variable name.
- VINFORMATD does not accept an expression as an argument. VINFORMATDX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, and type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x1 x2 x3 comma9.3;`
`input x1 x2 x3;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vinformatdx(vx(1));`
`put y=;`
	`y=3`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMATN Function

Returns the informat name that is associated with the specified variable.

Category: Variable Information

Syntax

VINFORMATN (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

If the VINFORMATN function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VINFORMATN returns only the informat name, which does not include the width or the period (for example, $CHAR).

Comparisons

- VINFORMATN returns the informat name that is associated with the specified variable. VINFORMATNX, however, evaluates the argument to determine the variable name. The function then returns the informat name that is associated with that variable name.
- VINFORMATN does not accept an expression as an argument. VINFORMATNX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, and length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x $char6.;`
`input x;`
`y=vinformatn(x);`
`put y=;`
	`y=$CHAR`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMATNX Function

Returns the informat name that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VINFORMATNX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VINFORMATNX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VINFORMATNX returns only the informat name, which does not include the width or the period (for example, $CHAR).

Comparisons

- VINFORMATN returns the informat name that is associated with the specified variable. VINFORMATNX, however, evaluates the argument to determine the variable name. The function then returns the informat name that is associated with that variable name.
- VINFORMATN does not accept an expression as an argument. VINFORMATNX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, and type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x1 x2 x3 $char6.;`
`input x1 x2 x3;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vinformatnx(vx(1));`
`put y=;`
	`y=$CHAR`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMATW Function

Returns the informat width that is associated with the specified variable.

Category: Variable Information

Syntax

VINFORMATW (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Comparisons

- VINFORMATW returns the informat width that is associated with the specified variable. VINFORMATWX, however, evaluates the argument to determine the variable name. The function then returns the informat width that is associated with that variable name.
- VINFORMATW does not accept an expression as an argument. VINFORMATWX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, type, and length, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x $char6.;`
`input x;`
`y=vinformatw(x);`
`put y=;`
	`y=6`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMATWX Function

Returns the informat width that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VINFORMATWX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Comparisons

- VINFORMATW returns the informat width that is associated with the specified variable. VINFORMATWX, however, evaluates the argument to determine the variable name. The function then returns the informat width that is associated with that variable name.
- VINFORMATW does not accept an expression as an argument. VINFORMATWX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, and type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x1 x2 x3 $char6.;`
`input x1 x2 x3;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vinformatwx(vx(1));`
`put y=;`
	`y=6`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VINFORMATX Function

Returns the informat that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VINFORMATX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VINFORMATX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VINFORMATX returns the complete informat name, which includes the width and the period (for example, $CHAR20.).

Comparisons

- VINFORMAT returns the informat that is associated with the specified variable. VINFORMATX, however, evaluates the argument to determine the variable name. The function then returns the informat that is associated with that variable name.
- VINFORMAT does not accept an expression as an argument. VINFORMATX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, length, and type, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`informat x1 x2 x3 $char6.;`
`input x1 x2 x3;`
`array vx(3) $6 vx1 vx2 vx3`
`('x1' 'x2' 'x3');`
`y=vinformatx(vx(1));`
`put y=;`
	`y=$CHAR6.`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VLABEL Function

Returns the label that is associated with the specified variable.

Category: Variable Information

Restriction: Use only with the DATA step

Syntax

VLABEL (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

If the VLABEL function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

If there is no label, VLABEL returns the variable name.

Comparisons

- VLABEL returns the label of the specified variable or the name of the specified variable, if no label exists. VLABELX, however, evaluates the argument to determine the variable name. The function then returns the label that is associated with that variable name, or the variable name if no label exists.
- VLABEL does not accept an expression as an argument. VLABELX accepts expressions, but the value of the specified expression cannot denote an array reference.
- VLABEL has the same functionality as CALL LABEL.
- Related functions return the value of other variable attributes, such as the variable name, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`label x1='Test1';`
`y=vlabel(x(1));`
`put y=;`
	`y=Test1`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VLABELX Function

Returns the label that is associated with the value of the specified argument.

Category: Variable Information

Syntax

VLABELX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VLABELX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

If there is no label, VLABELX returns the variable name.

Comparisons

- VLABEL returns the label of the specified variable, or the name of the specified variable if no label exists. VLABELX, however, evaluates the argument to determine the variable name. The function then returns the label that is associated with that variable name, or the variable name if no label exists.
- VLABEL does not accept an expression as an argument. VLABELX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`label x1='Test1';`
`y=vlabelx(vx(1));`
`put y=;`
	`y=Test1`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VLENGTH Function

Returns the compile-time (allocated) size of the specified variable.

Category: Variable Information
Restriction: Use only with the DATA step

Syntax

VLENGTH (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Comparisons

- LENGTH examines the variable at run-time, trimming trailing blanks to determine the length. VLENGTH returns a compile-time constant value, which reflects the maximum length.
- LENGTHC returns the same value as VLENGTH, but LENGTHC can be used in any calling environment and its argument can be any expression.
- VLENGTH returns the length of the specified variable. VLENGTHX, however, evaluates the argument to determine the variable name. The function then returns the compile-time size that is associated with that variable name.
- VLENGTH does not accept an expression as an argument. VLENGTHX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results
`length x $8;`	
`x='abc';`	
`y=vlength(x);`	
`z=length(x);`	
`put y=;`	`y=8`
`put z=;`	`z=3`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VLENGTHX Function

Returns the compile-time (allocated) size for the variable that has a name that is the same as the value of the argument.

Category: Variable Information

Syntax

VLENGTHX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Comparisons

- LENGTH examines the variable at run-time, trimming trailing blanks to determine the length. VLENGTHX, however, evaluates the argument to determine the variable name. The function then returns the compile-time size that is associated with that variable name.
- LENGTHC accepts an expression as the argument, but it returns the length of the value of the expression, not the length of the variable that has a name equal to the value of the expression.
- VLENGTH returns the length of the specified variable. VLENGTHX returns the length for the value of the specified expression.
- VLENGTH does not accept an expression as an argument. VLENGTHX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`length x1 $8;`
`x1='abc';`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vlengthx(vx(1));`
`z=length(x1);`	

`put y=;`	`y=8`
`put z=;`	`z=3`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VNAME Function

Returns the name of the specified variable.

Category: Variable Information

Restriction: Use only with the DATA step

Syntax

VNAME (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

If the VNAME function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

Comparisons

- VNAME returns the name of the specified variable. VNAMEX, however, evaluates the argument to determine a variable name. If the name is a known variable name, the function returns that name. Otherwise, the function returns a blank.
- VNAME does not accept an expression as an argument. VNAMEX accepts expressions, but the value of the specified expression cannot denote an array reference.
- VNAME has the same functionality as CALL VNAME.
- Related functions return the value of other variable attributes, such as the variable label, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345.

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`y=vname(x(1));`
`put y=;`
	`y=x1`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VNAMEX Function

Validates the value of the specified argument as a variable name.

Category: Variable Information

Syntax

VNAMEX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VNAMEX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

Comparisons

- VNAME returns the name of the specified variable. VNAMEX, however, evaluates the argument to determine a variable name. If the name is a known variable name, the function returns that name. Otherwise, the function returns a blank.
- VNAME does not accept an expression as an argument. VNAMEX accepts expressions, but the value of the specified variable cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable label, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345.

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vnamex(vx(1));`
`z=vnamex('x'||'1');`
`put y=;`
`put z=;`
	
	`y=x1`
	`z=x1`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VTYPE Function

Returns the type (character or numeric) of the specified variable.

Category: Variable Information

Restriction: Use only with the DATA step

Syntax

VTYPE (*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.
Restriction: You cannot use an expression as an argument.

Details

If the VTYPE function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 1.

VTYPE returns N for numeric variables and C for character variables.

Comparisons

- VTYPE returns the type of the specified variable. VTYPEX, however, evaluates the argument to determine the variable name. The function then returns the type (character or numeric) that is associated with that variable name.
- VTYPE does not accept an expression as an argument. VTYPEX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`y=vtype(x(1));`
`put y=;`
	
	`y=N`

See Also

Functions:
"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VTYPEX Function

Returns the type (character or numeric) for the value of the specified argument.

Category: Variable Information

Syntax

VTYPEX (*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VTYPEX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 1.

VTYPEX returns N for numeric variables and C for character variables.

Comparisons

- VTYPE returns the type of the specified variable. VTYPEX, however, evaluates the argument to determine the variable name. The function then returns the type (character or numeric) that is associated with that variable name.
- VTYPE does not accept an expression as an argument. VTYPEX accepts expressions, but the value of the specified expression cannot denote an array reference.
- Related functions return the value of other variable attributes, such as the variable name, informat, and format, among others. For a list, see the "Variable Information" functions in "Functions and CALL Routines by Category" on page 345 .

Examples

SAS Statements	Results

`array x(3) x1-x3;`
`array vx(3) $6 vx1 vx2 vx3`
`  ('x1' 'x2' 'x3');`
`y=vtypex(vx(1));`
`put y=;`
	`y=N`

See Also

Functions:

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VVALUE Function

Returns the formatted value that is associated with the variable that you specify.

Category: Variable Information

Restriction: Use only with the DATA step

Syntax

VVALUE(*var*)

Arguments

var
specifies a variable that is expressed as a scalar or as an array reference.

Restriction: You cannot use an expression as an argument.

Details

If the VVALUE function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VVALUE returns a character string that contains the current value of the variable that you specify. The value is formatted using the current format that is associated with the variable.

Comparisons

- VVALUE returns the value that is associated with the variable that you specify. VVALUEX, however, evaluates the argument to determine the variable name. The function then returns the value that is associated with that variable name.
- VVALUE does not accept an expression as an argument. VVALUEX accepts expressions, but the value of the expression cannot denote an array reference.
- VVALUE and an assignment statement both return a character string that contains the current value of the variable that you specify. With VVALUE, the value is formatted using the current format that is associated with the variable. With an assignment statement, however, the value is unformatted.
- The PUT function allows you to reformat a specified variable or constant. VVALUE uses the current format that is associated with the variable.

Examples

SAS Statements	Results

`y=9999;`
`format y comma10.2;`
`v=vvalue(y);`
`put v;`
	`9,999.00`

See Also

Functions:

"VVALUEX Function" on page 1225

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

VVALUEX Function

Returns the formatted value that is associated with the argument that you specify.

Category: Variable Information

Syntax

VVALUEX(*expression*)

Arguments

expression
specifies a character constant, variable, or expression that evaluates to a variable name.

Restriction: The value of the specified expression cannot denote an array reference.

Details

If the VVALUEX function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 200.

VVALUEX returns a character string that contains the current value of the argument that you specify. The value is formatted by using the format that is currently associated with the argument.

Comparisons

- VVALUE accepts a variable as an argument and returns the value of that variable. VVALUEX, however, accepts a character expression as an argument. The function then evaluates the expression to determine the variable name and returns the value that is associated with that variable name.
- VVALUE does not accept an expression as an argument, but it does accept array references. VVALUEX accepts expressions, but the value of the expression cannot denote an array reference.
- VVALUEX and an assignment statement both return a character string that contains the current value of the variable that you specify. With VVALUEX, the value is formatted by using the current format that is associated with the variable. With an assignment statement, however, the value is unformatted.
- The PUT function allows you to reformat a specified variable or constant. VVALUEX uses the current format that is associated with the variable.

Examples

SAS Statements	Results

`date1='31mar02'd;`
`date2='date1';`
`format date1 date7.;`
`datevalue=vvaluex(date2);`
`put datevalue;`
	`31MAR02`

See Also

Functions:

"VVALUE Function" on page 1224

"Variable Information" functions in "Functions and CALL Routines by Category" on page 345

WEEK Function

Returns the week-number value.

Category: Date and Time

Syntax

WEEK(<*sas-date*>, <*'descriptor'*>)

Arguments

sas-date
specifies the SAS data value. If the *sas-date* argument is not specified, the WEEK function returns the week-number value of the current date.

descriptor
specifies the value of the descriptor. The following descriptors can be specified in uppercase or lowercase characters.

U (default)
specifies the number-of-the-week within the year. Sunday is considered the first day of the week. The number-of-the-week value is represented as a decimal number in the range 0–53. Week 53 has no special meaning. The value of `week('31dec2006'd, 'u')` is 53.

Tip: The U and W descriptors are similar, except that the U descriptor considers Sunday as the first day of the week, and the W descriptor considers Monday as the first day of the week.

See: "The U Descriptor" on page 1227

V
specifies the number-of-the-week whose value is represented as a decimal number in the range 1–53. Monday is considered the first day of the week and week 1 of the year is the week that includes both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year.

See: "The V Descriptor" on page 1227

W
specifies the number-of-the-week within the year. Monday is considered the first day of the week. The number-of-the-week value is represented as a decimal number in the range 0–53. Week 53 has no special meaning.
The value of **`week('31dec2006'd, 'w')`** is 53.

Tip: The U and W descriptors are similar except that the U descriptor considers Sunday as the first day of the week, and the W descriptor considers Monday as the first day of the week.

See: "The W Descriptor" on page 1227

Details

The Basics

The WEEK function reads a SAS date value and returns the week number. The WEEK function is not dependent on locale, and uses only the Gregorian calendar in its computations.

The U Descriptor

The WEEK function with the U descriptor reads a SAS date value and returns the number of the week within the year. The number-of-the-week value is represented as a decimal number in the range 0–53, with a leading zero and maximum value of 53. Week 0 means that the first day of the week occurs in the preceding year. The fifth week of the year is represented as 05.

Sunday is considered the first day of the week. For example, the value of **`week('01jan2007'd, 'u')`** is 0.

The V Descriptor

The WEEK function with the V descriptor reads a SAS date value and returns the week number. The number-of-the-week is represented as a decimal number in the range 01–53. The decimal number has a leading zero and a maximum value of 53. Weeks begin on a Monday, and week 1 of the year is the week that includes both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. In the following example, 01jan2006 and 30dec2005 occur in the same week. The first day (Monday) of that week is 26dec2005. Therefore, **`week('01jan2006'd, 'v')`** and **`week('30dec2005'd, 'v')`** both return a value of 52. This means that both dates occur in week 52 of the year 2005.

The W Descriptor

The WEEK function with the W descriptor reads a SAS date value and returns the number of the week within the year. The number-of-the-week value is represented as a decimal number in the range 0–53, with a leading zero and maximum value of 53. Week 0 means that the first day of the week occurs in the preceding year. The fifth week of the year would be represented as 05.

Monday is considered the first day of the week. Therefore, the value of **`week('01jan2007'd, 'w')`** is 1.

Comparisons of Descriptors

U is the default descriptor. Its range is 0-53, and the first day of the week is Sunday. The V descriptor has a range of 1-53 and the first day of the week is Monday. The W descriptor has a range of 0-53 and the first day of the week is Monday.

The following list describes the descriptors and an associated week:

- Week 0:

  U indicates the days in the current Gregorian year before week 1.

  V does not apply.

  W indicates the days in the current Gregorian year before week 1.

- Week 1:

  U begins on the first Sunday in a Gregorian year.

  V begins on the Monday between December 29 of the previous Gregorian year and January 4 of the current Gregorian year. The first ISO week can span the previous and current Gregorian years.

  W begins on the first Monday in a Gregorian year.

- End of Year Weeks:

  U specifies that the last week (52 or 53) in the year can contain less than 7 days. A Sunday to Saturday period that spans 2 consecutive Gregorian years is designated as 52 and 0 or 53 and 0.

  V specifies that the last week (52 or 53) of the ISO year contains 7 days. However, the last week of the ISO year can span the current Gregorian and next Gregorian year.

  W specifies that the last week (52 or 53) in the year can contain less than 7 days. A Monday to Sunday period that spans two consecutive Gregorian years is designated as 52 and 0 or 53 and 0.

Examples

The following example shows the values of the U, V, and W descriptors for dates near the end of certain years and the beginning of the next year. Examining the full data set illustrates how the behavior can differ between the various descriptors depending on the day of the week for January 1. The output displays the first 20 observations:

```
options pageno=1 nodate ls=80 ps=64;

title 'Values of the U, V, and W Descriptors';
data a(drop=i date0 date1 y);
     date0 = '20dec2005'd;
     do y = 0 to 5;
        date1 = intnx("YEAR",date0,y,'s');
        do i = 0 to 20;
          date = intnx("DAY",date1,i);
          year = YEAR(date);
          week   = week(date);
          week_u = week(date, 'u');
          week_v = week(date, 'v');
          week_w = week(date, 'w');
```

```
          output;
        end;
      end;
      format date WEEKDATX17.;
run;
proc print;
run;
```

Output 4.97 Results of Identifying the Values of the U, V, and W Descriptors

Values of the U, V, and W Descriptors 1

Obs	date	year	week	week_u	week_v	week_w
1	Tue, 20 Dec 2005	2005	51	51	51	51
2	Wed, 21 Dec 2005	2005	51	51	51	51
3	Thu, 22 Dec 2005	2005	51	51	51	51
4	Fri, 23 Dec 2005	2005	51	51	51	51
5	Sat, 24 Dec 2005	2005	51	51	51	51
6	Sun, 25 Dec 2005	2005	52	52	51	51
7	Mon, 26 Dec 2005	2005	52	52	52	52
8	Tue, 27 Dec 2005	2005	52	52	52	52
9	Wed, 28 Dec 2005	2005	52	52	52	52
10	Thu, 29 Dec 2005	2005	52	52	52	52
11	Fri, 30 Dec 2005	2005	52	52	52	52
12	Sat, 31 Dec 2005	2005	52	52	52	52
13	Sun, 1 Jan 2006	2006	1	1	52	0
14	Mon, 2 Jan 2006	2006	1	1	1	1
15	Tue, 3 Jan 2006	2006	1	1	1	1
16	Wed, 4 Jan 2006	2006	1	1	1	1
17	Thu, 5 Jan 2006	2006	1	1	1	1
18	Fri, 6 Jan 2006	2006	1	1	1	1
19	Sat, 7 Jan 2006	2006	1	1	1	1
20	Sun, 8 Jan 2006	2006	2	2	1	1

See Also

Functions:

"INTNX Function" on page 848

Formats:

"WEEKU*w*. Format" on page 259
"WEEKV*w*. Format" on page 261
"WEEKW*w*. Format" on page 263

Informats:

"WEEKU*w*. Informat" on page 1400
"WEEKV*w*. Informat" on page 1402
"WEEKW*w*. Informat" on page 1404

WEEKDAY Function

From a SAS date value, returns an integer that corresponds to the day of the week.

Category: Date and Time

Syntax

WEEKDAY(*date*)

Arguments

date
specifies a SAS expression that represents a SAS date value.

Details

The WEEKDAY function produces an integer that represents the day of the week, where 1=Sunday, 2=Monday, ..., 7=Saturday.

Examples

SAS Statements	Results
`x=weekday('16mar97'd);` `put x;`	`1`

WHICHC Function

Searches for a character value that is equal to the first argument, and returns the index of the first matching value.

Category: Search

Syntax

WHICHC(*string*, *value-1* <, *value-2*, ...>)

Arguments

string
is a character constant, variable, or expression that specifies the value to search for.

value
is a character constant, variable, or expression that specifies the value to be searched.

Details

The WHICHC function searches the second and subsequent arguments for a value that is equal to the first argument, and returns the index of the first matching value.

If *string* is missing, then WHICHC returns a missing value. Otherwise, WHICHC compares the value of *string* with *value-1*, *value-2*, and so on, in sequence. If argument *value-i* equals *string*, then WHICHC returns the positive integer *i*. If *string* does not equal any subsequent argument, then WHICHC returns 0.

Using WHICHC is useful when the values that are being searched are subject to frequent change. If you need to perform many searches without changing the values that are being searched, using the HASH object is much more efficient.

Examples

The following example searches the array for the first argument and returns the index of the first matching value.

```
data _null_;
   array fruit (*) $12 fruit1-fruit3 ('watermelon' 'apple' 'banana');
   x1=whichc('watermelon', of fruit[*]);
   x2=whichc('banana', of fruit[*]);
   x3=whichc('orange', of fruit[*]);
   put x1= / x2= / x3=;
run;
```

SAS writes the following output to the log:

```
x1=1
x2=3
x3=0
```

See Also

Functions:

"WHICHN Function" on page 1231

"The IN Operator in Character Comparisons" in *SAS Language Reference: Concepts*.

"Using the HASH Object" in *SAS Language Reference: Concepts*.

WHICHN Function

Searches for a numeric value that is equal to the first argument, and returns the index of the first matching value.

Category: Search

Syntax

WHICHN(*argument*, *value-1* <, *value-2*, ...>)

Arguments

argument
is a numeric constant, variable, or expression that specifies the value to search for.

value
is a numeric constant, variable, or expression that specifies the value to be searched.

Details

The WHICHN function searches the second and subsequent arguments for a value that is equal to the first argument, and returns the index of the first matching value.

If *string* is missing, then WHICHN returns a missing value. Otherwise, WHICHN compares the value of *string* with *value-1*, *value-2*, and so on, in sequence. If argument *value-i* equals *string*, then WHICHN returns the positive integer *i*. If *string* does not equal any subsequent argument, then WHICHN returns 0.

Using WHICHN is useful when the values that are being searched are subject to frequent change. If you need to perform many searches without changing the values that are being searched, using the HASH object is much more efficient.

Examples

The following example searches the array for the first argument and returns the index of the first matching value.

```
data _null_;
   array dates[*] Columbus Hastings Nicea US_Independence missing
                  Magna_Carta Gutenberg
                  (1492 1066 325 1776 . 1215 1450);
   x0=whichn(., of dates[*]);
   x1=whichn(1492, of dates[*]);
   x2=whichn(1066, of dates[*]);
   x3=whichn(1450, of dates[*]);
   x4=whichn(1000, of dates[*]);
   put x0= / x1= / x2= / x3= / x4=;
run;
```

SAS writes the following output to the log:

```
x0=.
x1=1
x2=2
x3=7
x4=0
```

See Also

Functions:

"WHICHC Function" on page 1230

"The IN Operator in Numeric Comparisons" in *SAS Language Reference: Concepts*.

"Using the Hash Object" in *SAS Language Reference: Concepts*.

YEAR Function

Returns the year from a SAS date value.

Category: Date and Time

Syntax

YEAR(*date*)

Arguments

date
specifies a SAS expression that represents a SAS date value.

Details

The YEAR function produces a four-digit numeric value that represents the year.

Examples

SAS Statements	Results

`date='25dec97'd;`
`y=year(date);`
`put y;`
	`1997`

See Also

Functions:

"DAY Function" on page 637
"MONTH Function" on page 936

YIELDP Function

Returns the yield-to-maturity for a periodic cash flow stream, such as a bond.

Category: Financial

Syntax

YIELDP(*A,c,n,K,k$_0$,p*)

Arguments

A
specifies the face value.
Range: $A > 0$

c
specifies the nominal annual coupon rate, expressed as a fraction.
Range: $0 \leq c < 1$

n
specifies the number of coupons per year.
Range: $n > 0$ and is an integer

K
specifies the number of remaining coupons from settlement date to maturity.
Range: $K > 0$ and is an integer

k$_0$
specifies the time from settlement date to the next coupon as a fraction of the annual basis.
Range: $0 < k_0 \leq \frac{1}{n}$

p
specifies the price with accrued interest.
Range: $p > 0$

Details

The YIELDP function is based on the relationship

$$P = \sum_{k=1}^{K} c\left(k\right) \frac{1}{\left(1 + \frac{y}{n}\right)^{t_k}}$$

where

$$t_k = nk_0 + k - 1$$
$$c\left(k\right) = \frac{c}{n}A \quad for \;\; k = 1, \ldots, K-1$$
$$c\left(K\right) = \left(1 + \frac{c}{n}\right) A$$

The YIELDP function solves for y.

Examples

In the following example, the YIELDP function returns the yield-to-maturity of a bond that has a face value of 1000, an annual coupon rate of 0.01, 4 coupons per year, and 14 remaining coupons. The time from settlement date to next coupon date is 0.165, and the price with accrued interest is 800.

```
data _null_;
   y=yieldp(1000,.01,4,14,.165,800);
   put y;
run;
```

The value returned is 0.0775031248.

YRDIF Function

Returns the difference in years between two dates.

Category: Date and Time

Syntax

YRDIF(*sdate*,*edate*,*basis*)

Arguments

sdate
specifies a SAS date value that identifies the starting date.

edate
specifies a SAS date value that identifies the ending date.

basis
identifies a character constant or variable that describes how SAS calculates the date difference. The following character strings are valid:

'30/360'
specifies a 30-day month and a 360-day year in calculating the number of years. Each month is considered to have 30 days, and each year 360 days, regardless of the actual number of days in each month or year.

Alias: '360'

Tip: If either date falls at the end of a month, it is treated as if it were the last day of a 30-day month.

'ACT/ACT'
uses the actual number of days between dates in calculating the number of years. SAS calculates this value as the number of days that fall in 365-day years divided by 365 plus the number of days that fall in 366-day years divided by 366.

Alias: 'Actual'

'ACT/360'
uses the actual number of days between dates in calculating the number of years. SAS calculates this value as the number of days divided by 360, regardless of the actual number of days in each year.

'ACT/365'
uses the actual number of days between dates in calculating the number of years. SAS calculates this value as the number of days divided by 365, regardless of the actual number of days in each year.

Examples

In the following example, YRDIF returns the difference in years between two dates based on each of the options for *basis*.

```
data _null_;
   sdate='16oct1998'd;
   edate='16feb2003'd;
   y30360=yrdif(sdate, edate, '30/360');
   yactact=yrdif(sdate, edate, 'ACT/ACT');
   yact360=yrdif(sdate, edate, 'ACT/360');
   yact365=yrdif(sdate, edate, 'ACT/365');
   put y30360= yactact= yact360= yact365=;
run;
```

SAS Statements	Results
`put y30360=;`	`4.333333333`
`put yactact=;`	`4.3369863014`
`put yact360=;`	`4.4`
`put yact365=;`	`4.3397260274`

See Also

Functions:

"DATDIF Function" on page 632

YYQ Function

Returns a SAS date value from year and quarter year values.

Category: Date and Time

Syntax

YYQ(*year,quarter*)

Arguments

year
specifies a two-digit or four-digit integer that represents the year. The YEARCUTOFF= system option defines the year value for two-digit dates.

quarter
specifies the quarter of the year (1, 2, 3, or 4).

Details

The YYQ function returns a SAS date value that corresponds to the first day of the specified quarter. If either *year* or *quarter* is missing, or if the quarter value is not valid, the result is missing.

Examples

SAS Statements	Results
`DateValue=yyq(2001,3);`	
`put DateValue;`	`15157`
`put DateValue date7.;`	`01JUL01`
`put DateValue date9.;`	`01JUL2001`
`StartOfQtr=yyq(99,4);`	
`put StartOfQtr;`	`14518`
`put StartOfQtr=worddate.;`	`StartOfQtr=October 1, 1999`

See Also

Functions:

"QTR Function" on page 1063
"YEAR Function" on page 1233

System Option:

"YEARCUTOFF= System Option" on page 2058

ZIPCITY Function

Returns a city name and the two-character postal code that corresponds to a ZIP code.

Category: State and ZIPCode

Syntax

ZIPCITY(*zip-code*)

Arguments

zip-code
specifies a numeric or character expression that contains a five-digit ZIP code.

Tip: If the value of *zip-code* begins with leading zeros, you can enter the value without the leading zeros. For example, if you enter 1040, ZIPCITY assumes that the value is 01040.

Details

The Basics If the ZIPCITY function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

ZIPCITY returns a city name and the two-character postal code that corresponds to its five-digit zip code argument. ZIPCITY returns the character values in mixed-case. If the zip code is unknown, ZIPCITY returns a blank value.

Note: The SASHELP.ZIPCODE data set must be present when you use this function. If you remove the data set, ZIPCITY will return unexpected results. △

How the ZIP Code Is Translated to the State Postal Code To determine which state corresponds to a particular ZIP code, this function uses a zone table that consists of the start and end ZIP code values for each state. It then finds the corresponding state for that range of ZIP codes. The zone table consists of start and end ZIP code values for each state to allow for exceptions, and does not validate ZIP code values.

With very few exceptions, a zone does not span multiple states. The exceptions are included in the zone table. It is possible for new zones or new exceptions to be added by the U.S. Postal Service at any time. However, SAS software is updated only with each new release of the product.

Determining When the State Postal Code Table Was Last Updated The SASHELP.ZIPCODE data set contains postal code information that is used with the ZIPCITY and other ZIP code functions. This data set is updated with each new release of SAS software. To determine when this table was last updated, execute PROC CONTENTS:

```
proc contents data=SASHELP.ZIPCODE;
run;
```

Then view the label information for the SASHELP.ZIPCODE data set:

```
Label                zipcodedownload.com
                     April2004, UNIQUE-updated
                     (sorted) February
                     2006, Release 9.2
```

The label shows that **zipcodedownload.com** is the site from which the ZIP codes were downloaded, and that April 2004 is the date that the ZIP codes were last refreshed. February 2006 is the last date that modifications were made to SASHELP.ZIPCODE.

Note: You can download the latest version of the SASHELP.ZIPCODE file from the SAS external Web site at any time. The file is located at **http://support.sas.com/rnd/datavisualization/mapsonline/html/misc.html**. Select **Zipcode Dataset** from the Name column to begin the download process. You must execute the CIMPORT procedure after you download and unzip the data set. △

Comparisons

The ZIPCITY, ZIPNAME, ZIPNAMEL, and ZIPSTATE functions accept the same argument but return different values:

- ZIPCITY returns the name of the city in mixed-case and the two-character postal code that corresponds to its five-digit ZIP code argument.
- ZIPNAME returns the uppercase name of the state or U.S. territory that corresponds to its five-digit ZIP code argument.
- ZIPNAMEL returns the mixed case name of the state or U.S. territory that corresponds to its five-digit ZIP code argument.
- ZIPSTATE returns the uppercase two-character state postal code (or world-wide GSA geographic code for U.S. territories) that corresponds to its five-digit ZIP code argument.

Examples

The following SAS statements produce these results.

SAS Statements	Results
`city1=zipcity(27511);` `put city1;`	`Cary, NC`

`length zip $10.;`
`zip='90049-1392';`
`zip=substr(zip,1,5);`
`city2=zipcity(zip);`
`put city2;`
	`Los Angeles, CA`

`city3=zipcity(4338);` `put city3;`	`Augusta, ME`
`city4=zipcity(01040);` `put city4;`	`Holyoke, MA`

See Also

Functions:

"ZIPNAME Function" on page 1243
"ZIPNAMEL Function" on page 1245
"ZIPSTATE Function" on page 1246
"ZIPFIPS Function" on page 1241

ZIPCITYDISTANCE Function

Returns the geodetic distance between two ZIP code locations.

Category: Distance
Category: State and Zip Code

Syntax

ZIPCITYDISTANCE(*zip-code-1*, *zip-code-2*)

Arguments

zip-code
specifies a numeric or character expression that contains the ZIP code of a location in the United States of America.

Details

The ZIPCITYDISTANCE function returns the geodetic distance in miles between two ZIP code locations. The centroid of each ZIP code is used in the calculation.

The SASHELP.ZIPCODE data set must be present when you use this function. If you remove the data set, then ZIPCITYDISTANCE will return unexpected results.

The SASHELP.ZIPCODE data set contains postal code information that is used with ZIPCITYDISTANCE and other ZIP code functions. This data set is updated with each new release of SAS software. To determine when this table was last updated, execute PROC CONTENTS:

```
proc contents data=SASHELP.ZIPCODE;
run;
```

Then view the label information for the SASHELP.ZIPCODE data set:

```
Label                zipcodedownload.com
                     April2004, UNIQUE-updated
                     (sorted) February
                     2006, Release 9.2
```

The label shows that **zipcodedownload.com** is the site from which the ZIP codes were downloaded, and that April 2004 is the date that the ZIP codes were last refreshed. February 2006 is the last date that modifications were made to SASHELP.ZIPCODE.

Note: You can download the latest version of the SASHELP.ZIPCODE file from the SAS external Web site at any time. The file is located at **http://support.sas.com/rnd/datavisualization/mapsonline/html/misc.html**. Select **Zipcode Dataset** from the Name column to begin the download process. You must execute the CIMPORT procedure after you download and unzip the data set. △

Examples

In the following example, the first ZIP code identifies a location in San Francisco, CA, and the second ZIP code identifies a location in Bangor, ME. ZIPCITYDISTANCE returns the distance in miles between these two locations.

```
data _null_;
   distance=zipcitydistance('94103', '04401');
   put 'Distance from San Francisco, CA, to Bangor, ME: ' distance 4. ' miles';
run;
```

SAS writes the following output to the log:

```
Distance from San Francisco, CA, to Bangor, ME: 2782 miles
```

See Also

Functions:

"ZIPCITY Function" on page 1238

ZIPFIPS Function

Converts ZIP codes to two-digit FIPS codes.

Category: State and Zip Code

Syntax

ZIPFIPS(*zip-code*)

Arguments

zip-code
specifies a numeric or character expression that contains a five-digit ZIP code.

Tip: If the value of *zip-code* begins with leading zeros, you can enter the value without the leading zeros. For example, if you enter 1040, ZIPFIPS assumes that the value is 01040.

Details

The Basics The ZIPFIPS function returns the two-digit numeric U.S. Federal Information Processing Standards (FIPS) code that corresponds to its five-digit ZIP code argument.

How the Zip Code Is Translated to the State Postal Code To determine which state corresponds to a particular ZIP code, this function uses a zone table that consists of the start and end ZIP code values for each state. It then finds the corresponding state for that range of ZIP codes. The zone table consists of start and end ZIP code values for each state to allow for exceptions, and does not validate ZIP code values.

With very few exceptions, a zone does not span multiple states. The exceptions are included in the zone table. It is possible for new zones or new exceptions to be added by the U.S. Postal Service at any time. However, SAS software is updated only with each new release of the product.

Examples

The following SAS statements produce these results.

SAS Statements	Results
`fips1=zipfips('27511');` `put fips1;`	`37`
`fips2=zipfips('01040');` `put fips2;`	`25`
`fips3=zipfips(1040);` `put fips3;`	`25`
`fips4=zipfips(59017);` `put fips4;`	`30`
`fips5=zipfips(24862);` `put fips5;`	`54`

See Also

Functions:

"ZIPCITY Function" on page 1238

"ZIPNAME Function" on page 1243

"ZIPNAMEL Function" on page 1245

"ZIPSTATE Function" on page 1246

ZIPNAME Function

Converts ZIP codes to uppercase state names.

Category: State and Zip Code

Syntax

ZIPNAME(*zip-code*)

Arguments

zip-code
specifies a numeric or character expression that contains a five-digit ZIP code.

Tip: If the value of *zip-code* begins with leading zeros, you can enter the value without the leading zeros. For example, if you enter 1040, ZIPNAME assumes that the value is 01040.

Details

The Basics If the ZIPNAME function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

ZIPNAME returns the name of the state or U.S. territory that corresponds to its five-digit ZIP code argument. ZIPNAME returns character values up to 20 characters long, all in uppercase.

How the Zip Code Is Translated to the State Postal Code To determine which state corresponds to a particular ZIP code, this function uses a zone table that consists of the start and end ZIP code values for each state. It then finds the corresponding state for that range of ZIP codes. The zone table consists of start and end ZIP code values for each state to allow for exceptions, and does not validate ZIP code values.

With very few exceptions, a zone does not span multiple states. The exceptions are included in the zone table. It is possible for new zones or new exceptions to be added by the U.S. Postal Service at any time. However, SAS software is updated only with each new release of the product.

Comparisons

The ZIPCITY, ZIPNAME, ZIPNAMEL, and ZIPSTATE functions accept the same argument but return different values:

- ZIPCITY returns the mixed-case name of the city and the two-character postal code that corresponds to its five-digit ZIP code argument.
- ZIPNAME returns the upper-case name of the state or U.S. territory that corresponds to its five-digit ZIP code argument.
- ZIPNAMEL returns the mixed-case name of the state or U.S. territory that corresponds to its five-digit ZIP code argument.
- ZIPSTATE returns the uppercase two-character state postal code (or world-wide GSA geographic code for U.S. territories) that corresponds to its five-digit ZIP code argument.

Examples

The following SAS statements produce these results.

SAS Statements	Results
`state1=zipname('27511');` `put state1;`	**NORTH CAROLINA**
`state2=zipname('01040');` `put state2;`	**MASSACHUSETTS**
`state3=zipname(1040);` `put state3;`	**MASSACHUSETTS**
`state4=zipname('59017');` `put state4;`	**MONTANA**

`length zip $10.;`
`zip='90049-1392';`
`zip=substr(zip,1,5);`
`state5=zipname(zip);`
`put state5;`
	CALIFORNIA

See Also

Functions:

"ZIPCITY Function" on page 1238
"ZIPFIPS Function" on page 1241
"ZIPNAMEL Function" on page 1245
"ZIPSTATE Function" on page 1246

ZIPNAMEL Function

Converts zip codes to mixed case state names.

Category: State and Zip Code

Syntax

ZIPNAMEL(*zip-code*)

Arguments

zip-code
specifies a numeric or character expression that contains a five-digit zip code.

Tip: If the value of *zip-code* begins with leading zeros, you can enter the value without the leading zeros. For example, if you enter 1040, ZIPNAMEL assumes that the value is 01040.

Details

The Basics If the ZIPNAMEL function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

ZIPNAMEL returns the name of the state or U.S. territory that corresponds to its five-digit zip code argument. ZIPNAMEL returns mixed-case character values up to 20 characters long.

How the Zip Code Is Translated to the State Postal Code To determine which state corresponds to a particular zip code, this function uses a zone table that consists of the start and end zip code values for each state. It then finds the corresponding state for that range of zip codes. The zone table consists of start and end zip code values for each state to allow for exceptions, and does not validate zip code values.

With very few exceptions, a zone does not span multiple states. The exceptions are included in the zone table. It is possible for new zones or new exceptions to be added by the U.S. Postal Service at any time. However, SAS software is updated only with each new release of the product.

Comparisons

The ZIPCITY, ZIPNAME, ZIPNAMEL, and ZIPSTATE functions accept the same argument but return different values:

- ZIPCITY returns the name of the city in mixed-case and the two-character postal code that corresponds to its five-digit zip code argument.
- ZIPNAME returns the uppercase name of the state or U.S. territory that corresponds to its five-digit zip code argument.
- ZIPNAMEL returns the mixed-case name of the state or U.S. territory that corresponds to its five-digit zip code argument.
- ZIPSTATE returns the upper-case two-character state postal code (or world-wide GSA geographic code for U.S. territories) that corresponds to its five-digit zip code argument.

Examples

The following SAS statements produce these results.

SAS Statements	Results
`state1=zipnamel('27511');` `put state1;`	`North Carolina`
`state2=zipnamel('01040');` `put state2;`	`Massachusetts`
`state3=zipnamel(1040);` `put state3;`	`Massachusetts`
`state4=zipnamel(59017);` `put state4;`	`Montana`

`length zip $10.;`
`zip='90049-1392';`
`zip=substr(zip,1,5);`
`state5=zipnamel(zip);`
`put state5;`
	`California`

See Also

Functions:

"ZIPCITY Function" on page 1238
"ZIPFIPS Function" on page 1241
"ZIPNAME Function" on page 1243
"ZIPSTATE Function" on page 1246

ZIPSTATE Function

Converts ZIP codes to two-character state postal codes.

Category: State and Zip Code

Syntax

ZIPSTATE(*zip-code*)

Arguments

zip-code
specifies a numeric or character expression that contains a valid five-digit ZIP code.

Tip: If the value of *zip-code* begins with leading zeros, you can enter the value without the leading zeros. For example, if you enter 1040, ZIPSTATE assumes that the value is 01040.

Details

The Basics If the ZIPSTATE function returns a value to a variable that has not yet been assigned a length, by default the variable is assigned a length of 20.

ZIPSTATE returns the two-character state postal code (or world-wide GSA geographic code for U.S. territories) that corresponds to its five-digit ZIP code argument. ZIPSTATE returns character values in uppercase.

Note: ZIPSTATE does not validate the ZIP code. △

How the Zip Code Is Translated to the State Postal Code To determine which state corresponds to a particular ZIP code, this function uses a zone table that consists of the start and end ZIP code values for each state. It then finds the corresponding state for that range of ZIP codes. The zone table consists of start and end ZIP code values for each state to allow for exceptions, and does not validate ZIP code values.

With very few exceptions, a zone does not span multiple states. The exceptions are included in the zone table. It is possible for new zones or new exceptions to be added by the U.S. Postal Service at any time. However, SAS software is updated only with each new release of the product.

Army Post Office (APO) and Fleet Post Office (FPO) Postal Codes The ZIPSTATE function recognizes APO and FPO ZIP codes. The APO and FPO states correspond to their exit bases in the United States.

Determining When the State Postal Code Table Was Last Updated The SASHELP.ZIPCODE data set contains postal code information that is used with the ZIPSTATE and other ZIP code functions. This data set is updated with each new release of SAS software. To determine when this table was last updated, execute PROC CONTENTS:

```
proc contents data=SASHELP.ZIPCODE;
run;
```

Then view the label information for the SASHELP.ZIPCODE data set:

```
Label                zipcodedownload.com
                     April2004, UNIQUE-updated
                     (sorted) February
                     2006, Release 9.2
```

The label shows that **zipcodedownload.com** is the site from which the ZIP codes were downloaded, and that April 2004 is the date that the ZIP codes were last refreshed. February 2006 is the last date that modifications were made to SASHELP.ZIPCODE.

Note: You can download the latest version of the SASHELP.ZIPCODE file from the SAS external Web site at any time. The file is located at **http://support.sas.com/rnd/datavisualization/mapsonline/html/misc.html**. Select **Zipcode Dataset**

from the Name column to begin the download process. You must execute the CIMPORT procedure after you download and unzip the data set. △

Comparisons

The ZIPCITY, ZIPNAME, ZIPNAMEL, and ZIPSTATE functions accept the same argument but return different values:

- ZIPCITY returns the mixed-case name of the city and the two-character postal code that corresponds to its five-digit ZIP code argument.
- ZIPNAME returns the uppercase name of the state or U.S. territory that corresponds to its five-digit ZIP code argument.
- ZIPNAMEL returns the mixed-case name of the state or U.S. territory that corresponds to its five-digit ZIP code argument.
- ZIPSTATE returns the upper-case two-character state postal code (or world-wide GSA geographic code for U.S. territories) that corresponds to its five-digit ZIP code argument.

Examples

The following SAS statements produce these results.

SAS Statements	Results
`state1=zipstate('27511');` `put state1;`	**NC**
`state2=zipstate('01040');` `put state2;`	**MA**
`state3=zipstate(1040);` `put state3;`	**MA**
`state4=zipstate(59017);` `put state4;`	**MT**

`length zip $10.;`
`zip='90049-1392';`
`zip=substr(zip,1,5);`
`state5=zipstate(zip);`
`put state5;`
	CA

See Also

Functions:

"ZIPFIPS Function" on page 1241

"ZIPNAME Function" on page 1243

"ZIPNAMEL Function" on page 1245

"ZIPCITY Function" on page 1238

Functions and CALL Routines Documented in Other SAS Publications

In addition to functions and CALL routines documented in *SAS Language Reference: Dictionary*, functions and CALL routines are also documented in the following publications:

SAS Companion for Windows

The functions and CALL routines listed here are documented only in *SAS Companion for Windows*. Other functions and CALL routines in *SAS Companion for Windows* contain information specific to the Windows operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter functions and CALL routines are not listed here.

Function or CALL routine	Description
CALL SOUND	Generates a sound with a specific frequency and duration.
MCIPISLP	Causes SAS to wait for a piece of multimedia equipment to become active.
MCIPISTR	Submits an MCI string command to a piece of multimedia equipment.
MODULE	Calls a specific routine or module that resides in an external dynamic link library (DLL).
WAKEUP	Specifies the time a SAS DATA step begins execution.

SAS Companion for OpenVMS on HP Integrity SErvers

The functions and CALL routines listed here are documented only in *SAS Companion for OpenVMS on HP Integrity Servers*. Other functions and CALL routines in *SAS Companion for OpenVMS on HP Integrity Servers* contain information specific to the OpenVMS operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter functions and CALL routines are not listed here.

Function or CALL Routine	Description
ASCEBC	Converts an input character string from ASCII to EBCDIC.
CALL FINDEND	Releases resources that are associated with a directory search.
DELETE	Deletes a file.
EBCASC	Converts an input character string from EBCDIC to ASCII.
FILEATTR	Returns the attribute information for a specified file.
FINDFILE	Searches a directory for a file.
GETDVI	Returns a specified item of information from a device.
GETJPI	Retrieves job-process information.
GETLOG	Returns information about a DCL logical name.
GETMSG	Translates an OpenVMS error code into text.
GETQUOTA	Retrieves disk quota information.
GETSYM	Returns the value of a DCL symbol.
GETTERM	Returns the characteristics of your terminal device.
MODULE	Calls a specific routine or module that resides in a shareable image.
NODENAME	Returns the name of the current node.
PUTLOG	Creates an OpenVMS logical-name in your process-level logical name table.
PUTSYM	Creates a DCL symbol in the parent SAS process.
SETTERM	Modifies a characteristic of your terminal device.
TERMIN	Allows simple input from SYS$INPUT.
TERMOUT	Allows simple output to SYS$OUTPUT.
TTCLOSE	Closes a channel that was previously assigned by TTOPEN.
TTCONTRL	Modifies the characteristics of a channel that was previously assigned by TTOPEN.
TTOPEN	Assigns an I/O channel to a terminal.
TTREAD	Reads characters from the channel assigned by TTOPEN.
TTWRITE	Writes characters to the channel assigned by TTOPEN.
VMS	Spawns a subprocess and executes a DCL command.

SAS Companion for z/OS

The functions and CALL routines listed here are documented only in *SAS Companion for z/OS*. Other functions and CALL routines in *SAS Companion for z/OS* contain information specific to the z/OS operating environment, where the main documentation is in *SAS Language Reference: Dictionary* . These latter functions and CALL routines are not listed here.

Function or CALL routine	Description
CALL TSO	Issues a TSO command or invokes a CLIST or a REXX exec during a SAS session.
CALL WTO	Sends a message to the system console.
TSO	Issues a TSO command or invokes a CLIST or a REXX exec during a SAS session.
WTO	Sends a message to the system console.

SAS Data Quality Server: Reference

Function	**Description**
DQCASE	Returns a character value with standardized capitalization.
DQGENDER	Returns a gender determination from the name of an individual.
DQGENDERINFOGET	Returns the name of the parse definition that is associated with the specified gender definition.
DQGENDERPARSED	Returns a gender determination from the parsed name of an individual.
DQIDENTIFY	Returns a category name from a character value.
DQLOCALEGUESS	Returns the name of the locale that is most likely represented by a character value.
DQLOCALEINFOGET	Returns information about locales.
DQLOCALEINFOLIST	Displays the names of the definitions in a locale and returns a count of those definitions.
DQMATCH	Returns a match code from a character value.
DQMATCHINFOGET	Returns the name of the parse definition that is associated with a match definition.
DQMATCHPARSED	Returns a match code from a parsed character value.
DQPARSE	Returns a parsed character value.
DQPARSEINFOGET	Returns the token names in a parse definition.
DQPARSETOKENGET	Returns a token from a parsed character value.
DQPARSETOKENPUT	Inserts a token into a parsed character value and returns the updated parsed character value.
DQPATTERN	Returns a pattern analysis from an input character value.

Function	Description
DQSCHEMEAPPLY CALL ROUTINE	Applies a scheme and returns a transformed value and a transformation flag.
DQSCHEMEAPPLY	Applies a scheme and returns a transformed value.
DQSRVARCHJOB	Runs an dfPower Architect job on a DataFlux Integration Server and returns a job identifier.
DQSRVCOPYLOG	Copies a job's log from a DataFlux Integration Server.
DQSRVDELETELOG	Deletes a job's log file from a DataFlux Integration Server.
DQSRVJOBSTATUS	Returns the status of a job that was submitted to a DataFlux Integration Server.
DQSRVKILLJOB	Terminates a job that is running on a DataFlux Integration Server.
DQSRVPROFJOBFILE	Runs a file-type Profile job on a DataFlux Integration Server and returns a job identifier.
DQSRVPROFJOBREP	Runs a repository-type Profile job on a DataFlux Integration Server and returns a job identifier.
DQSRVUSER	Authenticates a user on a DataFlux Integration Server.
DQSTANDARDIZE	Returns a character value after standardizing casing, spacing, and format, and applies a common representation to certain words and abbreviations.
DQTOKEN	Returns a token from a character table.

SAS Logging Facility: Configuration and Programming Reference

Function or CALL Routine	Description
LOG4SAS_APPENDER	Creates a fileref appender that can be referenced by a logger.
LOG4SAS_LOGEVENT	Logs a message using a specific logger.
LOG4SAS_LOGGER	Creates a logger.

SAS Macro Language: Reference

SAS Macro	Description
%BQUOTE, %NRBQUOTE	Masks special characters and mnemonic operators in a resolved value at macro execution.
%EVAL	Evaluates arithmetic and logical expressions using integer arithmetic.
%INDEX	Returns the position of the first character of a string.
%LENGTH	Returns the length of a string.
%QUOTE, %NRQUOTE	Masks special characters and mnemonic operators in a resolved value at macro execution.
%SCAN, %QSCAN	Searches for a word that is specified by its position in a string.
%STR, %NRSTR	Masks special characters and mnemonic operators in constant text at macro compilation.
%SUBSTR, %QSUBSTR	Produces a substring of a character string.
%SUPERQ	Masks all special characters and mnemonic operators at macro execution but prevents further resolution of the value.
%SYMEXIST	Returns an indication of the existence of a macro variable.
%SYMGLOBL	Returns an indication as to whether a macro variable is global in scope.
%SYMLOCAL	Returns an indication as to whether a macro variable is local in scope,
%SYSEVALF	Evaluates arithmetic and logical expressions using floating-point arithmetic.
%SYSFUNC, %QSYSFUNC	Executes SAS functions or user-written functions.
%SYSGET	Returns the value of the specified operating environment variable.
%SYSPROD	Reports whether a SAS software product is licenses at the site.
%UNQUOTE	During macro execution, unmasks all special characters and mnemonic operators for a value.
%UPCASE, %QUPCASE	Converts values to uppercase.

SAS National Language Support (NLS): Reference Guide

Function or CALL Routine	Description
EUROCURR	Converts one European currency to another.
GETPXLANGUAGE	Returns the current two letter language code.
GETPXLOCALE	Returns the POSIX locale value for a SAS locale.
GETPXREGION	Returns the current two letter region code.
KCOMPARE	Returns the result of a comparison of character expressions.
KCOMPRESS	Removes specified characters from a character expression.
KCOUNT	Returns the number of double-byte characters in an expression.
KCVT	Converts data from one type of encoding data to another encoding data.
KINDEX	Searches a character expression for a string of characters.
KINDEXC	Searches a character expression for specified characters.
KLEFT	Left-aligns a character expression by removing unnecessary leading DBCS blanks and SO/SI.
KLENGTH	Returns the length of an argument.
KLOWCASE	Converts all letters in an argument to lowercase.
KREVERSE	Reverses a character expression.
KRIGHT	Right-aligns a character expression by trimming trailing DBCS blanks and SO/SI.
KSCAN	Selects a specified word from a character expression.
KSTRCAT	Concatenates two or more character expressions.
KSUBSTR	Extracts a substring from an argument.
KSUBSTRB	Extracts a substring from an argument according to the byte position of the substring in the argument.
KTRANSLATE	Replaces specific characters in a character expression.
KTRIM	Removes trailing DBCS blanks and SO/SI from character expressions.
KTRUNCATE	Truncates a numeric value to a specified length.
KUPCASE	Converts all single-byte letters in an argument to uppercase.
KUPDATE	Inserts, deletes, and replaces character value contents.
KUPDATEB	Inserts, deletes, and replaces the contents of the character value according to the byte position of the character value in the argument.
KVERIFY	Returns the position of the first character that is unique to an expression.
NLDATE	Converts the SAS date value to the date value of the specified locale by using the date format descriptors.

Function or CALL Routine	Description
NLDATM	Converts the SAS datetime value to the time value of the specified locale by using the datetime- format descriptors.
NLTIME	Converts the SAS time or the datetime value to the time value of the specified locale by using the NLTIME descriptors.
SORTKEY	Creates a linguistic sort key.
TRANTAB	Transcodes data by using the specified translation table.
VARTRANSCODE	Returns the transcode attribute of a SAS data set variable.
VTRANSCODE	Returns a value that indicates whether transcoding is enabled for the specified character variable.
VTRANSCODEX	Returns a value that indicates whether transcoding is enabled for the specified argument.
UNICODELEN	Specifies the length of the character unit for the Unicode data.
UNICODEWIDTH	Specifies the length of a display unit for the Unicode data.

References

Abramowitz, M. and Stegun, I. (1964), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables — National Bureau of Standards Applied Mathematics Series #55*, Washington, DC: U.S. Government Printing Office.

Amos, D.E., Daniel, S.L., and Weston, K. (1977), "CDC 6600 Subroutines IBESS and JBESS for Bessel Functions *I(v,x)* and *J(v,x)*, $x \geq 0$, $v \geq 0$," *ACM Transactions on Mathematical Software*, 3, 76–95.

Aho, A.V., Hopcroft, J.E., and Ullman, J.D., (1974), The Design and Analysis of Computer Algorithms, Reading, MA: Addison-Wesley Publishing Co.

Cheng, R.C.H. (1977), "The Generation of Gamma Variables," *Applied Statistics*, 26, 71–75.

Duncan, D.B. (1955), "Multiple Range and Multiple F Tests," *Biometrics*, 11, 1–42.

Dunnett, C.W. (1955), "A Multiple Comparisons Procedure for Comparing Several Treatments with a Control," *Journal of the American Statistical Association*, 50, 1096–1121.

Fishman, G.S. (1976), "Sampling from the Poisson Distribution on a Computer," *Computing*, 17, 145–156.

Fishman, G.S. (1978), *Principles of Discrete Event Simulation*, New York: John Wiley & Sons, Inc.

Fishman, G.S. and Moore, L.R. (1982), "A Statistical Evaluation of Multiplicative Congruential Generators with Modulus $(2^{31} - 1)$," *Journal of the American Statistical Association*, 77, 1 29–136.

Knuth, D.E. (1973), *The Art of Computer Programming, Volume 3. Sorting and Searching*, Reading, MA: Addison-Wesley.

Hochberg, Y. and Tamhane, A.C. (1987), *Multiple Comparison Procedures*, New York: John Wiley & Sons, Inc.

Williams, D.A. (1971), "A Test for Differences Between Treatment Means when Several Dose Levels are Compared with a Zero Dose Control," *Biometrics*, 27, 103–117.

Williams, D.A. (1972), "The Comparison of Several Dose Levels with a Zero Dose Control," *Biometrics*, 28, 519–531.

CHAPTER

5

Informats

Definition of Informats

An *informat* is an instruction that SAS uses to read data values into a variable. For example, the following value contains a dollar sign and commas:

```
$1,000,000
```

To remove the dollar sign ($) and commas (,) before storing the numeric value 1000000 in a variable, read this value with the COMMA11. informat.

Unless you explicitly define a variable first, SAS uses the informat to determine whether the variable is numeric or character. SAS also uses the informat to determine the length of character variables.

Syntax

SAS informats have the following form:

<$>*informat*<*w*>.<*d*>

where

$
indicates a character informat; its absence indicates a numeric informat.

informat
names the informat. The informat is a SAS informat or a user-defined informat that was previously defined with the INVALUE statement in PROC FORMAT. For more information about user-defined informats, see the FORMAT procedure in the *Base SAS Procedures Guide*.

w
specifies the informat width, which for most informats is the number of columns in the input data.

d
specifies an optional decimal scaling factor in the numeric informats. SAS divides the input data by 10 to the power of *d*.

Note: Even though SAS can read up to 32 digits when you specify some numeric informats, numbers with more than 15 significant digits might lose precision due to the limitations of the eight-byte floating-point representation used by most computers. △

Informats always contain a period (.) as a part of the name. If you omit the *w* and the *d* values from the informat, SAS uses default values. If the data contain decimal points, SAS ignores the *d* value and reads the number of decimal places that are actually in the input data.

If the informat width is too narrow to read all the columns in the input data, you might get unexpected results. The problem frequently occurs with the date and time informats. You must adjust the width of the informat to include blanks or special characters between the day, month, year, or time. For more information about date and time values, see the discussion on SAS date and time values in *SAS Language Reference: Concepts*.

When a problem occurs with an informat, SAS writes a note to the SAS log and assigns a missing value to the variable. Problems occur if you use an incompatible informat, such as a numeric informat to read character data, or if you specify the width of a date and time informat that causes SAS to read a special character in the last column.

Using Informats

Ways to Specify Informats

You can specify informats in the following ways:

- in an INPUT statement
- with the INPUT, INPUTC, and INPUTN functions
- in an INFORMAT statement in a DATA step or a PROC step
- in an ATTRIB statement in a DATA step or a PROC step.

INPUT Statement

The INPUT statement with an informat after a variable name is the simplest way to read values into a variable. For example, the following INPUT statement uses two informats:

```
input @15 style $3. @21 price 5.2;
```

The $*w*. character informat reads values into the variable STYLE. The *w*.*d* numeric informat reads values into the variable PRICE.

For a complete discussion of the INPUT statement, see "INPUT Statement" on page 1617.

INPUT Function

The INPUT function converts a SAS character expression using a specified informat. The informat determines whether the resulting value is numeric or character. Thus, the INPUT function is useful for converting data. For example,

```
TempCharacter='98.6';
TemperatureNumber=input(TempCharacter,4.);
```

Here, the INPUT function in combination with the *w*.*d* informat converts the character value of TempCharacter to a numeric value and assigns the numeric value 98.6 to TemperatureNumber.

Use the PUT function with a SAS format to convert numeric values to character values. See "PUT Function" on page 1056 for an example of a numeric-to-character conversion. For a complete discussion of the INPUT function, see "INPUT Function" on page 823.

INFORMAT Statement

The INFORMAT statement associates an informat with a variable. SAS uses the informat in any subsequent INPUT statement to read values into the variable. For example, in the following statements the INFORMAT statement associates the DATE*w*. informat with the variables Birthdate and Interview:

```
informat Birthdate Interview date9.;
input @63 Birthdate Interview;
```

An informat that is associated with an INFORMAT statement behaves like an informat that you specify with a colon (:) format modifier in an INPUT statement. (For details about using the colon (:) modifier, see the "INPUT Statement, List" on page 1639.) Therefore, SAS uses a modified list input to read the variable so that

- the *w* value in an informat does not determine column positions or input field widths in an external file
- the blanks that are embedded in input data are treated as delimiters unless you change the DLM= or DLMSTR= option in an INFILE statement
- for character informats, the *w* value in an informat specifies the length of character variables
- for numeric informats, the *w* value is ignored

- for numeric informats, the *d* value in an informat behaves in the usual way for numeric informats.

If you have coded the INPUT statement to use another style of input, such as formatted input or column input, that style of input is not used when you use the INFORMAT statement.

See "INPUT Statement, List" on page 1639 for more information about how to use modified list input to read data.

Note: Any time a text file originates from anywhere other than the local encoding environment, it might be necessary to specify the ENCODING= option in either ASCII or EBCDIC environments.

For example, when you read an EBCDIC text file on an ASCII platform, it is recommended that you specify the ENCODING= option in the FILENAME or INFILE statement. However, if you use the DSD and the DLM= or DLMSTR= options in the FILENAME or INFILE statement, the ENCODING= option is a requirement because these options require certain characters in the session encoding (such as quotation marks, commas, and blanks).

The use of encoding-specific informats should be reserved for use with true binary files. That is, they contain both character and non-character fields. △

ATTRIB Statement

The ATTRIB statement can also associate an informat, as well as other attributes, with one or more variables. For example, in the following statements, the ATTRIB statement associates the DATE*w*. informat with the variables Birthdate and Interview:

```
attrib Birthdate Interview informat=date9.;
input @63 Birthdate Interview;
```

An informat that is associated by using the INFORMAT= option in the ATTRIB statement behaves like an informat that you specify with a colon (:) format modifier in an INPUT statement. (For details about using the colon (:) modifier, see the "INPUT Statement, List" on page 1639.) Therefore, SAS uses a modified list input to read the variable in the same way as it does for the INFORMAT statement.

See "ATTRIB Statement" on page 1448 for more information.

Permanent versus Temporary Association

When you specify an informat in an INPUT statement, SAS uses the informat to read input data values during that DATA step. SAS, however, does not permanently associate the informat with the variable. To permanently associate a format with a variable, use an INFORMAT statement or an ATTRIB statement. SAS permanently associates an informat with the variable by modifying the descriptor information in the SAS data set.

User-Defined Informats

In addition to the informats that are supplied with Base SAS software, you can create your own informats. In Base SAS software, PROC FORMAT allows you to create your own informats and formats for both character and numeric variables. For more information about user-defined informats, see the FORMAT procedure in the *Base SAS Procedures Guide*.

When you execute a SAS program that uses user-defined informats, these informats should be available. The two ways to make these informats available are

- to create permanent, not temporary, informats with PROC FORMAT

- to store the source code that creates the informats (the PROC FORMAT step) with the SAS program that uses them.

If you execute a program that cannot locate a user-defined informat, the result depends on the setting of the FMTERR= system option. If the user-defined informat is not found, then these system options produce these results:

System Options	Results
FMTERR	SAS produces an error that causes the current DATA or PROC step to stop.
NOFMTERR	SAS continues processing by substituting a default informat.

Although using NOFMTERR enables SAS to process a variable, you lose the information that the user-defined informat supplies. This option can cause a DATA step to misread data, and it can produce incorrect results.

To avoid problems, make sure that users of your program have access to all the user-defined informats that are used.

Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms

Definitions

Integer values for integer binary data are typically stored in one of three sizes: one–byte, two–byte, or four–byte. The ordering of the bytes for the integer varies depending on the platform (operating environment) on which the integers were produced.

The ordering of bytes differs between the "big endian" and the "little endian" platforms. These colloquial terms are used to describe byte ordering for IBM mainframes (big endian) and for Intel-based platforms (little endian). In the SAS System, the following platforms are considered big endian: IBM mainframe, HP-UX, AIX, Solaris on SPARC, and Macintosh. In SAS, the following platforms are considered little endian: Intel ABI, Linux, OpenVMS Alpha, OpenVMS Integrity, Solaris on x64, Tru64 UNIX, and Windows.

How the Bytes Are Ordered

On big endian platforms, the value 1 is stored in binary and is represented here in hexadecimal notation. One byte is stored as 01, two bytes as 00 01, and four bytes as 00 00 00 01. On little endian platforms, the value 1 is stored in one byte as 01 (the same as big endian), in two bytes as 01 00, and in four bytes as 01 00 00 00.

If an integer is negative, the "two's complement" representation is used. The high-order bit of the most significant byte of the integer will be set on. For example, –2 would be represented in one, two, and four bytes on big endian platforms as FE, FF FE, and FF FF FF FE respectively. On little endian platforms, the representation would be FE, FE FF, and FE FF FF FF. These representations result from the output of the integer binary value –2 expressed in hexadecimal representation.

Reading Data Generated on Big Endian or Little Endian Platforms

SAS can read signed and unsigned integers regardless of whether they were generated on a big endian or a little endian system. Likewise, SAS can write signed and unsigned integers in both big endian and little endian format. The length of these integers can be up to eight bytes.

The following table shows which informat to use for various combinations of platforms. In the Sign? column, "no" indicates that the number is unsigned and cannot be negative. "Yes" indicates that the number can be either negative or positive.

Table 5.1 SAS Informats and Byte Ordering

Platform for Which the Data Was Created	Platform the Data Is Read on	Signed Integer	Informat
big endian	big endian	yes	IB or S370FIB
big endian	big endian	no	PIB, S370FPIB, S370FIBU
big endian	little endian	yes	IBR
big endian	little endian	no	PIBR
little endian	big endian	yes	IBR
little endian	big endian	no	PIBR
little endian	little endian	yes	IB or IBR
little endian	little endian	no	PIB or PIBR
big endian	either	yes	S370FIB
big endian	either	no	S370FPIB
little endian	either	yes	IBR
little endian	either	no	PIBR

Integer Binary Notation in Different Programming Languages

The following table compares integer binary notation according to programming language.

Table 5.2 Integer Binary Notation and Programming Languages

Language	2 Bytes or 8-Bit Systems	4 Bytes or 16-Bit Systems	8 Bytes or 64-Bit Systems
SAS	IB2., IBR2., PIB2.,PIBR2., S370FIB2., S370FIBU2., S370FPIB2.	IB4., IBR4., PIB4., PIBR4., S370FIB4., S370FIBU4., S370FPIB4.	IB8., IBR8., PIB8., PIBR8., S370FIB8., S370FIBU8., S370FPIB8.
C	short	int	long*
Java	short	int	long*
Visual Basic 6.0	short	long*	none

Language	2 Bytes or 8-Bit Systems	4 Bytes or 16-Bit Systems	8 Bytes or 64-Bit Systems
Visual Basic.NET	short	integer	long*
PL/I	fixed bin(15)	fixed bin(31)	fixed bin(63)
Fortran	integer*2	integer*4	integer*8
COBOL	comp pic 9(4)	comp pic 9(8)	comp pic 9(16)
IBM assembler	H	F	FD

* the size of integers declared as long depends on the operating environment

* the size of integers declared as long depends on the operating environment

Working with Packed Decimal and Zoned Decimal Data

Definitions

Packed decimal — specifies a method of encoding decimal numbers by using each byte to represent two decimal digits. Packed decimal representation stores decimal data with exact precision. The fractional part of the number is determined by the informat or format because there is no separate mantissa and exponent.

An advantage of using packed decimal data is that exact precision can be maintained. However, computations involving decimal data might become inexact due to the lack of native instructions.

Zoned decimal — specifies a method of encoding decimal numbers in which each digit requires one byte of storage. The last byte contains the number's sign as well as the last digit. Zoned decimal data produces a printable representation.

Nibble — specifies 1/2 of a byte.

Types of Data

Packed Decimal Data

A packed decimal representation stores decimal digits in each "nibble" of a byte. Each byte has two nibbles, and each nibble is indicated by a hexadecimal character. For example, the value 15 is stored in two nibbles, using the hexadecimal characters 1 and 5.

The sign indication is dependent on your operating environment. On IBM mainframes, the sign is indicated by the last nibble. With formats, C indicates a positive value, and D indicates a negative value. With informats, A, C, E, and F indicate positive values, and B and D indicate negative values. Any other nibble is invalid for signed packed decimal data. In all other operating environments, the sign is indicated in its own byte. If the high-order bit is 1, then the number is negative. Otherwise, it is positive.

The following applies to packed decimal data representation:

- You can use the S370FPD format on all platforms to obtain the IBM mainframe configuration.

- You can have unsigned packed data with no sign indicator. The packed decimal format and informat handles the representation. It is consistent between ASCII and EBCDIC platforms.
- Note that the S370FPDU format and informat expects to have an F in the last nibble, while packed decimal expects no sign nibble.

Zoned Decimal Data

The following applies to zoned decimal data representation:

- A zoned decimal representation stores a decimal digit in the low order nibble of each byte. For all but the byte containing the sign, the high-order nibble is the numeric zone nibble (F on EBCDIC and 3 on ASCII).
- The sign can be merged into a byte with a digit, or it can be separate, depending on the representation. But the standard zoned decimal format and informat expects the sign to be merged into the last byte.
- The EBCDIC and ASCII zoned decimal formats produce the same printable representation of numbers. There are two nibbles per byte, each indicated by a hexadecimal character. For example, the value 15 is stored in two bytes. The first byte contains the hexadecimal value F1 and the second byte contains the hexadecimal value C5.

Packed Julian Dates

The following applies to packed Julian dates:

- The two formats and informats that handle Julian dates in packed decimal representation are PDJULI and PDJULG. PDJULI uses the IBM mainframe year computation, while PDJULG uses the Gregorian computation.
- The IBM mainframe computation considers 1900 to be the base year, and the year values in the data indicate the offset from 1900. For example, 98 means 1998, 100 means 2000, and 102 means 2002. 1998 would mean 3898.
- The Gregorian computation allows for 2–digit or 4–digit years. If you use 2–digit years, SAS uses the setting of the YEARCUTOFF= system option to determine the true year.

Platforms Supporting Packed Decimal and Zoned Decimal Data

Some platforms have native instructions to support packed and zoned decimal data, while others must use software to emulate the computations. For example, the IBM mainframe has an Add Pack instruction to add packed decimal data, but the Intel-based platforms have no such instruction and must convert the decimal data into some other format.

Languages Supporting Packed Decimal and Zoned Decimal Data

Several languages support packed decimal and zoned decimal data. The following table shows how COBOL picture clauses correspond to SAS formats and informats.

IBM VS COBOL II Clauses	Corresponding S370Fxxx Formats/Informats
PIC S9(X) PACKED-DECIMAL	S370FPDw.
PIC 9(X) PACKED-DECIMAL	S370FPDUw.
PIC S9(W) DISPLAY	S370FZDw.
PIC 9(W) DISPLAY	S370FZDUw.
PIC S9(W) DISPLAY SIGN LEADING	S370FZDLw.
PIC S9(W) DISPLAY SIGN LEADING SEPARATE	S370FZDSw.
PIC S9(W) DISPLAY SIGN TRAILING SEPARATE	S370FZDTw.

For the packed decimal representation listed above, X indicates the number of digits represented, and W is the number of bytes. For PIC S9(X) PACKED-DECIMAL, W is **`ceil((x+1)/2)`**. For PIC 9(X) PACKED-DECIMAL, W is **`ceil(x/2)`**. For example, PIC S9(5) PACKED-DECIMAL represents five digits. If a sign is included, six nibbles are needed. **`ceil((5+1)/2)`** has a length of three bytes, and the value of W is 3.

Note that you can substitute COMP-3 for PACKED-DECIMAL.

In IBM assembly language, the P directive indicates packed decimal, and the Z directive indicates zoned decimal. The following shows an excerpt from an assembly language listing, showing the offset, the value, and the DC statement:

```
offset  value (in hex)      inst label     directive

+000000 00001C                2 PEX1       DC PL3'1'
+000003 00001D                3 PEX2       DC PL3'-1'
+000006 F0F0C1                4 ZEX1       DC ZL3'1'
+000009 F0F0D1                5 ZEX2       DC ZL3'1'
```

In PL/I, the FIXED DECIMAL attribute is used in conjunction with packed decimal data. You must use the PICTURE specification to represent zoned decimal data. There is no standardized representation of decimal data for the Fortran or the C languages.

Summary of Packed Decimal and Zoned Decimal Formats and Informats

SAS uses a group of formats and informats to handle packed and zoned decimal data. The following table lists the type of data representation for these formats and informats. Note that the formats and informats that begin with S370 refer to IBM mainframe representation.

Format	Data Type Representation	Corresponding Informat	Comments
PD	Packed decimal	PD	Local signed packed decimal
PK	Packed decimal	PK	Unsigned packed decimal; not specific to your operating environment
ZD	Zoned decimal	ZD	Local zoned decimal

Format	Data Type Representation	Corresponding Informat	Comments
none	Zoned decimal	ZDB	Translates EBCDIC blank (hex 40) to EBCDIC zero (hex F0), and then corresponds to the informat as zoned decimal
none	Zoned decimal	ZDV	Non-IBM zoned decimal representation
S370FPD	Packed decimal	S370FPD	Last nibble C (positive) or D (negative)
S370FPDU	Packed decimal	S370FPDU	Last nibble always F (positive)
S370FZD	Zoned decimal	S370FZD	Last byte contains sign in upper nibble: C (positive) or D (negative)
S370FZDU	Zoned decimal	S370FZDU	Unsigned; sign nibble always F
S370FZDL	Zoned decimal	S370FZDL	Sign nibble in first byte in informat; separate leading sign byte of hex C0 (positive) or D0 (negative) in format
S370FZDS	Zoned decimal	S370FZDS	Leading sign of - (hex 60) or + (hex 4E)
S370FZDT	Zoned decimal	S370FZDT	Trailing sign of - (hex 60) or + (hex 4E)
PDJULI	Packed decimal	PDJULI	Julian date in packed representation - IBM computation
PDJULG	Packed decimal	PDJULG	Julian date in packed representation - Gregorian computation
none	Packed decimal	RMFDUR	Input layout is: *mmssttt*F
none	Packed decimal	SHRSTAMP	Input layout is: *yyyyddd*F*hhmmssth*, where *yyyyddd*F is the packed Julian date; *yyyy* is a 0-based year from 1900
none	Packed decimal	SMFSTAMP	Input layout is: *xxxxxxxxyyyyddd*F, where *yyyyddd*F is the packed Julian date; *yyyy* is a 0-based year from 1900

Format	Data Type Representation	Corresponding Informat	Comments
none	Packed decimal	PDTIME	Input layout is: 0*hhmmss*F
none	Packed decimal	RMFSTAMP	Input layout is: 0*hhmmss*F*yyyyddd*F, where *yyyyddd*F is the packed Julian date; yyyy is a 0-based year from 1900

Reading Dates and Times Using the ISO 860 Basic and Extended Notations

ISO 8601 Formatting Symbols

The following list explains the formatting symbols that are used to notate the ISO 8601 dates, time, datetime, durations, and interval values:

n specifies a number that represents the number of years, months, or days

P indicates that the duration that follows is specified by the number of years, months, days, hours, minutes, and seconds

T indicates that a time value follows. Any value with a time must begin with T.

Requirement: Time values that are read by the extended notation informats that begin with the characters E8601 must use an uppercase T.

W indicates that the duration is specified in weeks.

Z indicates that the time value is the time in Greenwich, England, or UTC time.

+|- the + indicates the time zone offset to the east of Greenwich, England. The - indicates the time zone offset to the west of Greenwich, England.

yyyy specifies a four-digit year

mm as part of a date, specifies a two-digit month, 01 - 12

dd specifies a two-digit day, 01 - 31

hh specifies a two-digit hour, 00 - 24

mm as part of a time, specifies a two-digit minute, 00 - 59

ss specifies a two-digit second, 00 - 59

fff | *ffffff* specifies an optional fraction of a second using the digits 0 - 9:

fff use 1 - 3 digits for values read by the $N8601B informat and the $N8601E informat

ffffff	use 1 - 6 digits for informat other than the $N8601B informat and the $N8601E informat

Y	indicates that a year value proceeds this character in a duration
M	as part of a date, indicates that a month value proceeds this character in a duration
D	indicates that a day value proceeds this character in a duration
H	indicates that an hour value proceeds this character in a duration
M	as part of a time, indicates that a minute value proceeds this character in a duration
S	indicates that a seconds value proceeds this character in a duration

Reading ISO 8601 Date, Time, and Datetime Values

SAS reads ISO 8601 dates, times, and datetimes using various informats, and the resulting values are SAS date, time, or datetime values. The following table shows different date, time, and datetime forms and the informats you use to read them:

Table 5.3 Informats for Reading ISO 8601 Dates, Times, and Datetimes

Date, Time, or Datetime	ISO 8601 Notation	Example	Informat
Basic Notations			
Date	YYYYMMDD	20080915	B8601DA*w*.
Time	hhmmssnnnnnn	155300322348	B8601TM*w.d*
Time with time zone	hhmmss+|-hhmm	155300+0500	B8601TZ*w.d*
	hhmmssZ	155300Z	B8601TZ*w.d*
Convert to local time with time zone	hhmmss+|-hhmm	155300+0500	B8601TZ*w.d*
Datetime	YYYYMMDDThhmmss nnnnnn	20080915T155300	B8601DT*w.d*
Datetime with timezone	YYYYMMDDThhmmss +|-hhmm	20080915T155300+0500	B8601DZ*w.d*
	YYYYMMDDThhmmssZ	20080915T155300Z	B8601DZ*w.d*
Read the date from a datetime	YYYYMMD	20080915	B8601DN*w*.
Extended Notations			
Date	YYYY-MM-DD	2008-09-15	E8601DA*w*.
Time	hh:mm:ss.nnnnnn	15:53:00.322348	E8601TM*w.d*
Time with time zone	hh:mm:ss.nnnnnn+|-hh:mm	15:53:00+05:00	E8601TZ*w.d*
Convert to local time with time zone	hh:mm:ss.nnnnnn+|-hh:mm	15:53:00+05:00	E8601LZ*w.d*

Date, Time, or Datetime	ISO 8601 Notation	Example	Informat
Datetime	YYYY-MM-DDT hh:mm:ss.nnnnnn	2008-09-15T15:53:00	E8601DT*w.d*
Datetime with time zone	YYYY-MM-DDT hh:mm:ss.nnnnnn +|-hh:mm	2008-09-15T15:53:00+05:00	E8601DZ*w.d*
Read date from a datetime	YYYY-MM-DD	2008-09-15	E8601DN*w*.

Reading ISO 8601 Duration, Interval, and Datetime Values

Informats That Read Duration, Interval, and Datetime Values

SAS uses two informats that reads ISO datetime, duration, and interval values.

$N8601B informat
: reads duration, interval, and datetime values that are specified in either the basic notation or the extended notation

$N8601E informat
: reads duration, interval, and datetime values that are specified only in the extended notation

Use the $N8601E informat when you want to make sure that you are in compliance with the extended notation.

The datetime values that are read by these informats result in a SAS character representation. If you want a datetime value to be read as a numeric value, use the B8601DT informat, the B8601DZ informat, the E8601DT informat, or the E8601DZ informat.

Complete Duration, Interval, and Datetime Notations

The following table shows the formatting of duration, datetime, and interval values that can be read in the complete form:

Time Component	ISO 8601 Notation	Example
Duration - Basic Notation	PYYYYMMDDThhmmss	P20080915T155300
	-PYYYYMMDDThhmmss	-P20080915T155300
Duration - Extended Notation	PYYYY-MM-DDThh:mm:ss	P2008-09-15T15:53:00
	-PYYYY-MM-DDThh:mm:ss	-P2008-09-15T15:53:00
Duration - Basic and Extended Notation	PnYnMnDTnHnMnS	P2y10m14dT20h13m45s
	-PnYnMnDTnHnMnS	-P2n10m14dT20h13m45s
	PnW (weeks)	P6w
Interval - Basic Notation	YYYYMMDDThhmmss/ YYYYMMDDThhmmss	20080915T155300/ 20101113T000000
	PnYnMnDTnHnMnS/ YYYYMMDDThhmmss	P2y10M14dT20h13m45s/ 20080915T155300

Time Component	ISO 8601 Notation	Example
	YYYYMMDDThhmmss/ PnYnMnDTnHnMnS	20080915T155300/ P2y10M14dT20h13m45s
Interval- Extended Notation	YYYY-MM-DDThh:mm:ss/ YYYY-MM-DDThh:mm:ss	2008-09-15T15:53:00/ 2010-11-13T00:00:00
	PnYnMnDTnHnMnS/ YYYY-MM-DDThh:mm:ss	P2y10M14dT20h13m45s/ 2008-09-15T15:53:00
	YYYY-MM-DDThh:mm:ss/ PnYnMnDTnHnMnS	2008-09-15T15:53:00/ P2y10M14dT20h13m45s
Datetime-Basic Notation	YYYYMMDDThhmmss.fff+|-hhmm	20080915T155300
	(all blank)	
Datetime-Extended Notation	YYYY-MM-DDThh:mm:ss.fff+|-hhmm	2008-09-15T15:53:00 +04:30
	(all blank)	

Reading Omitted Components

One or more date or time components can be omitted from a datetime value or a duration value that is in the form P*yyyymmdd*. SAS reads omitted components using the $N8601B informat or the $N8601E informat, and the omitted component must be represented by a hyphen (-).

The following examples show duration, datetime, and interval values with omitted components:

p0003-02--T10:31:33
The omitted component is the number of days.

-p0003-02-02T-:31:33
The omitted component is the number of hours.

x-09-15T15:x:x
The omitted components are the number of years, minutes, and seconds.

2008-09-15T15:x:00/2010-09-15T15:x:00
The omitted components are the minutes.

When reading values that contain a time zone offset, omitted components are not allowed. Use 00 in place of omitted components.

Truncated Values

SAS reads truncated duration, datetime, and interval values, where one or more lower order components is truncated because the value is 0 or the value is not significant.

The following list shows examples of truncated values:

p00030202T1031

2008-09-15T15/2010-09-15T15:53

-p0003-03-03T-:-:-

P2y3m4dT5h6m

```
2008-09-xTx:x:x
2008
```

When reading values that contain a time zone offset, truncation is not allowed. Use 00 in place of truncated values.

Normalizing Duration Components

When a value for a duration component is greater than the largest standard value for a component, SAS normalizes the component except when the duration component is a single component. The following table shows examples of normalized duration components:

Duration	Extended Normalized Duration
p3y13m	**p0004-01**
pt24h24m65s	**P----01T-:25:05**
p3y13mT24h61m	**P0004-01-01T01:01**
p0004-13	**p0005-01**
p0003-02-61T15:61:61	**P0003-04-01T16:02:01**
p13m	**P13M**

If a component contains the largest value, such as 60 for minutes or seconds, SAS normalizes the value and replaces the value with a hyphen. For example, **pT12:60:13** becomes **PT13:-:13**.

Thirty days is used to normalize a month.

Dates and times in a datetime value that are greater than the standard value for the component are not normalized. They produce an error.

Fractions in Durations, Datetime, and Interval Values

Ending components can contain a fraction that consists of a period or a comma, followed by one to three digits. The following examples show the use of fractions in duration, datetime, and interval values:

```
200809.5
P2008-09-15T10.33
2008-09-15/P0003-03-03,333
```

Informats by Category

There are five categories of informats in this list:

Category	Description
Character	instructs SAS to read character data values into character variables.
Column Binary	instructs SAS to read data stored in column-binary or multipunched form into character and numeric variables.
Date and Time	instructs SAS to read date values into variables that represent dates, times, and datetimes.

Category	Description
ISO 8601	instructs SAS to read date, time, and datetime values that are written in the ISO 8601 standard into either numeric or character variables.
Numeric	instructs SAS to read numeric data values into numeric variables.

For information about reading column-binary data, see *SAS Language Reference: Concepts*. For information about creating user-defined informats, see the FORMAT procedure in the *Base SAS Procedures Guide*.

The following table provides brief descriptions of the SAS informats. For more detailed descriptions, see the dictionary entry for each informat.

Table 5.4 Categories and Descriptions of Informats

Category	Informats	Description
Character	"$ASCII*w*. Informat" on page 1280	Converts ASCII character data to native format.
	"$BASE64X*w*. Informat" on page 1281	Converts ASCII text into character data by using Base 64 encoding.
	"$BINARY*w*. Informat" on page 1283	Converts binary data to character data.
	"$CHAR*w*. Informat" on page 1285	Reads character data with blanks.
	"$CHARZB*w*. Informat" on page 1286	Converts binary 0s to blanks.
	"$EBCDIC*w*. Informat" on page 1287	Converts EBCDIC character data to native format.
	"$HEX*w*. Informat" on page 1288	Converts hexadecimal data to character data.
	"$OCTAL*w*. Informat" on page 1289	Converts octal data to character data.
	"$PHEX*w*. Informat" on page 1290	Converts packed hexadecimal data to character data.
	"$QUOTE*w*. Informat" on page 1295	Removes matching quotation marks from character data.
	"$UPCASE*w*. Informat" on page 1295	Converts character data to uppercase.
	"$VARYING*w*. Informat" on page 1296	Reads character data of varying length.
	"$*w*. Informat" on page 1298	Reads standard character data.
Column Binary	"$CB*w*. Informat" on page 1284	Reads standard character data from column-binary files.
	"CB*w*.*d* Informat" on page 1319	Reads standard numeric values from column-binary files.

Category	Informats	Description
	"PUNCH.*d* Informat" on page 1366	Reads whether a row of column-binary data is punched.
	"ROW*w.d* Informat" on page 1371	Reads a column-binary field down a card column.
Date and Time	"$N8601B*w.d* Informat" on page 1291	Reads complete, truncated, and omitted forms of ISO 8601 duration, datetime, and interval values that are specified in either the basic or extended notations.
	"$N8601E*w.d* Informat" on page 1293	Reads ISO 8601 duration, datetime, and interval values that are specified in the extended notation.
	"ANYDTDTE*w*. Informat" on page 1299	Reads and extracts the date value from various date, time, and datetime forms.
	"ANYDTDTM*w*. Informat" on page 1301	Reads and extracts datetime values from various date, time, and datetime forms.
	"ANYDTTME*w*. Informat" on page 1304	Reads and extracts time values from various date, time, and datetime forms.
	"B8601DA*w*. Informat" on page 1306	Reads date values that are specified in the ISO 8601 base notation *yyyymmdd*.
	"B8601DN*w*. Informat" on page 1308	Reads date values that are specified the ISO 8601 basic notation *yyyymmdd* and returns SAS datetime values where the time portion of the value is 000000.
	"B8601DT*w.d* Informat" on page 1309	Reads datetime values that are specified in the ISO 8601 basic notation *yyyymmdd*T*hhmmssffffff*.
	"B8601DZ*w.d* Informat" on page 1310	Reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime basic notation *yyyymmdd* T*hhmmss*+|-*hhmm* or *yyyymmdd*T *hhmmssffffff*Z.
	"B8601TM*w.d* Informat" on page 1312	Reads time values that are specified in the ISO 8601 basic notation *hhmmssffffff*.
	"B8601TZ*w.d* Informat" on page 1314	Reads time values that are specified in the ISO 8601 basic time notation *hhmmssfffff*+|-*hhmm* or *hhmmssffffff*Z.
	"DATE*w*. Informat" on page 1322	Reads date values in the form *ddmmmyy* or *ddmmmyyyy* .
	"DATETIME*w*. Informat" on page 1324	Reads datetime values in the form *ddmmmyy hh:mm:ss.ss* or *ddmmmyyyy hh:mm:ss.ss*.
	"DDMMYY*w*. Informat" on page 1326	Reads date values in the form *ddmmyy<yy>* or *dd-mm-yy<yy>*, where a special character, such as a hyphen (-), period (.), or slash (/), separates the day, month, and year; the year can be either 2 or 4 digits.
	"E8601DA*w*. Informat" on page 1328	Reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*.
	"E8601DN*w*. Informat" on page 1330	Reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd* and returns SAS datetime values where the time portion of the value is 000000.

Category	Informats	Description
	"E8601DT*w.d* Informat" on page 1331	Reads datetime values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.
	"E8601DZ*w.d* Informat" on page 1333	Reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime extended notation *hh*:*mm*:*ss*+|-*hh*:*mm*.*fffff* or *hh*:*mm*:*ss*.*fffff*Z.
	"E8601LZ*w.d* Informat" on page 1334	Reads Coordinated Universal Time (UTC) values that are specified in the ISO 8601 extended notation *hh*:*mm*:*ss* +|-*hh*:*mm*.*fffff* or *hh*:*mm*:*ss*.*fffff*Z and converts them to the local time.
	"E8601TM*w.d* Informat" on page 1336	Reads time values that are specified in the ISO 8601 extended notation *hh:mm:ss.ffffff*.
	"E8601TZ*w.d* Informat" on page 1338	Reads time values that are specified in the ISO 8601 extended time notation *hh*:*mm*:*ss*+|-*hh*:*mm*.*ffffff* or *hh*:*mm*:*ss* Z.
	"JULIAN*w*. Informat" on page 1346	Reads Julian dates in the form *yyddd* or *yyyyddd* .
	"MDYAMPM*w.d* Informat" on page 1347	Reads datetime values in the form *mm-dd-yy<yy> hh:mm:ss.ss* AM|PM, where a special character such as a hyphen (-), period (.), slash (/), or colon (:) separates the month, day, and year; the year can be either 2 or 4 digits.
	"MMDDYY*w*. Informat" on page 1349	Reads date values in the form *mmddyy* or *mmddyyyy* .
	"MONYY*w*. Informat" on page 1351	Reads month and year date values in the form *mmmyy* or *mmmyyyy*.
	"MSEC*w*. Informat" on page 1352	Reads TIME MIC values.
	"PDJULG*w*. Informat" on page 1357	Reads packed Julian date values in the hexadecimal form *yyyyddd*F for IBM.
	"PDJULI*w*. Informat" on page 1358	Reads packed Julian dates in the hexadecimal format *ccyyddd* F for IBM.
	"PDTIME*w*. Informat" on page 1360	Reads packed decimal time of SMF and RMF records.
	"RMFDUR*w*. Informat" on page 1368	Reads duration intervals of RMF records.
	"RMFSTAMP*w*. Informat" on page 1370	Reads time and date fields of RMF records.
	"SHRSTAMP*w*. Informat" on page 1389	Reads date and time values of SHR records.
	"SMFSTAMP*w*. Informat" on page 1390	Reads time and date values of SMF records.
	"STIMER*w*. Informat" on page 1392	Reads time values and determines whether the values are hours, minutes, or seconds; reads the output of the STIMER system option.

Category	Informats	Description
	"TIME*w*. Informat" on page 1393	Reads hours, minutes, and seconds in the form *hh:mm:ss.ss* , where special characters such as the colon (:) or the period (.) are used to separate the hours, minutes, and seconds.
	"TODSTAMP*w*. Informat" on page 1395	Reads an eight-byte time-of-day stamp.
	"TU*w*. Informat" on page 1397	Reads timer units.
	"WEEKU*w*. Informat" on page 1400	Reads the format of the number-of-week value within the year and returns a SAS date value by using the U algorithm.
	"WEEKV*w*. Informat" on page 1402	Reads the format of the number-of-week value within the year and returns a SAS date value using the V algorithm.
	"WEEKW*w*. Informat" on page 1404	Reads the format of the number-of-week value within the year and returns a SAS date value using the W algorithm.
	"YMDDTTM*w.d* Informat" on page 1408	Reads datetime values in the form *<yy>yy-mm-dd hh:mm:ss.ss* , where special characters such as a hyphen (-), period (.), slash (/), or colon (:) are used to separate the year, month, day, hour, minute, and seconds; the year can be either 2 or 4 digits.
	"YYMMDD*w*. Informat" on page 1410	Reads date values in the form *yymmdd* or *yyyymmdd* .
	"YYMMN*w*. Informat" on page 1411	Reads date values in the form *yyyymm* or *yymm* .
	"YYQ*w*. Informat" on page 1413	Reads quarters of the year in the form *yy*Q *q* or *yyyy*Q*q*.
ISO 8601	"$N8601B*w.d* Informat" on page 1291	Reads complete, truncated, and omitted forms of ISO 8601 duration, datetime, and interval values that are specified in either the basic or extended notations.
	"$N8601E*w.d* Informat" on page 1293	Reads ISO 8601 duration, datetime, and interval values that are specified in the extended notation.
	"B8601DA*w*. Informat" on page 1306	Reads date values that are specified in the ISO 8601 base notation *yyyymmdd*.
	"B8601DN*w*. Informat" on page 1308	Reads date values that are specified the ISO 8601 basic notation *yyyymmdd* and returns SAS datetime values where the time portion of the value is 000000.
	"B8601DT*w.d* Informat" on page 1309	Reads datetime values that are specified in the ISO 8601 basic notation *yyyymmdd*T*hhmmssffffff*.
	"B8601DZ*w.d* Informat" on page 1310	Reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime basic notation *yyyymmdd* T*hhmmss*+|-*hhmm* or *yyyymmdd*T *hhmmssffffff*Z.
	"B8601TM*w.d* Informat" on page 1312	Reads time values that are specified in the ISO 8601 basic notation *hhmmssffffff*.

Category	Informats	Description
	"B8601TZ*w.d* Informat" on page 1314	Reads time values that are specified in the ISO 8601 basic time notation *hhmmssfffff*+|-*hhmm* or *hhmmssffffff*Z.
	"E8601DA*w*. Informat" on page 1328	Reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*.
	"E8601DN*w*. Informat" on page 1330	Reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd* and returns SAS datetime values where the time portion of the value is 000000.
	"E8601DT*w.d* Informat" on page 1331	Reads datetime values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.
	"E8601DZ*w.d* Informat" on page 1333	Reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime extended notation *hh*: *mm*:*ss*+|-*hh*:*mm.fffff* or *hh*:*mm*:*ss.fffff*Z.
	"E8601LZ*w.d* Informat" on page 1334	Reads Coordinated Universal Time (UTC) values that are specified in the ISO 8601 extended notation *hh*:*mm*:*ss* +|-*hh*:*mm.fffff* or *hh*:*mm*:*ss.fffff*Z and converts them to the local time.
	"E8601TM*w.d* Informat" on page 1336	Reads time values that are specified in the ISO 8601 extended notation *hh:mm:ss.ffffff*.
	"E8601TZ*w.d* Informat" on page 1338	Reads time values that are specified in the ISO 8601 extended time notation *hh*:*mm*:*ss*+|-*hh*:*mm.ffffff* or *hh*:*mm*:*ss* Z.
Numeric	"BINARY*w.d* Informat" on page 1315	Converts positive binary values to integers.
	"BITS*w.d* Informat" on page 1316	Extracts bits.
	"BZ*w.d* Informat" on page 1317	Converts blanks to 0s.
	"COMMA*w.d* Informat" on page 1320	Removes embedded characters.
	"COMMAX*w.d* Informat" on page 1321	Removes embedded characters.
	"E*w.d* Informat" on page 1327	Reads numeric values that are stored in scientific notation and double-precision scientific notation.
	"FLOAT*w.d* Informat" on page 1339	Reads a native single-precision, floating-point value and divides it by 10 raised to the *d*th power.
	"HEX*w*. Informat" on page 1341	Converts hexadecimal positive binary values to either integer (fixed-point) or real (floating-point) binary values.
	"IB*w.d* Informat" on page 1341	Reads native integer binary (fixed-point) values, including negative values.
	"IBR*w.d* Informat" on page 1343	Reads integer binary (fixed-point) values in Intel and DEC formats.
	"IEEE*w.d* Informat" on page 1344	Reads an IEEE floating-point value and divides it by 10 raised to the *d* th power.

Category	Informats	Description
	"NUMX*w*.*d* Informat" on page 1353	Reads numeric values with a comma in place of the decimal point.
	"OCTAL*w*.*d* Informat" on page 1354	Converts positive octal values to integers.
	"PD*w*.*d* Informat" on page 1355	Reads data that are stored in IBM packed decimal format.
	"PERCENT*w*.*d* Informat" on page 1361	Reads percentages as numeric values.
	"PIB*w*.*d* Informat" on page 1362	Reads positive integer binary (fixed-point) values.
	"PIBR*w*.*d* Informat" on page 1363	Reads positive integer binary (fixed-point) values in Intel and DEC formats.
	"PK*w*.*d* Informat" on page 1365	Reads unsigned packed decimal data.
	"RB*w*.*d* Informat" on page 1367	Reads numeric data that are stored in real binary (floating-point) notation.
	"S370FF*w*.*d* Informat" on page 1373	Reads EBCDIC numeric data.
	"S370FIB*w*.*d* Informat" on page 1374	Reads integer binary (fixed-point) values, including negative values, in IBM mainframe format.
	"S370FIBU*w*.*d* Informat" on page 1376	Reads unsigned integer binary (fixed-point) values in IBM mainframe format.
	"S370FPD*w*.*d* Informat" on page 1378	Reads packed data in IBM mainframe format.
	"S370FPDU*w*.*d* Informat" on page 1379	Reads unsigned packed decimal data in IBM mainframe format.
	"S370FPIB*w*.*d* Informat" on page 1380	Reads positive integer binary (fixed-point) values in IBM mainframe format.
	"S370FRB*w*.*d* Informat" on page 1381	Reads real binary (floating-point) data in IBM mainframe format.
	"S370FZD*w*.*d* Informat" on page 1383	Reads zoned decimal data in IBM mainframe format.
	"S370FZDL*w*.*d* Informat" on page 1385	Reads zoned decimal leading-sign data in IBM mainframe format.
	"S370FZDS*w*.*d* Informat" on page 1386	Reads zoned decimal separate leading-sign data in IBM mainframe format.
	"S370FZDT*w*.*d* Informat" on page 1387	Reads zoned decimal separate trailing-sign data in IBM mainframe format.
	"S370FZDU*w*.*d* Informat" on page 1388	Reads unsigned zoned decimal data in IBM mainframe format.
	"TRAILSGN*w*. Informat" on page 1396	Reads a trailing plus (+) or minus (–) sign.
	"VAXRB*w*.*d* Informat" on page 1398	Reads real binary (floating-point) data in VMS format.

Category	Informats	Description
	"VMSZN*w.d* Informat" on page 1399	Reads VMS and MicroFocus COBOL zoned numeric data.
	"*w.d* Informat" on page 1407	Reads standard numeric data.
	"ZD*w.d* Informat" on page 1414	Reads zoned decimal data.
	"ZDB*w.d* Informat" on page 1416	Reads zoned decimal data in which zeros have been left blank.
	"ZDV*w.d* Informat" on page 1416	Reads and validates zoned decimal data.

Dictionary

$ASCII*w*. Informat

Converts ASCII character data to native format.

Category: Character

Syntax

$ASCII*w*.

Syntax Description

w
specifies the width of the input field.

Default: 1 if the length of the variable is undefined. Otherwise, the default is the length of the variable.

Range: 1–32767

Details

If ASCII is the native format, no conversion occurs.

Comparisons

- On an IBM mainframe system, $ASCII*w*. converts ASCII data to EBCDIC.
- On all other systems, $ASCII*w*. behaves like the $CHAR*w*. informat except that the default length is different.

Examples

```
input @1 name $ascii3.;
```

Data Line		Results*
`----+----1`	EBCDIC	ASCII
`abc`	`818283`	`616263`
`ABC`	`C1C2C3`	`414243`
`();`	`4D5D5E`	`28293B`

* The results are hexadecimal representations of codes for characters. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one character value.

$BASE64X*w*. Informat

Converts ASCII text into character data by using Base 64 encoding.

Category: Character
Alignment: left

Syntax

$BAS64X*w*.

Syntax Description

w
specifies the width of the input field.
Default: 1
Range: 1–32767

Details

Base 64 is an industry encoding method whose encoded characters are determined by using a positional scheme that uses only ASCII characters. Several Base 64 encoding schemes have been defined by the industry for specific uses, such as e-mail or content masking. SAS maps positions 0–61 to the characters A–Z, a–z, and 0–9. Position 62 maps to the character +, and position 63 maps to the character /.

The following are some uses of Base 64 encoding:

- embed binary data in an XML file
- encode passwords
- encode URLs

The '=' character in the encoded results indicates that the results have been padded with zero bits. In order for the encoded characters to be decoded, the '=' must be included in the value to be decoded.

Examples

```
input @1 b64exmpl $base64x64.;
```

Data Line	Results
RkNBMDFBNzk5M0JD	FCA01A7993BC
TXlQYXNzd29yZA==	MyPassword
d3d3Lm15ZG9tYWluLmNvbi9teWhpZGRlblVSTA==	www.mydomain.com/myhiddenURL

See Also

Format:

"$BASE64X*w*. Format" on page 109

The XMLDOUBLE option of the LIBNAME Statement for the XML engine, in the *SAS XML LIBNAME Engine: User's Guide*

$BINARY*w*. Informat

Converts binary data to character data.

Category: Character

Syntax

$BINARY*w*.

Syntax Description

w

specifies the width of the input field. Because eight bits of binary information represent one character, every eight characters of input that $BINARY*w*. reads becomes one character value stored in a variable.

If *w*< 8, $BINARY*w*. reads the data as *w* characters followed by 0s. Thus, $BINARY4. reads the characters 0101 as 01010000, which converts to an EBCDIC **&** or an ASCII **P**. If *w*> 8 but is not a multiple of 8, $BINARY*w*. reads up to the largest multiple of 8 that is less than *w* before converting the data.

Default: 8

Range: 1–32767

Details

The $BINARY*w*. informat does not interpret actual binary data, but it converts a string of characters that contains only 0s or 1s as if it is actual binary information. Therefore, use only the character digits 1 and 0 in the input, with no embedded blanks. $BINARY*w*. ignores leading and trailing blanks.

To read representations of binary codes for unprintable characters, enter an ASCII or EBCDIC equivalent for a particular character as a string of 0s and 1s. The $BINARY*w*. informat converts the string to its equivalent character value.

Comparisons

- The BINARY*w*. informat reads eight characters of input that contain only 0s or 1s as a binary representation of one byte of numeric data.
- The $HEX*w*. informat reads hexadecimal characters that represent the ASCII or EBCDIC equivalent of character data.

Examples

```
input @1 name $binary16.;
```

Data Line	Results	
`----+----1----+----2`	ASCII	EBCDIC
`0100110001001101`	**LM**	**<(**

$CB*w*. Informat

Reads standard character data from column-binary files.

Category: Column Binary

Syntax

$CB*w*.

Syntax Description

w
specifies the width of the input field.

Default: none

Range: 1–32767

Details

Column-binary data storage compresses data so that more than 80 items of data can be stored on a single "virtual" punch card.

The $CB*w*. informat reads standard character data from column-binary files, with each card column represented in two bytes. The $CB*w*. informat translates the data into standard character codes. If the combinations are invalid punch codes, SAS returns blanks and sets the automatic variable _ERROR_ to 1.

Examples

```
input @1 name $cb2.;
```

Data Line*		Results
`----+----1`	EBCDIC	ASCII
200A	**+**	**N**

* The data line is a hexadecimal representation of the column binary. The "virtual" punch card column for the example data has row 12, row 6, and row 8 punched. The binary representation is 0010 0000 0000 1010.

See Also

Informats:

"CB*w.d* Informat" on page 1319
"PUNCH.*d* Informat" on page 1366
"ROW*w.d* Informat" on page 1371

"How to Read Column-Binary Data" in *SAS Language Reference: Concepts*

$CHAR*w*. Informat

Reads character data with blanks.

Category: Character

Syntax

$CHAR*w*.

Syntax Description

w
specifies the width of the input field.

Default: 8 if the length of the variable is undefined. Otherwise, the default is the length of the variable

Range: 1–32767

Details

The $CHAR*w*. informat does not trim leading and trailing blanks or convert a single period in the input data field to a blank before storing values. If you use $CHAR*w*. in an INFORMAT or ATTRIB statement within a DATA step to read list input, then by default SAS interprets any blank embedded within data as a field delimiter, including leading blanks.

Comparisons

- The $CHAR*w*. informat is almost identical to the $*w*. informat. However $CHAR*w*. does not trim leading blanks or convert a single period in the input data field to a blank, while the $*w*. informat does.
- Use the table below to compare the SAS informat $CHAR8. with notation in other programming languages:

Language	Character Notation
SAS	$CHAR8.
IBM 370 assembler	CL8
C	char [8]
COBOL	PIC x(8)
Fortran	A8
PL/I	CHAR(8)

Examples

```
input @1 name $char5.;
```

Data Line	Results*
`----+----1`	
`XYZ`	`XYZ##`
`XYZ`	`#XYZ#`
`.`	`##.##`
`X YZ`	`#X#YZ`

* The character # represents a blank space.

$CHARZB*w*. Informat

Converts binary 0s to blanks.

Category: Character

Syntax

$CHARZB*w*.

Syntax Description

w
specifies the width of the input field.

Default: 1 if the length of the variable is undefined. Otherwise, the default is the length of the variable.

Range: 1–32767

Details

The $CHARZB*w*. informat does not trim leading and trailing blanks in character data before it stores values.

Comparisons

The $CHARZB*w*. informat is identical to the $CHAR*w*. informat except that $CHARZB*w*. converts any byte that contains a binary 0 to a blank character.

Examples

```
input @1 name $charzb5.;
```

Data Line*		Results
EBCDIC	ASCII	
E7E8E90000	**58595A0000**	**XYZ##**
00E7E8E900	**0058595A00**	**#XYZ#**
00E700E8E9	**005800595A**	**#X#YZ**

* The data lines are hexadecimal representations of codes for characters. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one character.
** The character # represents a blank space.

$EBCDIC*w*. Informat

Converts EBCDIC character data to native format.

Category: Character

Syntax

$EBCDIC*w*.

Syntax Description

w
specifies the width of the input field.

Default: 1 if the length of the variable is undefined. Otherwise, the default is the length of the variable.

Range: 1–32767

Details

If EBCDIC is the native format, no conversion occurs.

Note: Any time a text file originates from anywhere other than the local encoding environment, it might be necessary to specify the ENCODING= option on either ASCII or EBCDIC environments.

When you read an EBCDIC text file on an ASCII platform, it is recommended that you specify the ENCODING= option in the FILENAME or INFILE statement. However, if you use the DSD and the DLM= or DLMSTR= options in the FILENAME or INFILE statement, the ENCODING= option is a requirement because these options require certain characters in the session encoding (such as quotation marks, commas, and blanks).

The use of encoding-specific informats should be reserved for use with true binary files. That is, they contain both character and non-character fields. △

Comparisons

□ On an IBM mainframe system, $EBCDIC*w*. behaves like the $CHAR*w*. informat.

□ On all other systems, $EBCDIC*w*. converts EBCDIC data to ASCII.

Examples

```
input @1 name $ebcdic3.
```

Data Line	Results*	
`----+----1`	ASCII	EBCDIC
`qrs`	`717273`	`9899A2`
`QRS`	`515253`	`D8D9E2`
`+;>`	`2B3B3E`	`4E5E6E`

* The results are hexadecimal representations of codes for characters. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one character value.

$HEX*w*. Informat

Converts hexadecimal data to character data.

Category: Character

See: $HEX*w*. Informat in the documentation for your operating environment.

Syntax

$HEX*w*.

Syntax Description

w

specifies the number of digits of hexadecimal data.

If *w*=1, $HEX*w*. pads a trailing hexadecimal 0. If *w* is an odd number that is greater than 1, then $HEX*w*. reads *w*–1 hexadecimal characters.

Default: 2

Range: 1–32767

Details

The $HEX*w*. informat converts every two digits of hexadecimal data into one byte of character data. Use $HEX*w*. to encode hexadecimal values into a character variable when your input method is limited to printable characters.

Comparisons

The HEX*w*. informat reads two digits of hexadecimal data at a time and converts them into one byte of numeric data.

Examples

```
input @1 name $hex4.;
```

Data Line	Results	
`----+----1`	ASCII	EBCDIC
`6C6C`	`ll`	`%%`

$OCTAL*w*. Informat

Converts octal data to character data.

Category: Character

Syntax

$OCTAL*w*.

Syntax Description

w

specifies the width of the input field in bits. Because one digit of octal data represents three bits of binary information, increment the value of *w* by three for every column of octal data that $OCTAL*w*. will read.

Default: 3

Range: 1–32767

Details

Eight bits of binary data represent the code for one digit of character data. Therefore, you need at least three digits of octal data to represent one digit of character data, which includes an extra bit. $OCTAL*w*. treats every three digits of octal data as one digit of character data, ignoring the extra bit.

Use $OCTAL*w*. to read octal representations of binary codes for unprintable characters. Enter an ASCII or EBCDIC equivalent for a particular character in octal notation. Then use $OCTAL*w*. to convert it to its equivalent character value.

Use only the digits 0 through 7 in the input, with no embedded blanks. $OCTAL*w*. ignores leading and trailing blanks.

Comparisons

The OCTAL*w*. informat reads octal data and converts them into the numeric equivalents.

Examples

```
input @1 name $octal9.;
```

Data Line	Results	
`----+----1`	EBCDIC	ASCII
`114`	`<`	`L`

$PHEX*w*. Informat

Converts packed hexadecimal data to character data.

Category: Character

Syntax

$PHEX*w*.

Syntax Description

w
specifies the number of bytes in the input.

When you use $PHEX*w*. to read packed hexadecimal data, the length of the variable is the number of bytes that are required to store the resulting character value, not *w*. In general, a character variable whose length is implicitly defined with $PHEX*w*. has a length of 2*w*–1.

Default: 2

Range: 1–32767

Details

Packed hexadecimal data are like packed decimal data, except that all hexadecimal characters are valid. In packed hexadecimal data, the value of the low-order nibble has no meaning. In packed decimal data, the value of the low-order nibble indicates the sign of the numeric value that the data represent. The $PHEX*w*. informat returns a character value and treats the value of the sign nibble as if it were `X'F'`, regardless of its actual value.

Comparisons

The PD*w*.*d*. informat reads packed decimal data and converts them to numeric data.

Examples

```
input @1 devaddr $phex2.;
```

Data Line*	Results
0001111000001111	**1E0**

*The data line represents two bytes of actual binary data, with each half byte corresponding to a single hexadecimal digit. The equivalent hexadecimal representation for the data line is 1E0F.

$N8601B*w.d* Informat

Reads complete, truncated, and omitted forms of ISO 8601 duration, datetime, and interval values that are specified in either the basic or extended notations.

Category: Date and Time
ISO 8601

Alignment: left

Time Zone Informat: No

ISO 8601 Element: 5.4.4 Complete representation

Syntax

$N8601B*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601B informat reads ISO 8601 duration, datetime, and interval values as character data for the following basic notations:

Time Component	ISO 8601 Notation	Example
Duration	P*yyyy-mm-dd*T*hh:mm:ss.fff*	P2008-09-15T15:53:00
	P*yyyymmdd*T*hhmmss*	P00020304T050607
	P*n*Y*n*M*n*DT*n*H*n*M*n.fff*S	P2y10m14dT20h13m45.222s
	P*n*W	P6w
Interval	*yyyy-mm-dd*T*hh:mm:ss.fff*/ *yyyy-mm-dd*T*hh:mm:ss.fff*	2008-09-15T15:53:00/ 2010-11-13T00:00:00
	*yyyymmdd*T*hhmmss.fff*/ *yyyymmdd*T*hhmmss.fff*	20080915T155300/ 20101115T120000
	P*n*Y*n*M*n*DT*n*H*n*M*n.fff*S/ *yyyy-mm-dd*T*hh:mm:ss.fff*	P2y10M14dT20h13m45s/ 2008-09-15T15:53:00
	*yyyy-mm-dd*T*hh:mm:ss.fff*/ P*n*Y*n*M*n*DT*n*H*n*M*n.fff*S	2008-09-15T15:53:00/ P2y10M14dT20h13m45s
Datetime	*yyyy-mm-dd*T*hh:mm:ss.fff*	2008-09-15T15:53:00
	yyyymmddThhmmss.fff	20080915T155300

The $N8601B informat also reads ISO 8601 duration, interval, and datetime components that contain omitted components or truncated components. Omitted components must use a single hyphen (-)to represent the component.

Comparisons

The $N8601B informat reads durations, intervals, and datetimes that are specified in either the basic or extended notation. The $N8601E informat reads valid durations, intervals, and datetimes that are specified only in the extended notation. Use the $N8601E informat when you need to ensure compliance with the extended notation.

Examples

```
input @1 i860 $n8601b.;
```

Data Line	Results
`p0002-04-05t5:1:12`	`0002405050112FFC`
`2008-09-15T15:53:00/2010-09-15T00:00:00`	`2008915155300FFD2010915000000FFD`
`p0033-01-04T3:2:55/2008-09-15T15:53:00`	`0033104030255FFC2008915155300FFD`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

$N8601E*w.d* Informat

Reads ISO 8601 duration, datetime, and interval values that are specified in the extended notation.

Category: Date and Time
ISO 8601
Alignment: left
Time Zone Informat: No
ISO 8601 Element: 5.4.4 Complete representation

Syntax

$N8601E*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 50

Range: 1 - 200

Requirement: The minimum length for a duration value or a datetime value is 16. The minimum length for an interval value is 16.

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0 - 3

Details

The $N8601E informat reads ISO 8601 durations, datetime, and interval values that can be specified in the following the extended notations:

Time Component	ISO 8601 Notation	Example
Duration	P*yyyy-mm-dd*T*hh:mm:ss.fff*	P2008-09-15T15:53:00
	PnW	P6w
Interval	*yyyy-mm-dd*T*hh:mm:ss.fff*/ *yyyy-mm-dd*T*hh:mm:ss.fff*	2008-09-15T15:53:00/ 2010-11-13T00:00:00
	*yyyy-mm-dd*T*hh:mm:ss.fff*/ P*n*Y*n*M*n*DT*n*H*n*M*ns.fff*S	2008-09-15T15:53:00/ P2y10M14dT20h13m45s
Datetime	*yyyy-mm-dd*T*hh:mm:ss.fff*	2008-09-15T15:53:00

n	specifies a number that represents the number of years, months, or days
P	is the character that is used to indicate that the duration that follows is specified by the number of years, months, days, hours, minutes, and seconds
W	is the character that is used to designate that the duration is specified in weeks.
T	is the character used to designate that a time value follows. If all time values are 0, T is not required.
yyyy	specifies a four-digit year
mm	specifies a two-digit month, 01 - 12
dd	specifies a two-digit day, 01 - 31
hh	specifies a two-digit hour, 00 - 23
mm	specifies a two-digit minute, 00 - 59
ss	specifies a two-digit second, 00 - 59
fff	specifies an optional fraction of a second that can be 1 - 3 digits, 0 - 9
Y	is the character that is used to designate years in a duration
M	is the character used to designate months in a duration
D	is the character used to designate days in a duration
H	is the character used to designate hours in a duration
S	is the character used to designate minutes in a duration
S	is the character used to designate seconds in a duration

Comparisons

The $N8601E informat reads only valid durations, intervals, and datetimes that are specified in the extended notation. The $N8601B informat reads valid durations, intervals, and datetimes that are specified in either the basic or extended notation. Use the $N8601E informat when you need to ensure compliance with the extended notation.

Examples

```
input @1 i860 $n8601e.;
```

Data Line	Results
p0002-04-05t5:1:12s	0002405050112FFC
2008-09-15T15:53:00/2010-09-15T00:00:00	2008915155300FFD2010915000000FFD
p0033-01-04T3:2:55/2008-09-15T15:53:00	0033104030255FFC2008915155300FFD

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

$QUOTE*w*. Informat

Removes matching quotation marks from character data.

Category: Character

Syntax

$QUOTE*w*.

Syntax Description

w
specifies the width of the input field.

Default: 8 if the length of the variable is undefined. Otherwise, the default is the length of the variable.

Range: 1–32767

Examples

```
input @1 name $quote7.;
```

Data Line	Results
`----+----1`	
`'SAS'`	**SAS**
`"SAS"`	**SAS**
`"SAS's"`	**SAS's**

$UPCASE*w*. Informat

Converts character data to uppercase.

Category: Character

Syntax

$UPCASE*w.*

Syntax Description

w
specifies the width of the input field.

Default: 8 if the length of the variable is undefined. Otherwise, the default is the length of the variable.

Range: 1–32767

Details

Special characters, such as hyphens, are not altered.

Examples

```
input @1 name $upcase3.;
```

Data Line	Results
`----+----1`	
sas	**SAS**

$VARYING*w*. Informat

Reads character data of varying length.

Valid: in a DATA step

Category: Character

Syntax

$VARYING*w. length-variable*

Syntax Description

w
specifies the maximum width of a character field for all the records in an input file.

Default: 8 if the length of the variable is undefined. Otherwise, the default is the length of the variable.

Range: 1–32767

length-variable
specifies a numeric variable that contains the width of the character field in the current record. SAS obtains the value of *length-variable* by reading it directly from a field that is described in an INPUT statement or by calculating its value in the DATA step.

Requirement: You must specify *length-variable* immediately after $VARYING*w*. in an INPUT statement.

Restriction: *Length-variable* cannot be an array reference.

Tip: If the value of *length-variable* is 0, negative, or missing, SAS reads no data from the corresponding record. A value of 0 for *length-variable* enables you to read zero-length records and fields. If *length-variable* is greater than 0 but less than *w*, SAS reads the number of columns that are specified by *length-variable*. Then SAS pads the value with trailing blanks up to the maximum width that is assigned to the variable. If *length-variable* is greater than or equal to *w*, SAS reads *w* columns.

Details

Use $VARYING*w*. when the length of a character value differs from record to record. After reading a data value with $VARYING*w*., the pointer's position is set to the first column after the value.

Examples

Example 1: Obtaining a Current Record Length Directly

```
input fwidth 1. name $varying9. fwidth;
```

Data Line	Results
`----+----1`	
`5shark`	`shark`
`3sunfish`	`sun`
`8bluefish`	`bluefish`

* Notice the result of reading the second data line.

Example 2: Obtaining a Record Length Indirectly Use the LENGTH= option in the INFILE statement to obtain a record length indirectly. The input data lines and results follow the explanation of the SAS statements.

```
data one;
   infile file-specification length=reclen;
   input @;
   fwidth=reclen-9;
   input name $ 1-9
         @10 class $varying20. fwidth;
run;
```

The LENGTH= option in the INFILE statement assigns the internally stored record length to RECLEN when the first INPUT statement executes. The trailing @ holds the record for another INPUT statement. Next, the assignment statement calculates the value of the varying-length field by subtracting the fixed-length portion of the record from the total record length. The variable FWIDTH contains the length of the last field and becomes the *length-variable* argument to the $VARYING20. informat.

Data Line		Results	
----+----1----+----2			
PATEL	CHEMISTRY	PATEL	CHEMISTRY
JOHNSON	GEOLOGY	JOHNSON	GEOLOGY
WILCOX	ART	WILCOX	ART

$*w*. Informat

Reads standard character data.

Category: Character

Alias: $F*w*.

Syntax

$*w*.

Syntax Description

w

specifies the width of the input field. You must specify *w* because SAS does not supply a default value.

Range: 1–32767

Details

The $*w*. informat trims leading blanks and left aligns the values before storing the text. In addition, if a field contains only blanks and a single period, $*w*. converts the period to a blank because it interprets the period as a missing value. The $*w*. informat treats two or more periods in a field as character data.

Comparisons

The $*w*. informat is almost identical to the $CHAR*w*. informat. However, $CHAR*w*. does not trim leading blanks nor does it convert a single period in an input field to a blank, while $*w*. does both.

Examples

```
input @1 name $5.;
```

Data Line	Results*
`----+----1`	
`XYZ`	`XYZ##`
`XYZ`	`XYZ##`
`.`	
`X YZ`	`X#YZ#`

* The character # represents a blank space.

ANYDTDTE*w*. Informat

Reads and extracts the date value from various date, time, and datetime forms.

Category: Date and Time

Syntax

ANYDTDTE*w*.

Syntax Description

w
specifies the width of the input field.
Default: 9
Range: 5–32

Details

The ANYDTDTE informat reads input data that corresponds to any of the following informats or date, time, or datetime forms and extracts the date part from the derived value.

Informat or Form of Input	Example Data	Informat or Form of Input	Example Data
DATE	01JAN09	MONYY	JAN09
	01JAN2009		JAN2009
DATETIME	01JAN09 14:30:08	TIME	14:30
	01JAN2009 14:30:08.5		14:30:08.05
DDMMYY	010109	YMDDTTM	08-03-16 11:23
	01012009		

Informat or Form of Input	Example Data	Informat or Form of Input	Example Data
JULIAN	09001	YYMMDD	090101
	2009001		20090101
MDYAMPM	09-15-08 3:53 pm	YYQ	09Q1
			2009Q1
MMDDYY	010109	YY<YY>*x*MM*	09/01
	01012009		2009-01
MM*x*YY<YY>*	01/09	*month-day-year*	January 1, 2009
	01-2009		

* *x* is a special character that separates the month from the year

If the input value is a time-only value, then SAS assumes a date of 01JAN1960.

It is possible for input data such as 01-02-03 or 01-02 to be ambiguous with respect to the month, day, and year. In this case, the DATESTYLE system option indicates the order of the month, day, and year.

Comparisons

The ANYDTDTE informat extracts the date part from the derived value. The ANYDTDTM informat extracts the datetime part. The ANYDTTME informat extracts the time part.

Examples

```
input dateinfo anydtdte21.;
```

Data Line	Informat Form	Results	Formatted with the DATE*w*. Format
`----+----1----+----2`			
01JAN09	DATE	**17898**	**01JAN09**
01JAN2009 14:30:08.5	DATETIME	**17898**	**01JAN09**
01012009	DDMMYY	**17898**	**01JAN09**
2009001	JULIAN	**17898**	**01JAN09**
01/01/09	MMDDYY	**17898**	**01JAN09**
JAN2009	MONYY	**17898**	**01JAN09**
14:30	TIME	**0**	**01JAN60**
20090101	YYMMDD	**17898**	**01JAN09**
09q1	YYQ	**17898**	**01JAN09**
January 1, 2009	none	**17898**	**01JAN09**

See Also

Informats:

"ANYDTDTM*w*. Informat" on page 1301
"ANYDTTME*w*. Informat" on page 1304
"DATE*w*. Informat" on page 1322
"DATETIME*w*. Informat" on page 1324
"DDMMYY*w*. Informat" on page 1326
"JULIAN*w*. Informat" on page 1346
"MDYAMPM*w.d* Informat" on page 1347
"MMDDYY*w*. Informat" on page 1349
"MONYY*w*. Informat" on page 1351
"TIME*w*. Informat" on page 1393
"YMDDTTM*w.d* Informat" on page 1408
"YYMMDD*w*. Informat" on page 1410
"YYQ*w*. Informat" on page 1413

ANYDTDTM*w*. Informat

Reads and extracts datetime values from various date, time, and datetime forms.

Category: Date and Time

Syntax

ANYDTDTM*w*.

Syntax Description

w
specifies the width of the input field.
Default: 19
Range: 1–32

Details

The ANYDTDTM informat reads data that is in the form of any of the following informats or date/time forms, and extracts the datetime part from the derived value:

Informat or Form of Input	Example Data
DATE	01JAN09
	01JAN2009
DATETIME	01JAN09 14:30:08
	01JAN2009 14:30:08.5
DDMM<YY>YY	010109
	01012009

Informat or Form of Input	Example Data
JULIAN	09001
	2009001
MMDD<YY>YY*	010109
	01012009
MM*x*<YY>YY*	01/09
	01-2009
MDYAMPM**	01/01/09 02:30:08 AM
	01/01/2009 02:30:08 AM
MONYY	JAN09
	JAN2009
TIME	14.30
	14:30:08.05
<YY>YYMMDD	090101
	20090101
<YY>YYQ	09Q1
	2009Q1
<YY>YY*x*MM*	09/01
	2009/01
month-day-year	January 1, 2009

* *x* is a special character that separates the month from the year. <YY> indicates the century is optional.

** If AM | PM is not present and the month and day values are ambiguous, the value for the DATESYTLE= system option is used to determine the order.

If the input value is a time-only value, then SAS assumes a date of 01JAN1960. If the input value is a date-only value , then SAS assumes a time of 12:00 midnight. Input time values must include hours and minutes. If any part of a date in the input value is missing in the input value, or if the hour and minutes in a time value are missing or out of range, then the value read is a SAS missing value.

The input values for the preceding informats are mutually exclusive except for MMDDYY, DDMMYY, or YYMMDD when two-digit years are used. It is possible for input data such as 01-02-03 or 01-02 to be ambiguous with respect to the month, day, and year. In this case, the DATESTYLE system option indicates the order of the month, day, and year.

The ANYDTTME informat uses the following rules when reading colons and periods in time values:

Use of Colons and Periods	Example
a single colon in the value *h*:*m* indicates hours and minutes	14:30
two colons in the value h:m:s indicate hours, minutes, and seconds	14:30:08

Use of Colons and Periods	Example
a single period in the value m:s.ff, where ff is a fraction of a second, indicates that the number preceding the period is the number of seconds	2:39.66
multiple periods in the value indicate that the period is a delimiter for dates and the value is not a time value.	12.25.2009

Comparisons

The ANYDTDTE informat extracts the date part from the derived value. The ANYDTDTM informat extracts the datetime part. The ANYDTTME informat extracts the time part.

Examples

```
input dateinfo anydtdtm21.;
```

Data Line	Informat or Form of Data	Result	Formatted with DATETIME*w.d* Format
-+-1-+-2			
01JAN2009	DATE	**1546387200**	**01JAN09:00:00:00**
01JAN2009 14:30:08.5	DATETIME	**1546439408.5**	**01JAN09:14:30:09**
01012009	DDMMYY	**1546387200**	**01JAN09:00:00:00**
2009001	JULIAN	**1546387200**	**01JAN09:00:00:00**
01/01/09	MMDDYY	**1546387200**	**01JAN09:00:00:00**
01-09	MMxYY	**1546387200**	**01JAN09:00:00:00**
JAN2009	MONYY	**1546387200**	**01JAN09:00:00:00**
14:30	TIME	**52200**	**01JAN60:14:30:00**
20090101	YYMMDD	**1546387200**	**01JAN09:00:00:00**
09Q1	YYQ	**1546387200**	**01JAN09:00:00:00**
January 1, 2009	*month-day-year*	**1546387200**	**01JAN09:00:00:00**

See Also

Informats:

"ANYDTDTE*w*. Informat" on page 1299
"ANYDTTME*w*. Informat" on page 1304
"DATE*w*. Informat" on page 1322
"DATETIME*w*. Informat" on page 1324

"DDMMYY*w*. Informat" on page 1326
"JULIAN*w*. Informat" on page 1346
"MMDDYY*w*. Informat" on page 1349
"MONYY*w*. Informat" on page 1351
"TIME*w*. Informat" on page 1393
"YYMMDD*w*. Informat" on page 1410
"YYQ*w*. Informat" on page 1413

ANYDTTME*w*. Informat

Reads and extracts time values from various date, time, and datetime forms.

Category: Date and Time

Syntax

ANYDTTME*w*.

Syntax Description

w
specifies the width of the input field.
Default: 8
Range: 1-32

Details

The ANYDTTME informat reads input data that corresponds to any of the following informats or forms.

Informat or Form of Input	Example Data	Informat or Form of Input	Example Data
DATE	01JAN09	MONYY	JAN09
	01JAN2009		JAN2009
DATETIME	01JAN09 14:30:08	YYMMDD	090101
	01JAN2009 14:30:08.5		20090101
DDMMYY	010109	YYQ	09Q1
	01012009		2009Q1
JULIAN	09001	YYQ	09Q1
	2009001		2009Q1
MMDDYY	010109	*month-day-year*	January 1, 2009
	01012009		2009-01

If the input value is a time-only value, then SAS assumes a date of 01JAN1960. If the input value is a date value only, then SAS assumes a time of 12:00 midnight.

It is possible for input data such as 01-02-03 or 01-02 to be ambiguous with respect to the month, day, and year. In this case, the DATESTYLE system option indicates the order of the month, day, and year.

The ANYDTTME informat uses the following rules when reading colons and periods in time values:

Use of Colons and Periods	Example
a single colon in the value *h*:*m* indicates hours and minutes	14:30
two colons in the value h:m:s indicate hours, minutes, and seconds	14:30:08
a single period in the value m:s.ff, where ff is a fraction of a second, indicates that the number preceding the period is the number of seconds	2:39.66
multiple periods in the value indicate that the period is a delimiter for dates and the value is not a time value.	12.25.2009

Comparisons

The ANYDTDTE informat extracts the date part from the derived value. The ANYDTDTM informat extracts the datetime part. The ANYDTTME informat extracts the time part.

Examples

```
input dateinfo anydttme21.;
```

Data Line	Informat	Results	Formatted with the TIME*w.d* Format
----+----1----+----2			
01JAN09	DATE	**0**	**00:00:00**
01JAN2009 14:30:08.5	DATETIME	**52208.5**	**14:30:09**
010109	DDMMYY	**0**	**00:00:00**
2009001	JULIAN	**0**	**00:00:00**
01012009	MMDDYY	**0**	**00:00:00**
JAN2009	MONYY	**0**	**00:00:00**
14:30:08.5	TIME	**52208.5**	**14:30:09**
20090101	YYMMDD	**0**	**00:00:00**

Data Line	Informat	Results	Formatted with the TIME*w.d* Format
`09Q1`	YYQ	`0`	`00:00:00`
`January 1, 2009`	*month-day-year*	`0`	`00:00:00`

See Also

Informats:

"ANYDTDTE*w*. Informat" on page 1299
"ANYDTDTM*w*. Informat" on page 1301
"DATE*w*. Informat" on page 1322
"DATETIME*w*. Informat" on page 1324
"DDMMYY*w*. Informat" on page 1326
"JULIAN*w*. Informat" on page 1346
"MMDDYY*w*. Informat" on page 1349
"MONYY*w*. Informat" on page 1351
"TIME*w*. Informat" on page 1393
"YYMMDD*w*. Informat" on page 1410
"YYQ*w*. Informat" on page 1413

B8601DA*w*. Informat

Reads date values that are specified in the ISO 8601 base notation *yyyymmdd*.

Category: Date and Time
ISO 8601
Alignment: left
Alias: ND8601DA
Time Zone Informat: No
ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

B8601DA*w*.

Syntax Description

w
specifies the width of the input field.

Default: 10

Requirement: The width of the output field must be 10.

Details

The B8602DA informat reads date values that are specified in the ISO 8601 basic date notation *yyyymmdd*:

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 31

Examples

```
input @1 bda b8601da.;
```

Data Line	Results
`------+------1`	
`20080915`	`17790`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

B8601DN*w*. Informat

Reads date values that are specified the ISO 8601 basic notation *yyyymmdd* and returns SAS datetime values where the time portion of the value is 000000.

Category: Date and Time
ISO 8601
Alignment: left
Alias: ND8601DN
Time Zone Informat: No
ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

B8601DN*w*.

Syntax Description

w
specifies the width of the input field.
Default: 10
Requirement: The width of the input field must be 10.

Details

The B8602DN informat reads date values that are specified in the ISO 8601 basic date notation *yyyymmdd* and returns the date in a SAS datetime value:

yyyy	is a four-digit year, such as 2008.
mm	is a two-digit month (zero padded) between 01 and 12.
dd	is a two-digit day of the month (zero padded) between 01 and 31.

Examples

```
input @1 bdn b8601dn.;
```

Data Line	Results
`----+----1`	
`20080915`	`1537056000`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

B8601DT*w.d* Informat

Reads datetime values that are specified in the ISO 8601 basic notation *yyyymmdd*T*hhmmssffffff*.

Category: Date and Time
ISO 8601
Alignment: left
Alias: ND8601DT
Time Zone Format: No
ISO 8601 Element: 5.4.1 Complete representation

Syntax

B8601DT*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 19
Range: 19–26

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.
Default: 0
Range: 0–6

Details

The B8602DT informat reads datetime values that are specified in the ISO 8601 basic datetime notation *yyyymmdd*T*hhmmssffffff*:

yyyy
is a four-digit year, such as 2008

mm
is a two-digit month (zero padded) between 01 and 12

dd
is a two-digit day of the month (zero padded) between 01 and 31

hh
is a two-digit hour (zero padded), between 00 - 23

mm
is a two-digit minute (zero padded), between 00 - 59

ss
is a two-digit second (zero padded), between 00 - 59

ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 - 9

Examples

```
input @1 bdt b8601dt;
```

Data Line	Results
`------+------1`	
`20080915T155300`	`1537113180`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

B8601DZ*w.d* Informat

Reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime basic notation *yyyymmdd*T*hhmmss*+|-*hhmm* or *yyyymmdd*T*hhmmssffffff*Z.

Category: Date and Time
ISO 8601

Alignment: left

Alias: ND8601DZ

Time Zone Informat: Yes

ISO 8601 Element: 5.4.1 Complete representation

Syntax

B8601DZ*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 26

Range: 20–35

d
specifies the number of digits to the right of the seconds value, which represents a fraction of a second. This argument is optional.

Default: 0

Range: 0–6

Details

UTC values specify a time and a time zone based on the zero meridian in Greenwich, England. The B8602DZ informat reads datetime values that are specified in one of the following ISO 8601 basic datetime notations:

*yyyymmdd*T*hhmmss*+|–*hhmm*

*yyyymmdd*T*hhmmssffffff*Z

where

yyyy
is a four-digit year, such as 2008

mm
is a two-digit month (zero padded) between 01 and 12

dd
is a two-digit day of the month (zero padded) between 01 and 31

hh
is a two-digit hour (zero padded), between 00 and 24

mm
is a two-digit minute (zero padded), between 00 and 59

ss
is a two-digit second (zero padded), between 00 and 59

.ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 and 9.

Z
indicates that the time is for zero meridian (Greenwich, England) or UTC time.

+|–*hhmm*
is an hour and minute signed offset from zero meridian time. Note that the offset must be +|–*hhmm* (that is, + or – and four characters).

Use + for time zones east of the zero meridian and use – for time zones west of the zero meridian. For example, +0200 indicates a two hour time difference to the east of the zero meridian, and –0600 indicates a six hour time differences to the west of the zero meridian.

Restriction: The shorter form +|–*hh* is not supported.

Examples

```
input @1 bdz b8601dz.;
```

Data Line	Results
`----+----1`	
`20080915T155300+0500`	`1537095180`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

B8601TM*w.d* Informat

Reads time values that are specified in the ISO 8601 basic notation *hhmmssffffff*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: ND8601TM

Time Zone Informat: No

ISO 8601 Element: 5.3.1.1 Complete representation and 5,3,1,3 Representation of decimal fractions

Syntax

B8601TM*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 8

Range: 6–15

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0–6

Details

The B8601TM informat reads time values that are specified in the ISO 8601 basic time notation *hhmmssffffff*:

hh
is a two-digit hour (zero padded), between 00 and 23

mm
is a two-digit minute (zero padded), between 00 and 59

ss
is a two-digit second (zero padded), between 00 and 59

.ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 - 9

Examples

```
input @1 btm b8601tm;
```

Data Line	Results
`----+----1`	
`155300`	`57180`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

B8601TZ*w.d* Informat

Reads time values that are specified in the ISO 8601 basic time notation *hhmmssfffff*+|-*hhmm* or *hhmmssfffffZ*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: ND8601TZ

Time Zone Informat: Yes

ISO 8601 Element: 5.3.1.1 Complete representation

Syntax

B8601TZ*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 14

Range: 9–20

d
(optional) specifies the number of digits to the right of the decimal point in the seconds value.

Default: 0

Range: 0 - 6

Details

UTC time values specify a time and a time zone based on the zero meridian in Greenwich, England. The B8602TZ informat reads time values that are specified in the following ISO 8601 basic time notations:

hhmmssffffff+|–*hhmm*

*hhmmssffffff*Z

where

hh
is a two-digit hour (zero padded), between 00 and 23

mm
is a two-digit minute (zero padded), between 00 and 59

ss
is a two-digit second (zero padded), between 00 and 59

ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 and 9

Z
indicates that the time is for zero meridian (Greenwich, England) or UTC time

+|–*hh:mm*
is an hour and minute signed offset from zero meridian time. Note that the offset must be +|–*hhmm* (that is, + or – and four characters).

Use + for time zones east of the zero meridian and use – for time zones west of the zero meridian. For example, +0200 indicates a two hour time difference to the east of the zero meridian, and –0600 indicates a six hour time differences to the west of the zero meridian.

Restriction: The shorter form +|–*hh* is not supported.

When SAS reads a UTC time by using the B8601TZ informat and the adjusted time is greater than 240000 or less than 000000, SAS adjusts the time so that it represents a time between 000000 and 240000. For example, if SAS reads the UTC time 234344-0500 using the B8601TZ informat, SAS adds five hours to the time so that the value is 284344, and then makes the time adjustment. The value stored represents the time 044344+0000.

Examples

```
input @1 btz b8601tz.;
```

Data Line	Results
`----+----1`	
`202401-0500`	`5041`
`202401z`	`73441`
`202401+0000`	`73441`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

BINARY*w.d* Informat

Converts positive binary values to integers.

Category: Numeric

Syntax

BINARY*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 8

Range: 1–64

d
specifies the power of 10 by which to divide the value. SAS uses the *d* value even if the data contain decimal points. This argument is optional.

Range: 0–31

Details

Use only the character digits 1 and 0 in the input, with no embedded blanks. BINARY*w.d* ignores leading and trailing blanks.

BINARY*w.d* cannot read negative values. It treats all input values as positive (unsigned).

Examples

```
input @1 value binary8.1;
```

Data Line	Results
`----+----1----+`	
`00001111`	`1.5`

BITS*w.d* Informat

Extracts bits.

Category: Numeric

Syntax

BITS*w.d*

Syntax Description

w
specifies the number of bits to read.

Default: 1

Range: 1–64

d
specifies the zero-based offset.

Range: 0–63

Details

The BITS*w.d* informat extracts particular bits from an input stream and assigns the numeric equivalent of the extracted bit string to a variable. Together, the *w* and *d* values specify the location of the string you want to read.

This informat is useful for extracting data from system records that have many pieces of information packed into single bytes.

Examples

```
input @1 value bits4.1;
```

Data Line	Results*
`----+----1----+`	
`B`	`8`

*The EBCDIC binary code for a capital B is 11000010, and the ASCII binary code is 01000010.

The input pointer moves to column 2 (*d*=1). Then the INPUT statement reads four bits (*w*=4) which is the bit string 1000 and stores the numeric value 8, which is equivalent to this binary combination.

BZ*w.d* Informat

Converts blanks to 0s.

Category: Numeric

Syntax

BZ*w*.*d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–32

d
specifies the power of 10 by which to divide the value. If the data contain decimal points, the *d* value is ignored. This argument is optional.

Range: 0–31

Details

The BZ*w*.*d* informat reads numeric values, converts any trailing or embedded blanks to 0s, and ignores leading blanks.

The BZ*w*.*d* informat can read numeric values that are located anywhere in the field. Blanks can precede or follow the numeric value, and a minus sign must precede negative values. The BZ*w*.*d* informat ignores blanks between a minus sign and a numeric value in an input field.

The BZ*w*.*d* informat interprets a single period in a field as a 0. The informat interprets multiple periods or other nonnumeric characters in a field as a missing value.

To use BZ*w*.*d* in a DATA step with list input, change the delimiter for list input with the DLM= or DLMSTR= option in the INFILE statement. By default, SAS interprets blanks between values in the data line as delimiters rather than 0s.

Comparisons

The BZ*w*.*d* informat converts trailing or embedded blanks to 0s. If you do not want to convert trailing blanks to 0s (for example, when reading values in E-notation), use either the *w*.*d* informat or the E*w*.*d* informat instead.

Examples

```
input @1 x bz4.;
```

Data Line	Result
`----+----1`	
`34`	`3400`
`-2`	`-200`
`-2 1`	`-201`

CB*w.d* Informat

Reads standard numeric values from column-binary files.

Category: Column Binary

Syntax

CB*w.d*

Syntax Description

w
specifies the width of the input field.

Range: 1–32

d
specifies the power of 10 by which to divide the value. SAS uses the *d* value even if the data contain decimal points. This argument is optional.

Details

Column-binary data storage compresses data so that more than 80 items of data can be stored on a single "virtual" punch card.

The CB*w.d* informat reads standard numeric values from column-binary files and translates the data into standard binary format.

SAS first stores each column of column-binary data you read with CB*w.d* in two bytes and ignores the two high-order bits of each byte. If the punch codes are valid, then SAS stores the equivalent numeric value in the variable that you specify. If the combinations are not valid, then SAS assigns the variable a missing value and sets the automatic variable _ERROR_ to 1.

Examples

```
input @1 x cb8.;
```

Data Line*	Results
`----+----1`	
0009	**9**

* The data line is a hexadecimal representation of the column binary. The "virtual" punch card column for the example data has row 9 punched. The binary representation is 0000 0000 0000 1001.

See Also

Informats:

"$CB*w*. Informat" on page 1284

"PUNCH.*d* Informat" on page 1366

"ROW*w.d* Informat" on page 1371

"How to Read Column-Binary Data" in *SAS Language Reference: Concepts*

COMMA*w.d* Informat

Removes embedded characters.

Category: Numeric

Alias: DOLLAR*w.d*

Syntax

COMMA*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 1

Range: 1–32

d

specifies the power of 10 by which to divide the value. If the data contain decimal points, the *d* value is ignored. This argument is optional.

Range: 0–31

Details

The COMMA*w.d* informat reads numeric values and removes embedded commas, blanks, dollar signs, percent signs, dashes, and close parentheses from the input data. The COMMA*w.d* informat converts an open parenthesis at the beginning of a field to a minus sign.

Comparisons

The COMMA*w.d* informat operates like the COMMAX*w.d* informat, but it reverses the roles of the decimal point and the comma. This convention is common in European countries.

Examples

```
input @1 x comma10.;
```

Data Line	Results
`----+----1----+`	
`$1,000,000`	`1000000`
`(500)`	`-500`

COMMAX*w.d* Informat

Removes embedded periods, blanks, dollar signs, percent signs, dashes, and closing parenthesis from the input data. An open parenthesis at the beginning of a field is converted to a minus sign. The COMMAX informat reverses the roles of the decimal point and the comma.

Category: Numeric

Alias: DOLLARX*w.d*

Syntax

COMMAX*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 1

Range: 1–32

d

specifies the power of 10 by which to divide the value. If the data contain a comma, which represents a decimal point, the *d* value is ignored. This argument is optional.

Range: 0–31

Details

The COMMAX*w*.*d* informat reads numeric values and removes embedded periods, blanks, dollar signs, percent signs, dashes, and close parentheses from the input data. The COMMAX*w*.*d* informat converts an open parenthesis at the beginning of a field to a minus sign.

Comparisons

The COMMAX*w*.*d* informat operates like the COMMA*w*.*d* informat, but it reverses the roles of the decimal point and the comma. This convention is common in European countries.

Examples

```
input @1 x commax10.;
```

Data Line	Results
----+----1----+	
$1.000.000	1000000
1.234,56	1234.56
(500)	-500

DATE*w*. Informat

Reads date values in the form *ddmmmyy* or *ddmmmyyyy*.

Category: Date and Time

Syntax

DATE*w*.

Syntax Description

w
specifies the width of the input field.

Default: 7

Range: 7–32

Tip: Use a width of 9 to read a 4–digit year.

Details

The date values must be in the form *ddmmmyy* or *ddmmmyyyy*, where

dd
is an integer from 01 through 31 that represents the day of the month.

mmm
is the first three letters of the month name.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

You can separate the year, month, and day values by blanks or by special characters. Make sure the width of the input field allows space for blanks and special characters.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input calendar_date date11.;
```

Data Line	Results
`----+----1----+`	
`16mar99`	`14319`
`16 mar 99`	`14319`
`16-mar-1999`	`14319`

See Also

Format:
"DATE*w*. Format" on page 151
Function:
"DATE Function" on page 634
System Option:
"YEARCUTOFF= System Option" on page 2058

DATETIME*w*. Informat

Reads datetime values in the form *ddmmmyy hh:mm:ss.ss* or *ddmmmyyyy hh:mm:ss.ss*.

Category: Date and Time

Syntax

DATETIME*w*.

Syntax Description

w
specifies the width of the input field.
Default: 18
Range: 13–40

Details

The datetime values must be in the following form: *ddmmmyy* or *ddmmmyyyy*, followed by a blank or special character, followed by *hh:mm:ss.ss* (the time). In the date,

dd
is an integer from 01 through 31 that represents the day of the month.

mmm
is the first three letters of the month name.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

In the time,

hh
is the number of hours ranging from 00 through 23.

mm
is the number of minutes ranging from 00 through 59.

ss.ss
is the number of seconds ranging from 00 through 59 with the fraction of a second following the decimal point.

DATETIME*w*. requires values for both the date and the time. However, the *ss.ss* portion is optional.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Note: SAS can read time values with AM and PM in them. △

Comparisons

The DATETIME*w.d* informat reads datetime values with optional separators in the form *dd-mmm-yy<yy> hh:mm:ss.ss* AM|PM, and the date and time can be separated by a special character.

The MDYAMPM*w.d* in format reads datetime values with optional separators in the form *mm-dd-yy<yy> hh:mm:ss.ss* AM | PM, and requires a space between the date and the time.

The YMDDTTM*w.d* informat reads datetime values with required separators in the form *<yy>yy-mm-dd*/*hh:mm:ss.ss*.

Examples

```
input date_and_time datetime20.;
```

Data Line	Results
`----+----1----+----2`	
`16mar08:11:23:07.4`	`1521285787.4`
`16mar2008/11:23:07.4`	`1521285787.4`
`16mar2008/11:23 PM`	`1521328980`

See Also

Formats:

- "DATE*w*. Format" on page 151
- "DATETIME*w.d* Format" on page 154
- "TIME*w.d* Format" on page 245

Function:

- "DATETIME Function" on page 637

Informats:

- "DATE*w*. Informat" on page 1322
- "MDYAMPM*w.d* Informat" on page 1347
- "TIME*w*. Informat" on page 1393
- "YMDDTTM*w.d* Informat" on page 1408

System Option:

- "YEARCUTOFF= System Option" on page 2058

See the discussion on using SAS date and time values in *SAS Language Reference: Concepts*

DDMMYY*w*. Informat

Reads date values in the form *ddmmyy<yy>* or *dd-mm-yy<yy>*, where a special character, such as a hyphen (-), period (.), or slash (/), separates the day, month, and year; the year can be either 2 or 4 digits.

Category: Date and Time

Syntax

DDMMYY*w.*

Syntax Description

w
specifies the width of the input field.

Default: 6

Range: 6–32

Details

The date values must be in the form *ddmmyy<yy>* or *ddxmmxyy<yy>*, where

dd
is an integer from 01 through 31 that represents the day of the month.

mm
is an integer from 01 through 12 that represents the month.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

x
is a separators that can be any special character or a blank.:

If you use separators, place them between all the values. Blanks can also be placed before and after the date. Make sure the width of the input field allows space for blanks and special characters.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input calendar_date ddmmyy10.;
```

Data Line	Results
`----+----1----+`	
`160308`	`17607`
`16/03/08`	`17607`
`16-03-2008`	`17607`
`16 03 2008`	`17607`

See Also

Formats:

"DATE*w*. Format" on page 151

"DDMMYY*w*. Format" on page 157

"MMDDYY*w*. Format" on page 196

"YYMMDD*w*. Format" on page 273

Function:

"MDY Function" on page 924

Informats:

"DATE*w*. Informat" on page 1322

"MMDDYY*w*. Informat" on page 1349

"YYMMDD*w*. Informat" on page 1410

System Option:

"YEARCUTOFF= System Option" on page 2058

E*w.d* Informat

Reads numeric values that are stored in scientific notation and double-precision scientific notation.

Category: Numeric

See: E*w.d* Informat in the documentation for your operating environment.

Syntax

E*w.d*

Syntax Description

w
specifies the width of the field that contains the numeric value.

Default: 12

Range: 1–32

d
specifies the number of digits to the right of the decimal point in the numeric value. If the data contain decimal points, the *d* value is ignored. This argument is optional.

Range: 0–31

Comparisons

The E*w*.*d* informat is not used extensively because the SAS informat for standard numeric data, the *w*.*d* informat, can read numbers in scientific notation. Use E*w*.*d* to permit only scientific notation in your input data.

Examples

```
input @1 x e7.;
```

Data Line	Results
`----+----1----+`	
`1.257E3`	`1257`
`12d3`	`12000`

E8601DA*w*. Informat

Reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601DA

Time Zone Informat: No

ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

E8601DA*w*.

Syntax Description

w
specifies the width of the input field.

Default: 10

Requirement: The width of the input field must be 10.

Details

The E8601DA informat reads date values that are specified in the ISO 8601 extended date notation *yyyy-mm-dd*:

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 12

Examples

```
input eda e8601da.;
```

Data Line	Results
`------+------1`	
`2008-09-15`	`17790`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

E8601DN*w*. Informat

Reads date values that are specified in the ISO 8601 extended notation *yyyy-mm-dd* and returns SAS datetime values where the time portion of the value is 000000.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601DN

Time Zone Informat: No

ISO 8601 Element: 5.2.1.1 Complete representation

Syntax

E8601DN*w*.

Syntax Description

w
specifies the width of the input field.

Default: 10

Requirement: The width of the input field must be 10.

Details

The E8601DN informat reads date values that are specified in the ISO 8601 extended date notation is *yyyy-mm-dd* and returns the date in a SAS datetime value:

yyyy	is a four-digit year, such as 2008
mm	is a two-digit month (zero padded) between 01 and 12
dd	is a two-digit day of the month (zero padded) between 01 and 31

Examples

```
input edn is8601dn.;
```

Data Line	Results
`------+------1`	
`2008-09-15`	`1537056000`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

E8601DT*w.d* Informat

Reads datetime values that are specified in the ISO 8601 extended notation *yyyy-mm-dd*T*hh:mm:ss.ffffff*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601DT

Time Zone Informat: No

ISO 8601 Element: 5.4.1 Complete representation

Syntax

E8601DT*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 19

Range: 19–26

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0–6

Details

The E8601DT informat reads datetime values that are specified in the ISO 8601 extended datetime notation*yyyy-mm-dd*T*hh:mm:ss.ffffff*:

yyyy
is a four-digit year, such as 2008.

mm
is a two-digit month (zero padded) between 01 and 12.

dd
is a two-digit day of the month (zero padded) between 01 and 31.

hh
is a two-digit hour (zero padded), between 00 and 23.

mm
is a two-digit minute (zero padded), between 00 and 59.

ss
is a two-digit second (zero padded), between 00 and 59.

.ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 and 9.

Examples

```
input @1 edt e8601dt.;
```

Data Line	Results
`------+------1------+------2------+------3`	
`2008-09-15T15:53:00`	`1537113180`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

E8601DZ*w.d* Informat

Reads datetime values that are specified in the Coordinated Universal Time (UTC) time scale using ISO 8601 datetime extended notation *yyyy-mm-dd*T*hh:mm:ss*+|-*hh:mm.fffff* or *yyyy-mm-dd*T*hh:mm:ss.fffff*Z.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601DZ

Time Zone Informat: Yes

ISO 8601 Element: 5.4.1 Complete representation

Syntax

E8601DZ*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 26

Range: 20–35

d
specifies the number of digits to the right of the decimal point in the value for the lowest order component. This argument is optional.

Default: 0

Range: 0–6

Details

UTC values specify a time and a time zone based on the zero meridian in Greenwich, England. The E8602DZ informat reads datetime values contain UTC time offsets and that are specified in one of the following ISO 8601 extended datetime notations:

*yyyy-mm-dd*T*hh:mm:ss.ffffff*+|–*hh:mm*

*yyyy-mm-dd*T*hh:mm:ss.ffffff*Z

where

yyyy
is a four-digit year, such as 2008

mm
is a two-digit month (zero padded) between 01 and 12

dd
is a two-digit day of the month (zero padded) between 01 and 31

hh
is a two-digit hour (zero padded), between 00 and 24

mm
is a two-digit minute (zero padded), between 00 and 59

ss
is a two-digit second (zero padded), between 00 and 59

.ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 and 9

Z
indicates that the time is UTC time at the zero meridian (Greenwich, England)

+|–*hh:mm*
is an hour and minute signed offset from zero meridian time. Note that the offset must be +|–*hh:mm* (that is, + or – and five characters).

Use + for time zones east of the zero meridian and use – for time zones west of the zero meridian. For example, +02:00 indicates a two hour time difference to the east of the zero meridian, and –06:00 indicates a six hour time differences to the west of the zero meridian.

Restriction: The shorter form +|–*hh* is not supported.

Examples

Input Statement	Data Line	Results
	`------+------1------+------2------+`	
`input edz e8601dz.;`	`2008-09-15T15:53:00Z`	`1537113180`
`input edz  e8601dz28.2;`	`2008-09-15T15:53:00+03:00`	`1537102380`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

E8601LZ*w.d* Informat

Reads Coordinated Universal Time (UTC) values that are specified in the ISO 8601 extended notation *hh*:*mm*:*ss*+|-*hh*:*mm*.*fffff* or *hh*:*mm*:*ss*.*fffff*Z and converts them to the local time.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601LZ

Time Zone Informat: Yes

ISO 8601 Element: 5.3.1.1 Complete representation

Syntax

E8601LZ*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 14

Range: 9–20

Requirement: To read a time with the Z time zone indicator, the width of the input field must be 9 if data follows on the same line of data.

d
specifies the number of digits to the right of the decimal point in the value for the lowest order component. This argument is optional.

Default: 0

Range: 0–6

Details

UTC values specify a time and a time zone based on the zero meridian in Greenwich, England. The E8602LZ informat reads UTC time values that are specified in one of the following ISO 8601 extended time notations and return a SAS time value for the local time:

hh:mm:ss.ffffff+|–00:00

*hh:mm:ss.ffffff*Z

where

hh
is a two-digit hour (zero padded), between 00 and 23

mm
is a two-digit minute (zero padded), between 00 and 59

ss
is a two-digit second (zero padded), between 00 and 59

.ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 and 9

Z
indicate zero meridian or UTC time.

+|–*hh:mm*
is an hour and minute signed offset from zero meridian or UTC time. Note that the offset must be +|–*hh:mm* (that is, + or – and five characters).

Use the + for time zones east of the zero meridian and use the – for time zones west of the zero meridian.

Restriction: The shorter form +|–*hh* is not supported.

When SAS reads a UTC time by using the E8601LZ informat and the adjusted time is greater than 24:00:00 or less than 00:00:00, SAS adjusts the time so that it represents a time between 00:00:00 and 24:00:00. For example, if SAS reads the UTC time 23:43:44-05:00 using the E8601LZ informat, SAS adds 5 hours to the time so that the value is 28:43:44, and then makes the time adjustment. The value stored represents the time 04:43:44+00:00.

Examples

```
input elz e8601lz.;
```

Data Line	Results
`09:13:21+02:00`	`26001`
`23:43:44Z`	`85424`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

E8601TM*w.d* Informat

Reads time values that are specified in the ISO 8601 extended notation *hh:mm:ss.ffffff*.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601TM

Time Zone Informat: No

ISO 8601 Element: 5.3.1.1 Complete representation and 5.3.1.3 Representation of decimal fractions

Syntax

E8601TM*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 8

Range: 8–15

d
specifies the number of digits to the right of the decimal point in the seconds value. This argument is optional.

Default: 0

Range: 0–6

Details

The E8601TM informat reads time values that are specified in the following ISO 8601 extended time notation:

hh:mm:ss.ffffff

hh
is a two-digit hour (zero padded), between 00 and 23

mm
is a two-digit minute (zero padded), between 00 and 59

ss
is a two-digit second (zero padded), between 00 and 59

.ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 and 9

Examples

```
input @1 etm e8601tm.
```

Data Line	Results
`------+------1`	
`15:53:00`	`57180`

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

E8601TZ*w.d* Informat

Reads time values that are specified in the ISO 8601 extended time notation *hh*:*mm*:*ss*+|-*hh*:*mm*.*ffffff* or *hh*:*mm*:*ss*Z.

Category: Date and Time
ISO 8601

Alignment: left

Alias: IS8601TZ

Time Zone Informat: Yes

ISO 8601 Element: 5.3.1.1 Complete representation

Syntax

E8601TZ*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 14

Range: 9–20

Requirement: To read a time with the Z time zone indicator, the width of the input field must be 9 if data follows on the same line of data.

d
(optional) specifies the number of digits to the right of the decimal point in the value for the lowest order component.

Default: 0

Range: 0–6

Details

UTC time values specify a time and a time zone based on the zero meridian in Greenwich, England. The E8602TZ informat reads UTC time values that are specified in one of the following ISO 8601 extended notations:

hh:mm:ss+|–*hh:mm.ffffff*

hh:mm:ss

The following list explains the UTC time variables:

hh
is a two-digit hour (zero padded), between 00 and 23

mm
is a two-digit minute (zero padded), between 00 and 59

ss
is a two-digit second (zero padded), between 00 and 59

.ffffff
are optional fractional seconds, with a precision of up to six digits, where each digit is between 0 - 9

Z
indicate zero meridian or UTC time

+|–*hh:mm*
is an hour and minute signed offset from zero meridian. Note that the offset must be +|–*hh:mm* (that is, + or – and five characters).

Use the + for time zones east of the zero meridian and use the – for time zones west of the zero meridian.

Restriction: The shorter form +|–*hh* is not supported.

When SAS reads a UTC time by using the E8601TZ informat and the adjusted time is greater than 24:00:00 or less than 00:00:00, SAS adjusts the time so that it represents a time between 00:00:00 and 24:00:00. For example, if SAS reads the UTC time 23:43:44-05:00 using the E8601TZ informat, SAS adds 5 hours to the time so that the value is 28:43:44, and then makes the time adjustment. The value stored represents the time 04:43:44+00:00.

Examples

```
input @1 etz e8601tz.;
```

Data Line	Results
------+------1------+------2	
23:43:44-05:00	17024
23:43:44Z	85424

See Also

"Reading Dates and Times Using the ISO 860 Basic and Extended Notations" on page 1269

FLOAT*w.d* Informat

Reads a native single-precision, floating-point value and divides it by 10 raised to the *d*th power.

Category: Numeric

Syntax

FLOAT*w.d*

Syntax Description

w
specifies the width of the input field.
Requirement: *w* must be 4.

d
specifies the power of 10 by which to divide the value. This argument is optional.

Details

The FLOAT*w.d* informat is useful in operating environments where a float value is not the same as a truncated double.

On the IBM mainframe systems, a four-byte floating-point number is the same as a truncated eight-byte floating-point number. However, in operating environments that use the IEEE floating-point standard, such as the IBM PC-based operating environments and most UNIX platforms, a four-byte floating-point number is not the same as a truncated double. Therefore, the RB4. informat does not produce the same results as FLOAT4. Floating-point representations other than IEEE might have this same characteristic. Values read with FLOAT4. typically come from some other external program that is running in your operating environment.

Comparisons

The following table compares the names of float notation in several programming languages:

Language	Float Notation
SAS	FLOAT4.
Fortran	REAL*4
C	float
IBM 370 ASM	E
PL/I	FLOAT BIN(21)

Examples

```
input x float4.;
```

Data Line*	Results
`----+----1----+----2`	
`3F800000`	`1`

* The data line is a hexadecimal representation of a binary number that is stored in IEEE form.

HEX*w*. Informat

Converts hexadecimal positive binary values to either integer (fixed-point) or real (floating-point) binary values.

Category: Numeric

See: HEX*w*. Informat in the documentation for your operating environment.

Syntax

HEX*w*.

Syntax Description

w

specifies the field width of the input value and also specifies whether the final value is fixed-point or floating-point.

Default: 8

Range: 1–16

Tip: If *w*<16, HEX*w*. converts the input value to positive integer binary values, treating all input values as positive (unsigned). If *w* is 16, HEX*w*. converts the input value to real binary (floating-point) values, including negative values.

Details

Note: Different operating environments store floating-point values in different ways. However, HEX16. reads hexadecimal representations of floating-point values with consistent results if the values are expressed in the same way that your operating environment stores them. △

The HEX*w*. informat ignores leading or trailing blanks.

Examples

```
input @1 x hex3. @5 y hex16.;
```

Data Line*	Results
`----+----1----+----2`	
`88F 4152000000000000`	`2191 5.125`

* The data line shows IBM mainframe hexadecimal data.

IB*w.d* Informat

Reads native integer binary (fixed-point) values, including negative values.

Category: Numeric

See: IB*w.d* Informat in the documentation for your operating environment.

Syntax

IB*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 4

Range: 1–8

d
specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

The IB*w.d* informat reads integer binary (fixed-point) values, including negative values represented in two's complement notation. IB*w.d* reads integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 1263. △

Comparisons

The IB*w.d* and PIB*w.d* informats are used to read native format integers. (Native format allows you to read and write values created in the same operating environment.) The IBR*w.d* and PIBR*w.d* informats are used to read little endian integers in any operating environment.

To view a table that shows the type of informat to use with big endian and little endian integers, see Table 5.1 on page 1264.

To view a table that compares integer binary notation in several programming languages, see Table 5.2 on page 1264.

Examples

You can use the INPUT statement and specify the IB informat. However, these examples use the informat with the INPUT function, where binary input values are described using a hexadecimal literal.

```
x=input('0080'x,ib2.);
y=input('8000'x,ib2.);
```

SAS Statement	Results on Big Endian Platforms	Results on Little Endian Platforms
`put x=;`	`128`	`-32768`
`put y=;`	`-32768`	`128`

See Also

Informat:

"IBR*w.d* Informat" on page 1343

IBR*w.d* Informat

Reads integer binary (fixed-point) values in Intel and DEC formats.

Category: Numeric

Syntax

IBR*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 4

Range: 1–8

d

specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

The IBR*w.d* informat reads integer binary (fixed-point) values, including negative values that are represented in two's complement notation. IBR*w.d* reads integer binary values that are generated by and for Intel and DEC platforms. Use IBR*w.d* to read integer binary data from Intel or DEC environments in other operating environments. The IBR*w.d* informat in SAS code allows for a portable implementation for reading the data in any operating environment.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 1263. △

Comparisons

The IB*w.d* and PIB*w.d* informats are used to read native format integers. (Native format allows you to read and write values that are created in the same operating environment.) The IBR*w.d* and PIBR*w.d* informats are used to read little endian integers in any operating environment.

On Intel and DEC operating environments, the IB*w.d* and IBR*w.d* informats are equivalent.

To view a table that shows the type of informat to use with big endian and little endian integers, see Table 5.1 on page 1264.

To view a table that compares integer binary notation in several programming languages, see Table 5.2 on page 1264.

Examples

You can use the INPUT statement and specify the IBR informat. However, in these examples we use the informat with the INPUT function, where binary input values are described using a hexadecimal literal.

```
x=input('0100'x,ibr2.);
y=input('0001'x,ibr2.);
```

SAS Statement	Results on Big Endian Platforms	Results on Little Endian Platforms
`put x=;`	`1`	`1`
`put y=;`	`256`	`256`

See Also

Informat:

"IB*w.d* Informat" on page 1341

IEEE*w.d* Informat

Reads an IEEE floating-point value and divides it by 10 raised to the *d* th power.

Category: Numeric

Syntax

IEEE*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 8

Range: 2–8

Tip: If *w* is 8, an IEEE double-precision, floating-point number is read. If *w* is 5, 6, or 7, an IEEE double-precision, floating-point number is read, which assumes truncation of the appropriate number of bytes. If *w* is 4, an IEEE single-precision, floating-point number is read. If *w* is 3, an IEEE single-precision, floating-point number is read, which assumes truncation of one byte.

d
specifies the power of 10 by which to divide the value.

Details

The IEEE*w.d* informat is useful in operating environments where IEEE is the floating-point representation that is used. In addition, you can use the IEEE*w.d* informat to read files that are created by programs on operating environments that use the IEEE floating-point representation.

Typically, programs generate IEEE values in single precision (4 bytes) or double precision (8 bytes). Truncation is performed by programs solely to save space on output files. Machine instructions require that the floating-point number be of one of the two lengths. The IEEE*w.d* informat allows other lengths, which enables you to read data from files that contain space-saving truncated data.

Examples

```
input test1 ieee4.;
input test2 ieee5.;
```

Data Line*	Results
`----+----1----+`	
`3F800000`	`1`
`3FF0000000`	`1`

* The data lines are hexadecimal representations of binary numbers that are stored in IEEE format.

The first INPUT statement reads the first data line, and the second INPUT statement reads the next data line.

JULIAN*w*. Informat

Reads Julian dates in the form *yyddd* or *yyyyddd*.

Category: Date and Time

Syntax

JULIAN*w*.

Syntax Description

w
specifies the width of the input field.

Default: 5

Range: 5–32

Details

The date values must be in the form *yyddd* or *yyyyddd*, where

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

dd or *ddd*
is an integer from 01 through 365 that represents the day of the year.

Julian dates consist of strings of contiguous numbers, which means that zeros must pad any space between the year and the day values.

Julian dates that contain year values before 1582 are invalid for the conversion to Gregorian dates.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input julian_date julian7.;
```

Data Line	Results*
`----+----1`	
`99075`	`14319`
`1999075`	`14319`

* The input values correspond to the 75th day of 1999, which is March 16.

See Also

Format:

"JULIAN*w*. Format" on page 194

Functions:

"DATEJUL Function" on page 635

"JULDATE Function" on page 866

System Option:

"YEARCUTOFF= System Option" on page 2058

MDYAMPM*w.d* Informat

Reads datetime values in the form *mm-dd-yy<yy> hh:mm:ss.ss* AM|PM, where a special character such as a hyphen (-), period (.), slash (/), or colon (:) separates the month, day, and year; the year can be either 2 or 4 digits.

Category: Date and Time

Alignment: right

Default Time Period: AM

Requirement: A space must separate the date and the time.

Syntax

MDYAMPM*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 19

Range: 8–40

d

specifies the number of digits to the right of the decimal point in the seconds value. The digits to the right of the decimal point specify a fraction of a second. This argument is optional.

Default: 0

Range: 0–39

Details

The MDYAMPMw.d format reads SAS datetime values in the following form:

mm-dd-yy<*yy*> *hh*:*mm*<:*ss*<.*ss*>> <AM | PM>

where:

mm
is an integer from 01 through 12 that represents the month.

dd
is an integer from 01 through 31 that represents the day of the month.

yy or *yyyy*
specifies a two-digit or four-digit integer that represents the year.

hh
is the number of hours that range from 00 through 23.

mm
is the number of minutes that range from 00 through 59.

ss.*ss*
is the number of seconds that range from 00 through 59 with the fraction of a second following the decimal point.

Requirement: If a fraction of a second is specified, the decimal point can be represented only by a period and is required.

AM | PM
specifies either the time period 00:01–12:00 noon (AM) or the time period 12:01 – 12:00 midnight (PM)

- or :
represents one of several special characters, such as the slash (/), hyphen (-), colon (:), or a blank character that can be used to separate date and time components. Special characters can be used as separators between any date or time component and between the date and the time.

Comparisons

The MDYAMPM*w.d* informat reads datetime values with optional separators in the form *mm-dd-yy*<*yy*> *hh:mm:ss.ss* AM | PM, and requires a space between the date and the time.

The DATETIME*w.d* informat reads datetime values with optional separators in the form *dd-mmm-yy*<*yy*> *hh:mm:ss.ss* AM|PM, and the date and time can be separated by a special character.

The YMDDTTM*w.d* informat reads datetime values with required separators in the form <*yy*>*yy-mm-dd*/*hh:mm:ss.ss*.

Examples

```
input @1 dt mdyampm25.2.;
```

Data Line	Results
`09.15.2008 03:53:00 pm`	`1537113180`
`09-15-08 3.53 pm`	`1537113180`

See Also

Informat:

"DATETIME*w*. Informat" on page 1324
"YMDDTTM*w.d* Informat" on page 1408

MMDDYY*w*. Informat

Reads date values in the form *mmddyy* or *mmddyyyy*.

Category: Date and Time

Syntax

MMDDYY*w*.

Syntax Description

w
specifies the width of the input field.
Default: 6
Range: 6–32

Details

The date values must be in the form *mmddyy* or *mmddyyyy*, where

mm
is an integer from 01 through 12 that represents the month.

dd
is an integer from 01 through 31 that represents the day of the month.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

You can separate the month, day, and year fields by blanks or by special characters. However, if you use delimiters, place them between all fields in the value. Blanks can also be placed before and after the date.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input calendar_date mmddyy8.;
```

Data Line	Results
----+----1----+	
031699	14319
03/16/99	14319
03 16 99	14319
03161999	14319

See Also

Formats:

"DATE*w*. Format" on page 151
"DDMMYY*w*. Format" on page 157
"MMDDYY*w*. Format" on page 196
"YYMMDD*w*. Format" on page 273

Functions:

"DAY Function" on page 637
"MDY Function" on page 924
"MONTH Function" on page 936
"YEAR Function" on page 1233

Informats:

"DATE*w*. Informat" on page 1322
"DDMMYY*w*. Informat" on page 1326
"YYMMDD*w*. Informat" on page 1410

System Option:

"YEARCUTOFF= System Option" on page 2058

MONYY*w*. Informat

Reads month and year date values in the form *mmmyy* or *mmmyyyy*.

Category: Date and Time

Syntax

MONYY*w*.

Syntax Description

w
specifies the width of the input field.
Default: 5
Range: 5–32

Details

The date values must be in the form *mmmyy* or *mmmyyyy*, where

mmm
is the first three letters of the month name.

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

A value read with the MONYY*w*. informat results in a SAS date value that corresponds to the first day of the specified month.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOF= system option. △

Examples

```
input month_and_year monyy7.;
```

Data Line	Results
`----+----1----+`	
`mar 99`	`14304`
`mar1999`	`14304`

See Also

Formats:
"DDMMYY*w*. Format" on page 157

"MMDDYY*w*. Format" on page 196
"MONYY*w*. Format" on page 206
"YYMMDD*w*. Format" on page 273
Functions:
"MONTH Function" on page 936
"YEAR Function" on page 1233
Informats:
"DDMMYY*w*. Informat" on page 1326
"MMDDYY*w*. Informat" on page 1349
"YYMMDD*w*. Informat" on page 1410
System Option:
"YEARCUTOFF= System Option" on page 2058

MSEC*w*. Informat

Reads TIME MIC values.

Category: Date and Time

Syntax

MSEC*w*.

Syntax Description

w
specifies the width of the input field.

Requirement: *w* must be 8 because the OS TIME macro or the STCK System/370 instruction on IBM mainframes each return an eight-byte value.

Details

The MSEC*w*. informat reads time values that are produced by IBM mainframe operating environments and converts the time values to SAS time values.

Use the MSEC*w*. informat to find the difference between two IBM mainframe TIME values, with precision to the nearest microsecond.

Comparisons

The MSEC*w*. and TODSTAMP*w*. informats both read IBM time-of-day clock values, but the MSEC*w*. informat assigns a time value to a variable, and the TODSTAMP*w*. informat assigns a datetime value.

Examples

```
input btime msec8.;
```

Data Line*	Results
0000EA044E65A000	**62818.412122**

* The data line is a hexadecimal representation of a binary 8-byte time-of-day clock value. Each byte occupies one column of the input field. The result is a SAS time value corresponding to 5:26:58.41 p.m.

See Also

Informat:

"TODSTAMP*w*. Informat" on page 1395

NUMX*w.d* Informat

Reads numeric values with a comma in place of the decimal point.

Category: Numeric

Syntax

NUMX*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 12

Range: 1–32

d

specifies the number of digits to the right of the decimal. If the data contain decimal points, the *d* value is ignored. This argument is optional.

Range: 0–31

Details

The NUMX*w.d* informat reads numeric values and interprets a comma as a decimal point.

Comparisons

The NUMX*w.d* informat is similar to the *w.d* informat except that it reads numeric values that contain a comma in place of the decimal point.

Examples

```
input @1 x numx10.;
```

Data Line	Results
----+----1----+	
896,48	896.48
3064,1	3064.1
6489	6489

See Also

Formats:

"NUMX*w.d* Format" on page 209
"*w.d* Format" on page 254

OCTAL*w.d* Informat

Converts positive octal values to integers.

Category: Numeric

Syntax

OCTAL*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 3

Range: 1–24

d

specifies the power of 10 by which to divide the value. This argument is optional.

Range: 1–31

Restriction: must be greater than or equal to the *w* value.

Details

Use only the digits 0 through 7 in the input, with no embedded blanks. The OCTAL*w.d* informat ignores leading and trailing blanks.

OCTAL*w.d* cannot read negative values. It treats all input values as positive (unsigned).

Examples

```
input @1 value octal3.1;
```

Data Line	Results
`----+----1`	
`177`	`12.7`

PD*w.d* Informat

Reads data that are stored in IBM packed decimal format.

Category: Numeric

See: PB*w.d* Informat in the documentation for your operating environment.

Syntax

PD*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–16

d
specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

The PD*w.d* informat is useful because many programs write data in packed decimal format for storage efficiency, fitting two digits into each byte and using only a half byte for a sign.

Note: Different operating environments store packed decimal values in different ways. However, PD*w.d* reads packed decimal values with consistent results if the values are created on the same type of operating environment that you use to run SAS. △

The PD*w.d* format writes missing numerical data as -0. When the PD*w.d* informat reads -0, it stores it as 0.

Comparisons

The following table compares packed decimal notation in several programming languages:

Language	Notation
SAS	PD4.
COBOL	COMP-3 PIC S9(7)
IBM 370 Assembler	PL4
PL/I	FIXED DEC

Examples

Example 1: Reading Packed Decimal Data

```
input @1 x pd4.;
```

Data Line*	Results
`----+----1`	
0000128C	**128**

* The data line is a hexadecimal representation of a binary number stored in packed decimal form. Each byte occupies one column of the input field.

Example 2: Creating a SAS Date with Packed Decimal Data

```
input mnth pd4.;
date=input(put(mnth,6.),mmddyy6.);
```

Data Line*	Results
`----+----1`	
0122599C	**14603**

* The data line is a hexadecimal representation of a binary number that is stored in packed decimal form on an IBM mainframe operating environment. Each byte occupies one column of the input field. The result is a SAS date value that corresponds to December 25, 1999.

PDJULG*w*. Informat

Reads packed Julian date values in the hexadecimal form *yyyyddd*F for IBM.

Category: Date and Time

Syntax

PDJULG*w*.

Syntax Description

w
specifies the width of the input field.

Default: 4

Range: 4

Details

The PDJULG*w*. informat reads IBM packed Julian date values in the form of *yyyyddd*F, converting them to SAS date values, where

yyyy
is the two-byte representation of the four-digit Gregorian year.

ddd
is the one-and-a-half byte representation of the three-digit integer that corresponds to the Julian day of the year, 1–365 (or 1–366 for leap years).

F
is the half byte that contains all binary 1s, which assigns the value as positive.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input date pdjulg4.;
```

Data Line	Results*
`----+----1`	
`1999003F`	**`14247`**

* SAS date value 14247 represents January 3, 1999.

See Also

Formats:

"JULDAY*w*. Format" on page 193
"JULIAN*w*. Format" on page 194
"PDJULG*w*. Format" on page 213
"PDJULI*w*. Format" on page 214

Functions:

"DATEJUL Function" on page 635
"JULDATE Function" on page 866

Informats:

"JULIAN*w*. Informat" on page 1346
"PDJULI*w*. Informat" on page 1358

System Option:

"YEARCUTOFF= System Option" on page 2058

PDJULI*w*. Informat

Reads packed Julian dates in the hexadecimal format *ccyddd*F for IBM.

Category: Date and Time

Syntax

PDJULI*w*.

Syntax Description

w
specifies the width of the input field.
Default: 4
Range: 4

Details

The PDJULI*w*. informat reads IBM packed Julian date values in the form *ccyyddd*F, converting them to SAS date values, where

cc
is the one-byte representation of a two-digit integer that represents the century.

yy
is the one-byte representation of a two-digit integer that represents the year. The PDJULI*w* informat makes an adjustment to the one-byte century representation by adding 1900 to the two-byte *ccyy* value in order to produce the correct four–digit Gregorian year. This adjustment causes *ccyy* values of 0098 to become 1998, 0101 to become 2001, and 0218 to become 2118.

ddd
is the one-and-a-half bytes representation of the three-digit integer that corresponds to the Julian day of the year, 1–365 (or 1–366 for leap years).

F
is the half byte that contains all binary 1s, which assigns the value as positive.

Examples

```
input date pdjuli4.;
```

Data Line	Results*
`----+----1`	
0099001F	**14245**
0110015F	**18277**

* SAS date value 14245 is January 1, 1999. SAS date value 18277 is January 15, 2010.

See Also

Formats:
- "JULDAY*w*. Format" on page 193
- "JULIAN*w*. Format" on page 194
- "PDJULG*w*. Format" on page 213
- "PDJULI*w*. Format" on page 214

Functions:
- "DATEJUL Function" on page 635
- "JULDATE Function" on page 866

Informats:
- "JULIAN*w*. Informat" on page 1346
- "PDJULG*w*. Informat" on page 1357

System Option:
- "YEARCUTOFF= System Option" on page 2058

PDTIME*w*. Informat

Reads packed decimal time of SMF and RMF records.

Category: Date and Time

Syntax

PDTIME*w*.

Syntax Description

w
specifies the width of the input field.

Requirement: *w* must be 4 because packed decimal time values in RMF and SMF records contain four bytes of information.

Details

The PDTIME*w*. informat reads packed decimal time values that are contained in SMF and RMF records that are produced by IBM mainframe systems and converts the values to SAS time values.

The general form of a packed decimal time value in hexadecimal notation is 0*hhmmss*F, where

0
is a half byte that contains all 0s.

hh
is one byte that represents two digits that correspond to hours.

mm
is one byte that represents two digits that correspond to minutes.

ss
is one byte that represents two digits that correspond to seconds.

F
is a half byte that contains all 1s.

If a field contains all 0s, PDTIME*w*. treats it as a missing value.

PDTIME*w*. enables you to read packed decimal time values from files that are created on an IBM mainframe on any operating environment.

Examples

```
input begin pdtime4.;
```

Data Line*	Results
0142225F	**51745**

* The data line is a hexadecimal representation of a binary time value that is stored in packed decimal form. Each byte occupies one column of the input field. The result is a SAS time value that corresponds to 2:22.25 p.m.

PERCENT*w.d* Informat

Reads percentages as numeric values.

Category: Numeric

Syntax

PERCENT*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 6
Range: 1–32

d
specifies the power of 10 by which to divide the value. If the data contain decimal points, the *d* value is ignored. This argument is optional.
Range: 0–31

Details

The PERCENT*w.d* informat converts the numeric portion of the input data to a number using the same method as the COMMA*w.d* informat. If a percent sign (%) follows the number in the input field, PERCENT*w.d* divides the number by 100.

Examples

```
input @1 x percent3. @4 y percent5.;
```

Data Line	Results
----+----1----+	
1% (20%)	**0.01 -0.2**

PIB*w.d* Informat

Reads positive integer binary (fixed-point) values.

Category: Numeric

See: PIB*w.d* Informat in the documentation for your operating environment.

Syntax

PIB*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–8

d
specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

All values are treated as positive. PIB*w.d* reads positive integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Note: Different operating environments store positive integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 1263. △

Comparisons

- Positive integer binary values are the same as integer binary values except that the sign bit is part of the value, which is always a positive integer. The PIB*w.d* informat treats all values as positive and includes the sign bit as part of the value.
- The PIB*w.d* informat with a width of 1 results in a value that corresponds to the binary equivalent of the contents of a byte. The binary equivalent of the contents of a byte is useful if your data contain values between hexadecimal 80 and hexadecimal FF, where the high-order bit can be misinterpreted as a negative sign.
- The IB*w.d* and PIB*w.d* informats are used to read native format integers. (Native format allows you to read and write values that are created in the same operating environment.) The IBR*w.d* and PIBR*w.d* informats are used to read little endian integers in any operating environment.

To view a table that shows the type of informat to use with big endian and little endian integers, see Table 5.1 on page 1264.

To view a table that compares integer binary notation in several programming languages, see Table 5.2 on page 1264.

Examples

You can use the INPUT statement and specify the PIB informat. However, in these examples we use the informat with the INPUT function, where binary input values are described by using a hexadecimal literal.

```
x=input('0100'x,pib2.);
y=input('0001'x,pib2.);
```

SAS Statement	Results on Big Endian Platforms	Results on Little Endian Platforms
`put x=;`	`256`	`1`
`put y=;`	`1`	`256`

See Also

Informat:

"PIBR*w.d* Informat" on page 1363

PIBR*w.d* Informat

Reads positive integer binary (fixed-point) values in Intel and DEC formats.

Category: Numeric

Syntax

PIBR*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–8

d
specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

All values are treated as positive. PIBR*w.d* reads positive integer binary values that have been generated by and for Intel and DEC operating environments. Use PIBR*w.d* to read positive integer binary data from Intel or DEC environments on other operating environments. The PIBR*w.d* informat in SAS code allows for a portable implementation for reading the data in any operating environment.

Note: Different operating environments store positive integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 1263. △

Comparisons

- Positive integer binary values are the same as integer binary values except that the sign bit is part of the value, which is always a positive integer. The PIBR*w.d* informat treats all values as positive and includes the sign bit as part of the value.
- The PIBR*w.d* informat with a width of 1 results in a value that corresponds to the binary equivalent of the contents of a byte. This is useful if your data contain values between hexadecimal 80 and hexadecimal FF, where the high-order bit can be misinterpreted as a negative sign.
- On Intel and DEC platforms, the PIB*w.d* and PIBR*w.d* informats are equivalent.
- The IB*w.d* and PIB*w.d* informats are used to read native format integers. (Native format allows you to read and write values that are created in the same operating environment.) The IBR*w.d* and PIBR*w.d* informats are used to read little endian integers in any operating environment.

  To view a table that shows the type of informat to use with big endian and little endian integers, see Table 5.1 on page 1264.

  To view a table that compares integer binary notation in several programming languages, see Table 5.2 on page 1264.

Examples

You can use the INPUT statement and specify the PIBR informat. However, these examples use the informat with the INPUT function, where binary input values are described using a hexadecimal literal.

```
x=input('0100'x,pibr2.);
y=input('0001'x,pibr2.);
```

SAS Statement	Results on Big Endian Platforms	Results on Little Endian Platforms
`put x=;`	`1`	`1`
`put y=;`	`256`	`256`

See Also

Informat:

"PIB*w.d* Informat" on page 1362

PK*w.d* Informat

Reads unsigned packed decimal data.

Category: Numeric

Syntax

PK*w.d*

Syntax Description

w
specifies the number of bytes of unsigned packed decimal data, each of which contains two digits.

Default: 1

Range: 1–16

d
specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

Each byte of unsigned packed decimal data contains two digits.

Comparisons

The PK*w.d* informat is the same as the PD*w.d* informat, except that PK*w.d* treats the sign half of the field's last byte as part of the value, not as the sign of the value.

Examples

```
input @1 x pk3.;
```

Data Line*	Results
`----+----1`	
`001234`	`1234`

* The data line is a hexadecimal representation of a binary number stored in unsigned packed decimal form. Each byte occupies one column of the input field.

PUNCH.*d* Informat

Reads whether a row of column-binary data is punched.

Category: Column Binary

Syntax

PUNCH.*d*

Syntax Description

d
specifies which row in a card column to read.

Range: 1–12

Details

Column-binary data storage compresses data so that more than 80 items of data can be stored on a single "virtual" punch card.

This informat assigns the value 1 to the variable if row *d* of the current card column is punched, or 0 if row *d* of the current card column is not punched. After PUNCH.*d* reads a field, the pointer does not advance to the next column.

Examples

Data Line*	SAS Statement	Results
`12-7-8`	`input x punch.12`	`1`
	`input x punch.11`	`0`
	`input x punch0.7`	`1`

* The data line is "virtual" punched card code. The punch card column for the example data has row 12, row 7, and row 8 punched.

See Also

Informats:

"$CB*w*. Informat" on page 1284
"CB*w.d* Informat" on page 1319
"ROW*w.d* Informat" on page 1371

"How to Read Column-Binary Data" in *SAS Language Reference: Concepts*

RB*w.d* Informat

Reads numeric data that are stored in real binary (floating-point) notation.

Category: Numeric
See: RB*w.d* Informat in the documentation for your operating environment.

Syntax

RB*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 4
Range: 2–8

d
specifies the power of 10 by which to divide the value. This argument is optional.
Range: 0–10

Details

Note: Different operating environments store real binary values in different ways. However, the RB*w.d* informat reads real binary values with consistent results if the

values are created on the same type of operating environment that you use to run SAS. △

Comparisons

The following table compares the names of real binary notation in several programming languages:

	Real Binary Notation	
Language	4 Bytes	8 Bytes
SAS	RB4.	RB8.
Fortran	REAL*4	REAL*8
C	float	double
IBM 370 assembler	F	D
PL/I	FLOAT BIN(21)	FLOAT BIN(53)

CAUTION:
Using the RBw.d informat to read real binary information on equipment that conforms to the IEEE standard for floating-point numbers results in a truncated eight-byte number (double-precision), rather than in a true four-byte floating-point number (single-precision). △

Examples

```
input @1 x rb8.;
```

Data Line*	Results
`----+----1`	
`4280000000000000`	`128`

* The data line is a hexadecimal representation of a real binary (floating-point) number on an IBM mainframe operating environment. Each byte occupies one column of the input field.

See Also

Informat:

"IEEE*w.d* Informat" on page 1344

RMFDURw. Informat

Reads duration intervals of RMF records.

Category: Date and Time

Syntax

RMFDUR*w*.

Syntax Description

w

specifies the width of the input field.

Requirement: *w* must be 4 because packed decimal duration values in RMF records contain four bytes of information.

Details

The RMFDUR*w*. informat reads the duration of RMF measurement intervals of RMF records that are produced as packed decimal data by IBM mainframe systems and converts them to SAS time values.

The general form of the duration interval data in an RMF record in hexadecimal notation is *mmssttt*F, where

mm
is the one-byte representation of two digits that correspond to minutes.

ss
is the one-byte representation of two digits that correspond to seconds.

ttt
is the one-and-a-half-bytes representation of three digits that correspond to thousandths of a second.

F
is a half byte that contains all binary 1s, which assigns the value as positive.

If the field does not contain packed decimal data, then RMFDUR*w*. results in a missing value.

Comparisons

- Both the RMFDUR*w*. informat and the RMFSTAMP*w*. informat read packed decimal information from RMF records that are produced by IBM mainframe systems.
- The RMFDUR*w*. informat reads duration data and results in a time value.
- The RMFSTAMP*w*. informat reads time-of-day data and results in a datetime value.

Examples

```
input dura rmfdur4.;
```

Data Line*	Results
`----+----1----+`	
`3552226F`	**`2152.226`**

* The data line is a hexadecimal representation of a binary duration value that is stored in packed decimal form as it would appear in an RMF record. Each byte occupies one column of the input field. The result is a SAS time value corresponding to 00:35:52.226.

See Also

Informats:

"RMFSTAMP*w*. Informat" on page 1370
"SMFSTAMP*w*. Informat" on page 1390

RMFSTAMP*w*. Informat

Reads time and date fields of RMF records.

Category: Date and Time

Syntax

RMFSTAMP*w*.

Syntax Description

w
specifies the width of the input field.

Requirement: *w* must be 8 because packed decimal time and date values in RMF records contain eight bytes of information: four bytes of time data that are followed by four bytes of date data.

Details

The RMFSTAMP*w*. informat reads packed decimal time and date values of RMF records that are produced by IBM mainframe systems, and converts the time and date values to SAS datetime values.

The general form of the time and date information in an RMF record in hexadecimal notation is 0*hhmmss*F*ccyyddd*F, where

0
is the half byte that contains all binary 0s.

hh
is the one-byte representation of two digits that correspond to the hour of the day.

mm
is the one-byte representation of two digits that correspond to minutes.

ss
is 1 byte that represents two digits that correspond to seconds.

cc
is the one-byte representation of two digits that correspond to the century.

yy
is the one-byte representation of two digits that correspond to the year.

ddd
is the one-and-a-half bytes that contain three digits that correspond to the day of the year.

F
is the half byte that contains all binary 1s.
The century indicators 00 correspond to 1900, 01 to 2000, and 02 to 2100.

RMFSTAMP*w*. enables you to read, on any operating environment, packed decimal time and date values from files that are created on an IBM mainframe.

Comparisons

Both the RMFSTAMP*w*. informat and the PDTIME*w*. informat read packed decimal values from RMF records. The RMFSTAMP*w*. informat reads both time and date values and results in a SAS datetime value. The PDTIME*w*. informat reads only time values and results in a SAS time value.

Examples

```
input begin rmfstamp8.;
```

Data Line*	Results
`----+----1----+----2`	
`0142225F0102286F`	`1350138145`

* The data line is a hexadecimal representation of a binary time and date value that is stored in packed decimal form as it would appear in an RMF record. Each byte occupies one column of the input field. The result is a SAS datetime value that corresponds to October 13, 2002, 2:22.25 PM.

ROW*w.d* Informat

Reads a column-binary field down a card column.

Category: Column Binary

Syntax

ROW*w.d*

Syntax Description

w

specifies the row where the field begins.

Range: 0–12

d

specifies the length in rows of the field.

Default: 1

Range: 1–25

Details

Column-binary data storage compresses data so that more than 80 items of data can be stored on a single "virtual" punch card.

The ROW*w.d* informat assigns the relative position of the punch in the field to a numeric variable.

If the field that you specify has more than one punch, then ROW*w.d* assigns the variable a missing value and sets the automatic variable _ERROR_ to 1. If the field has no punches, then ROW*w.d* assigns the variable a missing value.

ROW*w.d* can read fields across columns, continuing with row 12 of the new column and going down through the rest of the rows. After ROW*w.d* reads a field, the pointer moves to the next row.

Examples

```
input x row5.3
input x row7.1
input x row5.2
input x row3.5
```

Data Line*	Results
`----+----1`	
`00`	
`04`	`3`
	`1`
	`.`
	`5`

* The data line is a hexadecimal representation of the column binary. The "virtual" punch card column for the example data has row 7 punched. The binary representation is 0000 0000 0000 0100.

See Also

Informats:

"$CB*w*. Informat" on page 1284
"CB*w.d* Informat" on page 1319
"PUNCH.*d* Informat" on page 1366

"How to Read Column-Binary Data" in *SAS Language Reference: Concepts*

S370FF*w.d* Informat

Reads EBCDIC numeric data.

Category: Numeric

Syntax

S370FF*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 12
Range: 1–32

d
specifies the power of 10 by which to divide the value. This argument is optional.
Range: 0–31

Details

The S370FF*w.d* informat reads numeric data that are represented in EBCDIC and converts the data to native format. If EBCDIC is the native format, S370FF*w.d* performs no conversion.

S370FF*w.d* reads EBCDIC numeric values that are represented with one byte per digit. Use S370FF*w.d* on other operating environments to read numeric data from IBM mainframe files.

S370FF*w.d* reads numeric values located anywhere in the input field. EBCDIC blanks can precede or follow a numeric value with no effect. If a value is negative, an EBCDIC minus sign should immediately precede the value. S370FF*w.d* reads values with EBCDIC decimal points and values in scientific notation, and it interprets a single EBCDIC period as a missing value.

Comparisons

The S370FF*w.d* informat performs the same role for numeric data that the $EBCDIC*w.d* informat does for character data. That is, on an IBM mainframe system, S370FF*w.d* has the same effect as the standard *w.d* informat. On all other systems, using S370FF*w.d* is equivalent to using $EBCDIC*w.d* as well as using the standard *w.d* informat.

Examples

```
input @1 x s370ff3.;
```

Data Line*	Results
`----+----1`	
`F1F2F3`	`123`
`F2F4F0`	`240`

* The data lines are hexadecimal representations of codes for characters. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one character value.

S370FIB*w.d* Informat

Reads integer binary (fixed-point) values, including negative values, in IBM mainframe format.

Category: Numeric

Syntax

S370FIB*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 4

Range: 1–8

d

specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

The S370FIB*w.d* informat reads integer binary (fixed-point) values that are stored in IBM mainframe format, including negative values that are represented in two's complement notation. S370FIB*w.d* reads integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Use S370FIB*w.d* for integer binary data that are created in IBM mainframe format for reading in other operating environments.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 1263. △

Comparisons

- If you use SAS on an IBM mainframe, S370FIB*w.d* and IB*w.d* are identical.
- S370FPIB*w.d*, S370FIBU*w.d*, and S370FIB*w.d* are used to read big endian integers in any operating environment.

  To view a table that shows the type of informat to use with big endian and little endian integers, see Table 5.1 on page 1264.

  To view a table that compares integer binary notation in several programming languages, see Table 5.2 on page 1264.

Examples

You can use the INPUT statement and specify the S370FIB informat. However, this example uses the informat with the INPUT function, where the binary input value is described by using a hexadecimal literal.

```
x=input('0080'x,s370fib2.);
```

SAS Statement	Results
put x=;	**128**

See Also

Informats:

"S370FIBU*w.d* Informat" on page 1376
"S370FPIB*w.d* Informat" on page 1380

S370FIBU*w.d* Informat

Reads unsigned integer binary (fixed-point) values in IBM mainframe format.

Category: Numeric

Syntax

S370FIBU*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 4

Range: 1–8

d
specifies the power of 10 by which to divide the value. SAS uses the *d* value even if the data contain decimal points. This argument is optional.

Range: 0–10

Details

The S370FIBU*w.d* informat reads unsigned integer binary (fixed-point) values that are stored in IBM mainframe format, including negative values that are represented in two's complement notation. Unsigned integer binary values are the same as integer binary values, except that all values are treated as positive. S370FIBU*w.d* reads integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Use S370FIBU*w.d* for unsigned integer binary data that are created in IBM mainframe format for reading in other operating environments.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 1263. △

Comparisons

- The S370FIBU*w.d* informat is equivalent to the COBOL notation PIC 9(*n*) BINARY, where *n* is the number of digits.
- The S370FIBU*w.d* and S370FPIB*w.d* informats are identical.
- S370FPIB*w.d*, S370FIBU*w.d*, and S370FIB*w.d* are used to read big endian integers in any operating environment.

  To view a table that shows the type of informat to use with big endian and little endian integers, see Table 5.1 on page 1264.

  To view a table that compares integer binary notation in several programming languages, see Table 5.2 on page 1264.

Examples

You can use the INPUT statement and specify the S370FIBU informat. However, these examples use the informat with the INPUT function, where binary input values are described by using a hexadecimal literal.

```
x=input('7F'x,s370fibu1.);
y=input('F6'x,s370fibu1.);
```

SAS Statement	Results
put x=;	**127**
put y=;	**246**

See Also

Informats:

"S370FIB*w.d* Informat" on page 1374

"S370FPIB*w.d* Informat" on page 1380

S370FPD*w.d* Informat

Reads packed data in IBM mainframe format.

Category: Numeric

Syntax

S370FPD*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–16

d
specifies the power of 10 by which to divide the value. This argument is optional.

Default: 0

Range: 0–31

Details

Packed decimal data contain two digits per byte, but only one digit in the input field represents the sign. The last half of the last byte indicates the sign: a C or an F for positive numbers and a D for negative numbers.

Use S370FPD*w.d* to read packed decimal data from IBM mainframe files on other operating environments.

Comparisons

- If you use SAS on an IBM mainframe, the S370FPD*w.d* and the PD*w.d* informats are identical.
- The following table compares the equivalent packed decimal notation by programming language:

Language	Packed Decimal Notation
SAS	S370FPD4.
PL/I	FIXED DEC(7,0)
COBOL	COMP-3 PIC 9(7)
assembler	PL4

S370FPDU*w.d* Informat

Reads unsigned packed decimal data in IBM mainframe format.

Category: Numeric

Syntax

S370FPDU*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–16

d
specifies the power of 10 by which to divide the value. This argument is optional.

Default: 0

Range: 0–31

Details

Packed decimal data contain two digits per byte. The last half of the last byte, which indicates the sign for signed packed data, is always F for unsigned packed data.

Use S370FPDU*w.d* on other operating environments to read unsigned packed decimal data from IBM mainframe files.

Comparisons

- The S370FPDU*w.d* informat is similar to the S370FPD*w.d* informat except that the S370FPDU*w.d* informat rejects all sign digits except F.
- The S370FPDU*w.d* informat is equivalent to the COBOL notation PIC 9(*n*) PACKED-DECIMAL, where the *n* value is the number of digits.

Examples

```
input @1 x s370fpdu3.;
```

Data Line*	Results
`----+----1`	
12345F	**12345**

* The data line is a hexadecimal representation of a binary number that is stored in packed decimal form. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one column of the input field.

S370FPIB*w.d* Informat

Reads positive integer binary (fixed-point) values in IBM mainframe format.

Category: Numeric

Syntax

S370FPIB*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 4

Range: 1–8

d
specifies the power of 10 by which to divide the value. This argument is optional.

Default: 0

Range: 0–10

Details

Positive integer binary values are the same as integer binary values, except that all values are treated as positive. S370FPIB*w.d* reads integer binary values with consistent results if the values are created in the same type of operating environment that you use to run SAS.

Use S370FPIB*w.d* for positive integer binary data that are created in IBM mainframe format for reading in other operating environments.

Note: Different operating environments store integer binary values in different ways. This concept is called byte ordering. For a detailed discussion about byte ordering, see "Byte Ordering for Integer Binary Data on Big Endian and Little Endian Platforms" on page 1263. △

Comparisons

- If you use SAS on an IBM mainframe, S370FPIB*w.d* and PIB*w.d* are identical.
- S370FPIB*w.d*, S370FIBU*w.d*, and S370FIB*w.d* are used to read big endian integers in any operating environment.

  To view a table that shows the type of informat to use with big endian and little endian integers, see Table 5.1 on page 1264.

  To view a table that compares integer binary notation in several programming languages, see Table 5.2 on page 1264.

Examples

You can use the INPUT statement and specify the S370FPIB informat. However, this example uses the informat with the INPUT function, where the binary input value is described using a hexadecimal literal.

```
x=input('0100'x,s370fpib2.);
```

SAS Statement	Results
put x=4;	**256**

See Also

Informats:

"S370FIB*w.d* Informat" on page 1374

"S370FIBU*w.d* Informat" on page 1376

S370FRB*w.d* Informat

Reads real binary (floating-point) data in IBM mainframe format.

Category: Numeric

Syntax

S370FRB*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 6

Range: 2–8

d

specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–10

Details

Real binary values are represented in two parts: a mantissa that gives the value, and an exponent that gives the value's magnitude.

Use S370FRB*w.d* to read real binary data from IBM mainframe files on other operating environments.

Comparisons

- If you use SAS on an IBM mainframe, S370FRB*w.d* and RB*w.d* are identical.
- The following table shows the equivalent real binary notation for several programming languages:

	Real Binary Notation	
Language	4 Bytes	8 Bytes
SAS	S370FRB4.	S370FRB8.
PL/I	FLOAT BIN(21)	FLOAT BIN(53)
Fortran	REAL*4	REAL*8
COBOL	COMP-1	COMP-2
assembler	E	D
C	float	double

See Also

Informat:

"RB*w.d* Informat" on page 1367

S370FZDB*w.d* Informat

Reads zoned decimal data in which zeros have been left blank.

Category: Numeric
See: ZBDw.d Informat in *SAS Companion for z/OS*

Syntax

S370FZBD*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 8
Range: 1–32

d
specifies the power of 10 by which to divide the value. This argument is optional.
Default: 0
Range: 0–31

Details

Use the S370ZFDB*w.d* informat on other operating environments to read zoned decimal data from IBM mainframe files.

Examples

```
input @1 x s370fzdb8.;
```

Data Line *	Results
—-+—-1	
40404040F14040C0	1000
4040404040F1F2D3	–123

* The data lines contain a hexadecimal representation of a binary number that is stored in zoned decimal format on an IBM mainframe operating environment. Two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one column of the input field.

S370FZD*w.d* Informat

Reads zoned decimal data in IBM mainframe format.

Category: Numeric

Syntax

S370FZD*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 8
Range: 1–32

d
specifies the power of 10 by which to divide the value. If the data contain decimal points, the *d* value is ignored. This argument is optional.
Default: 0
Range: 0–31

Details

Zoned decimal data are similar to standard decimal data in that every digit requires one byte. However, the value's sign is stored in the last byte, along with the last digit.

Use S370FZD*w.d* on other operating environments to read zoned decimal data from IBM mainframe files.

Comparisons

- If you use SAS on an IBM mainframe, S370FZD*w.d* and ZD*w.d* are identical.
- The following table shows the equivalent zoned decimal notation for several programming languages:

Language	Zoned Decimal Notation
SAS	S370FZD3.
PL/I	PICTURE'99T'
COBOL	PIC S9(3) DISPLAY
assembler	ZL3

Examples

```
input @1 x s370fzd3.;
```

Data Line*	Results
`----+----1`	
`F1F2C3`	`123`
`F1F2D3`	`-123`

* The data line contains a hexadecimal representation of a binary number stored in zoned decimal format on an IBM mainframe operating environment. Each two hexadecimal characters

correspond to one byte of binary data, and each byte corresponds to one column of the input field.

See Also

Informat:

"ZD*w.d* Informat" on page 1414

S370FZDL*w.d* Informat

Reads zoned decimal leading-sign data in IBM mainframe format.

Category: Numeric

Syntax

S370FZDL*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 8

Range: 1–32

d

specifies the power of 10 by which to divide the value. This argument is optional.

Default: 0

Range: 0–31

Details

Use S370FZDL*w.d* on other operating environments to read zoned decimal data from IBM mainframe files.

Comparisons

- Zoned decimal leading-sign data is similar to standard zoned decimal data except that the sign of the value is stored in the first byte of zoned decimal leading-sign data, along with the first digit.
- The S370FZDL*w.d* informat is equivalent to the COBOL notation PIC S9(*n*) DISPLAY SIGN LEADING, where the *n* value is the number of digits.

Examples

```
input @1 x s370fzdl3.;
```

Data Line*	Results
`----+----1`	
`C1F2F3`	`123`
`D1F2F3`	`-123`

* The data lines contain a hexadecimal representation of a binary number stored in zoned decimal format on an IBM mainframe operating environment. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one column of the input field.

S370FZDS*w.d* Informat

Reads zoned decimal separate leading-sign data in IBM mainframe format.

Category: Numeric

Syntax

S370FZDS*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 8

Range: 2–32

d
specifies the power of 10 by which to divide the value. This argument is optional.

Default: 0

Range: 0–31

Details

Use S370FZDS*w.d* on other operating environments to read zoned decimal data from IBM mainframe files.

Comparisons

- Zoned decimal separate leading-sign data is similar to standard zoned decimal data except that the sign of the value is stored in the first byte of zoned decimal leading sign data, and the first digit of the value is stored in the second byte.
- The S370FZDS*w.d* informat is equivalent to the COBOL notation PIC S9(*n*) DISPLAY SIGN LEADING SEPARATE, where the *n* value is the number of digits.

Examples

```
input @1 x s370fzds4.;
```

Data Line*	Results
`----+----1`	
`4EF1F2F3`	`123`
`60F1F2F3`	`-123`

* The data line contains a hexadecimal representation of a binary number that is stored in zoned decimal format on an IBM mainframe operating environment. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one column of the input field.

S370FZDT*w.d* Informat

Reads zoned decimal separate trailing-sign data in IBM mainframe format.

Category: Numeric

Syntax

S370FZDT*w.d*

Syntax Description

w
specifies the width of the input field.

Default: 8

Range: 2–32

d
specifies the power of 10 by which to divide the value. This argument is optional.

Default: 0

Range: 0–31

Details

Use S370FZDT*w.d* on other operating environments to read zoned decimal data from IBM mainframe files.

Comparisons

- Zoned decimal separate trailing-sign data are similar to zoned decimal separate leading-sign data except that the sign of the value is stored in the last byte of zoned decimal separate trailing-sign data.

- The S370FZDT*w.d* informat is equivalent to the COBOL notation PIC S9(*n*) DISPLAY SIGN TRAILING SEPARATE, where the *n* value is the number of digits.

Examples

```
input @1 x s370fzdt4.;
```

Data Line*	Results
`----+----1`	
`F1F2F34E`	`123`
`F1F2F360`	`-123`

* The data line contains a hexadecimal representation of a binary number that is stored in zoned decimal format on an IBM mainframe operating environment. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one column of the input field.

S370FZDU*w.d* Informat

Reads unsigned zoned decimal data in IBM mainframe format.

Category: Numeric

Syntax

S370FZDU*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 8
Range: 1–32

d
specifies the power of 10 by which to divide the value. This argument is optional.
Default: 0
Range: 0–31

Details

Use S370FZDU*w.d* on other operating environments to read unsigned zoned decimal data from IBM mainframe files.

Comparisons

- The S370FZDU*w.d* informat is similar to the S370FZD*w.d* informat except that the S370FZDU*w.d* informat rejects all sign digits except F.

- The S370FZDU*w*.*d* informat is equivalent to the COBOL notation PIC 9(*n*) DISPLAY, where the *n* value is the number of digits.

Examples

```
input @1 x s370fzdu3.;
```

Data Line*	Results
`----+----1`	
`F1F2F3`	`123`

* The data line contains a hexadecimal representation of a binary number that is stored in zoned decimal format on an IBM mainframe operating environment. Each two hexadecimal characters correspond to one byte of binary data, and each byte corresponds to one column of the input field.

SHRSTAMP*w*. Informat

Reads date and time values of SHR records.

Category: Date and Time

Syntax

SHRSTAMP*w*.

Syntax Description

w

specifies the width of the input field.

Requirement: *w* must be 8 because packed decimal date and time values in SHR records contain eight bytes of information: four bytes of date data that are followed by four bytes of time data.

Details

The SHRSTAMP*w*. informat reads packed decimal date and time values of SHR records that are produced by IBM mainframe environments and converts the date and time values to SAS datetime values.

The general form of the date and time information in an SHR record in hexadecimal notation is *ccyyddd*F*hhmmssth*, where

ccyy

is the two byte representation of the year. The *cc* portion is the one byte representation of a two-digit integer that represents the century. The *yy* portion is the one byte representation of two digits that correspond to the year.

The *cc* portion is the century indicator where 00 indicates 19yy, 01 indicates 20yy, 02 indicates 21yy, and so on. A hexadecimal year value of 0115 is equal to the year 2015.

ddd
is the one-and-a-half bytes that contain three digits that correspond to the day of the year.

F
is the half byte that contains all binary 1s.

hh
is the one byte representation of two digits that correspond to the hour of the day.

mm
is the one byte representation of two digits that correspond to minutes.

ss
is the one byte representation of two digits that correspond to seconds.

th
is the one byte representation of two digits that correspond to a 100th of a second.

The SHRSTAMP*w*. informat enables you to read, on any operation environment, packed decimal date and time values from files that are created on an IBM mainframe.

Examples

```
input begin shrstamp8.;
```

Data Line*	Results
`----+----1----+----2`	
`0097239F12403576`	`1188304835.8`

* The data line is a hexadecimal representation of a packed decimal date and time value that is stored as it would appear in an SHR record. Each byte occupies one column of the input field. The result is a SAS datetime value that corresponds to Aug. 27, 1997 12:40:36 p.m.

SMFSTAMP*w*. Informat

Reads time and date values of SMF records.

Category: Date and Time

Syntax

SMFSTAMP*w*.

Syntax Description

w
specifies the width of the input field.

Requirement: *w* must be 8 because time and date values in SMF records contain eight bytes of information: four bytes of time data that are followed by four bytes of date data.

Tip: The time portion of an SMF record is a four-byte integer binary number that represents time as the number of hundredths of a second past midnight.

Details

The SMFSTAMP*w*. informat reads integer binary time values and packed decimal date values of SMF records that are produced by IBM mainframe systems and converts the time and date values to SAS datetime values.

The date portion of an SMF record in hexadecimal notation is *ccyyddd*F, where

cc
is the one-byte representation of two digits that correspond to the century.

yy
is the one-byte representation of two digits that correspond to the year.

ddd
is the one-and-a-half bytes that contain three digits that correspond to the day of the year.

F
is the half byte that contains all binary 1s.

The SMFSTAMP*w*. informat enables you to read, on any operating environment, integer binary time values and packed decimal date values from files that are created on an IBM mainframe.

Examples

```
input begin smfstamp8.;
```

Data Line*	Results
`----+----1----+----2`	
`0058DC0C0098200F`	`1216483835`

* The data line is a hexadecimal representation of a binary time and date value that is stored as it would appear in an SMF record. Each byte occupies one column of the input field. The result is a SAS datetime value that corresponds to July 19, 1998 4:10:35 PM.

STIMER*w*. Informat

Reads time values and determines whether the values are hours, minutes, or seconds; reads the output of the STIMER system option.

Category: Date and Time

Syntax

STIMER*w*.

Syntax Description

w
specifies the width of the input field.

Details

The STIMER informat reads performance statistics that the STIMER system option writes to the SAS log.

The informat reads time values and determines whether the values are hours, minutes, or seconds based on the presence of decimal points and colons:

- If no colon is present, the value is the number of seconds.
- If a single colon is present, the value before the colon is the number of minutes. The value after the colon is the number of seconds.
- If two colons are present, the sequence of time is hours, minutes, and then seconds.

In all cases, the result is a SAS time value.

The input values for STIMER must be in one of the following forms:

- *ss*
- *ss.ss*
- *mm:ss*
- *mm:ss.ss*
- *hh:mm:ss*
- *hh:mm:ss.ss*

where

ss
is an integer that represents the number of seconds.

mm
is an integer that represents the number of minutes.

hh
is an integer that represents the number of hours.

TIME*w*. Informat

Reads hours, minutes, and seconds in the form *hh:mm:ss.ss*, where special characters such as the colon (:) or the period (.) are used to separate the hours, minutes, and seconds.

Category: Date and Time

Syntax

TIME*w*.

Syntax Description

w
specifies the width of the input field.

Default: 8

Range: 5–32

Details

The TIME*w*. informat reads SAS time values in the following form:

hh:mm:ss<.*ss*> <AM | PM>

where

hh
is an integer that represents the number of hours.

:
represents a special character that separates hours, minutes, and seconds.

mm
is the number of minutes that range from 00 through 59.

ss<.*ss*>
is an integer that represents the number of seconds, and if needed, tenths of a second. Seconds and tenths of a second must always be separated by a period.

AM | PM
AM indicates time between 12:00 midnight and 11:59 in the morning. PM indicates time between 12:00 noon and 11:59 at night.

Separate *hh*, *mm*, and *ss* with a special character. When the period is used as the special character, the time is interpreted in the order hours, minutes, and seconds. For example, 23.22 is 23 hours and 22 minutes, not 23 minutes and 22 seconds, or 23 seconds and 22 tenths of a second.

If you do not enter a value for seconds, SAS assumes a value of 0.

The stored value is the total number of seconds in the time value.

Examples

```
input begin time10.;
```

Data Line	Results	Formatted with TIME*w*.
`----+----1----+`		
`12.56`	`46560`	`12:56:00`
`120:120`	`439200`	`122:00:00`
`1:13 pm`	`47580`	`13:13:00`

See Also

Formats:

"HHMM*w*.*d* Format" on page 185
"HOUR*w*.*d* Format" on page 188
"MMSS*w*.*d* Format" on page 200
"TIME*w*.*d* Format" on page 245

Functions:

"HOUR Function" on page 807
"MINUTE Function" on page 928
"SECOND Function" on page 1122
"TIME Function" on page 1161

TODSTAMP*w*. Informat

Reads an eight-byte time-of-day stamp.

Category: Date and Time

Syntax

TODSTAMP*w*.

Syntax Description

w
specifies the width of the input field.

Requirement: *w* must be 8 because the OS TIME macro or the STCK instruction on IBM mainframes each return an eight-byte value.

Details

The TODSTAMP*w*. informat reads time-of-day clock values that are produced by IBM mainframe operating systems and converts the clock values to SAS datetime values.

If the time-of-day value is all 0s, TODSTAMP*w*. results in a missing value.

Use TODSTAMP*w*. on other operating environments to read time-of-day values that are produced by an IBM mainframe.

Examples

```
input btime todstamp8.;
```

Data Line*	Results
`----+----1----+----2`	
`B361183D5FB80000`	`1262303998`

* The data line is a hexadecimal representation of a binary, 8-byte time-of-day clock value. Each byte occupies one column of the input field. The result is a SAS datetime value that corresponds to December 31, 1999, 11:59:58 p.m.

TRAILSGN*w*. Informat

Reads a trailing plus (+) or minus (–) sign.

Category: Numeric

Syntax

TRAILSGN*w*.

Syntax Description

w
specifies the width of the input field.

Default: 6

Range: 1–32

Details

If the data contains a decimal point, the TRAILSGN informat honors the number of decimal places that are in the input data. If the data contains a comma, the TRAILSGN informat reads the value, ignoring the comma.

Examples

```
input x trailsgn8.;
```

Data Line	Results
`----+----1----+`	
`1`	`1`
`1,000`	`1000`
`1+`	`1`
`1-`	`-1`
`1.2`	`1.2`
`1.2+`	`1.2`
`1.2-`	`-1.2`

TU*w*. Informat

Reads timer units.

Category: Date and Time

Syntax

TU*w*.

Syntax Description

w
specifies the width of the input field.

Requirement: *w* must be 4 because the OS TIME macro returns a four-byte value.

Details

The TU*w*. informat reads timer unit values that are produced by IBM mainframe operating environments and converts the timer unit values to SAS time values.

There are exactly 38,400 software timer units per second. The low-order bit in a timer unit value represents approximately 26.041667 microseconds.

Use the TU*w*. informat to read timer unit values that are produced by an IBM mainframe on other operating environments.

Examples

```
input btime tu4.;
```

Data Line*	Results
`----+----1----+`	
`8FC7A9BC`	**`62818.411563`**

* The data line is a hexadecimal representation of a binary, four-byte timer unit value. Each byte occupies one column of the input field. The result is a SAS time value that corresponds to 5:26:58.41 p.m.

VAXRB*w.d* Informat

Reads real binary (floating-point) data in VMS format.

Category: Numeric

Syntax

VAXRB*w.d*

Syntax Description

w
specifies the width of the input field.
Default: 4
Range: 2–8

d
specifies the power of 10 by which to divide the value. This argument is optional.
Range: 0–10

Details

Use the VAXRB*w.d* informat to read floating-point data from VMS files on other operating environments.

Comparisons

If you use SAS that is running under VMS, the VAXRB*w.d* and the RB*w.d* informats are identical.

See Also

Informat:
"RB*w.d* Informat" on page 1367

VMSZN*w.d* Informat

Reads VMS and MicroFocus COBOL zoned numeric data.

Category: Numeric

Width range: 1 to 32

Default width: 1

Syntax

VMSZN*w.d*

w
specifies the width of the output field.

d
specifies the number of digits to the right of the decimal point in the numeric value. This argument is optional.

Details

The VMSZN*w.d* informat is similar to the ZD*w.d* informat. Both read a string of ASCII digits, and the last digit is a special character denoting the magnitude of the last digit and the sign of the entire number. The difference between the VMSZN*w.d* informat and the ZD*w.d* informat is in the special character used for the last digit. The following table shows the special characters used by the VMSZN*w.d* informat.

Desired Digit	Special Character	Desired Digit	Special Character
0	0	-0	p
1	1	-1	q
2	2	-2	r
3	3	-3	s
4	4	-4	t
5	5	-5	u
6	6	-6	v
7	7	-7	w
8	8	-8	x
9	9	-9	y

Data formatted using the VMSZN*w.d* informat are ASCII strings.

Examples

```
input @1 vmszn4.;
```

Data line	Results
`------+------1`	
`1234`	`1234`
`123t`	`-1234`

See Also

Format:

"VMSZN*w*.*d* Format" on page 253

Informat:

"ZD*w*.*d* Informat" on page 1414

WEEKU*w*. Informat

Reads the format of the number-of-week value within the year and returns a SAS date value by using the U algorithm.

Category: Date and Time

Syntax

WEEKU*w*.

Syntax Description

w

specifies the width of the input field.

Default: 11

Range: 3–200

Details

The WEEKU*w*. informat reads the format of the number-of-week within the year, and then returns a SAS date value by using the U algorithm. If the input does not contain a year expression, then WEEKU*w*. uses the current year as the year expression, which is the default. If the input does not contain a day expression, then WEEKU*w*. uses the first day of the week as the day expression, which is the default.

The U Algorithm calculates the SAS date value using the number-of-week value within the year (Sunday is considered the first day of the week). The number-of-week value is represented as a decimal number in the range 0–53, with a leading zero and maximum value of 53. For example, the fifth week of the year would be represented as 05.

The inputs to the WEEKU*w*. informat are the same date for the following example. The current year is 2003.

Widths	Formats	Examples
3-4	Www	w01
5-6	yyWww	03W01
7-8	yyWwwdd	03W0101
9-10	yyyyWwwdd	2003W0101
11-200	yyyy-Www-dd	2003-W01-01

Comparisons

The WEEKU*w*. informat reads the number-of-week value within the year. Sunday is the first day of the week, as a decimal number in the range 0–53, with a leading zero. The WEEKV*w*. informat reads the number-of-week value as a decimal number in the range 01–53. Weeks begin on a Monday and week 1 of the year is the week that includes both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. The WEEKW*w*. informat reads the week-number-of-year value as a decimal number in the range 00–53, with Monday as the first day of week 1.

Examples

The current year is 2003 in the following examples.

Statements	Results
	----+----1----+
`v=input('W01',weeku3.);`	
`w=input('03W01',weeku5.);`	
`x=input('03W0101',weeku7.);`	
`y=input('2003W0101',weeku9.);`	
`z=input('2003-W01-01',weeku11.);`	
`put v;`	15710
`put w;`	15710
`put x;`	15710
`put y;`	15710
`put z;`	15710

See Also

Formats:

"WEEKU*w*. Format" on page 259
"WEEKV*w*. Format" on page 261
"WEEKW*w*. Format" on page 263

Functions:

"WEEK Function" on page 1226

Informats:

"WEEKV*w*. Informat" on page 1402
"WEEKW*w*. Informat" on page 1404

WEEKV*w*. Informat

Reads the format of the number-of-week value within the year and returns a SAS date value using the V algorithm.

Category: Date and Time

Syntax

WEEKV*w*.

Syntax Description

w

specifies the width of the input field.

Default: 11

Range: 3–200

Details

The WEEKV*w*. informat reads a format of the number-of-week value. If the input does not contain a year expression, WEEKV*w*. uses the current year as the year expression, which is the default. If the input does not contain a day expression, WEEKV*w*. uses the first day of the week as the day expression, which is the default.

The V algorithm calculates the SAS date value. The number-of-week value is represented as a decimal number in the range 01–53, with a leading zero and maximum value of 53. Weeks begin on a Monday and week 1 of the year is the week that includes both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. For example, the fifth week of the year would be represented as 06.

The inputs to the WEEKV*w*. informat are the same date for the following example. The current year is 2003.

Widths	Formats	Examples
3-4	Www	w01
5-6	yyWww	03W01
7-8	yyWwwdd	03W0101
9-10	yyyyWwwdd	2003W0101
11-200	yyyy-Www-dd	2003-W01-01

Comparisons

The WEEKU*w*. informat reads the number-of-week value within the year. Sunday is the first day of the week, as a decimal number in the range 0–53, with a leading zero. The WEEKV*w*. informat reads the number-of-week value as a decimal number in the range 01–53. Weeks begin on a Monday and week 1 of the year is the week that includes both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. The WEEKW*w*. informat reads the week-number-of-year value as a decimal number in the range 00–53, with Monday as the first day of week 1.

Examples

The current year is 2003 in the following examples.

Statements	Results
	`----+----1----+`
`v=input('W01',weekv3.);`	
`w=input('03W01',weekv5.);`	
`x=input('03W0101',weekv7.);`	
`y=input('2003W0101',weekv9.);`	
`z=input('2003-W01-01',weekv11.);`	
`put v;`	`15704`
`put w;`	`15704`
`put x;`	`15704`
`put y;`	`15704`
`put z;`	`15704`

See Also

Formats:

"WEEKU*w*. Format" on page 259

"WEEKV*w*. Format" on page 261

"WEEKW*w*. Format" on page 263

Functions:

"WEEK Function" on page 1226

Informats:

"WEEKU*w*. Informat" on page 1400

"WEEKW*w*. Informat" on page 1404

WEEKW*w*. Informat

Reads the format of the number-of-week value within the year and returns a SAS date value using the W algorithm.

Category: Date and Time

Syntax

WEEKW*w*.

Syntax Description

w
specifies the width of the input field.

Default: 11

Range: 3–200

Details

The WEEKW*w*. informat reads a format of the number-of-week value. If the input does not contain a year expression, the WEEKW*w*. informat uses the current year as the year expression, which is the default. If the input does not contain a day expression, the WEEKW*w*. informat uses the first day of the week as the day expression, which is the default. Algorithm W calculates the SAS date value using the number of the week within the year (Monday is considered the first day of the week). The number-of-week value is represented as a decimal number in the range 0–53, with a leading zero and maximum value of 53. For example, the fifth week of the year would be represented as 05.

The inputs to the WEEKW*w*. informat are the same date for the following example. The current year is 2003.

Widths	Formats	Examples
3-4	Www	w01
5-6	yyWww	03W01
7-8	yyWwwdd	03W0101
9-10	yyyyWwwdd	2003W0101
11-200	yyyy-Www-dd	2003-W01-01

Comparisons

The WEEKU*w*. informat reads the number-of-week value within the year. Sunday is the first day of the week, as a decimal number in the range 0–53, with a leading zero. The WEEKV*w*. informat reads the number-of-week value as a decimal number in the range 01–53. Weeks begin on a Monday and week 1 of the year is the week that includes both January 4th and the first Thursday of the year. If the first Monday of January is the 2nd, 3rd, or 4th, the preceding days are part of the last week of the preceding year. The WEEKW*w*. informat reads the week-number-of-year value as a decimal number in the range 00–53, with Monday as the first day of week 1.

Examples

The current year is 2003 in the following examples.

Statements	Results
	`----+----1----+`
`v=input('W01',weekw3.);`	
`w=input('03W01',weekw5.);`	
`x=input('03W0101',weekw7.);`	
`y=input('2003W0101',weekw9.);`	
`z=input('2003-W01-01',weekw11.);`	
`put v;`	`15711`
`put w;`	`15711`
`put x;`	`15711`
`put y;`	`15711`
`put z;`	`15711`

See Also

Formats:

"WEEKU*w*. Format" on page 259
"WEEKV*w*. Format" on page 261
"WEEKW*w*. Format" on page 263

Function:

"WEEK Function" on page 1226

Informats:

"WEEKU*w*. Informat" on page 1400
"WEEKV*w*. Informat" on page 1402

w.d Informat

Reads standard numeric data.

Category: Numeric

Alias: BEST*w.d*, D*w.d*, E*w.d*, F*w.d*

Syntax

w.d

Syntax Description

w
specifies the width of the input field.
Range: 1–32

d
specifies the power of 10 by which to divide the value. If the data contain decimal points, the *d* value is ignored. This argument is optional.
Range: 0–31

Details

The *w.d* informat reads numeric values that are located anywhere in the field. Blanks can precede or follow a numeric value with no effect. A minus sign with no separating blank should immediately precede a negative value. The *w.d* informat reads values with decimal points and values in scientific E-notation, and it interprets a single period as a missing value.

Comparisons

- The *w.d* informat is identical to the BZ*w.d* informat, except that the *w.d* informat ignores trailing blanks in the numeric values. To read trailing blanks as 0s, use the BZ*w.d* informat.
- The *w.d* informat can read values in scientific E-notation exactly as the E*w.d* informat does.

Examples

```
input @1 x 6. @10 y 6.2;
put x @7 y;
```

Data Line	Results	
----+----1----+----+		
23 2300	23	23
23 2300	23	0
23 -2300	23	-23
23.0 23.	23	23
2.3E1 2.3	23	2.3
-23 0	-23	.

YMDDTTM*w.d* Informat

Reads datetime values in the form *<yy>yy-mm-dd hh:mm:ss.ss*, where special characters such as a hyphen (-), period (.), slash (/), or colon (:) are used to separate the year, month, day, hour, minute, and seconds; the year can be either 2 or 4 digits.

Category: Date and Time

Alignment: right

Syntax

YMDDTTM*w.d*

Syntax Description

w

specifies the width of the output field.

Default: 19

Range: 13–40

d

specifies the number of digits to the right of the decimal point in the seconds value. The digits to the right of the decimal point specify a fraction of a second. This argument is optional.

Default: 0

Range: 0–39

Details

The YMDDTTMw.d format reads SAS datetime values in the following form:

<yy>yy-mm-dd hh:*mm*:<*ss*<*.ss*>>

The following list explains the datetime variables:

yy or *yyyy*
specifies a two- or four-digit integer that represents the year.

mm
is an integer from 01 through 12 that represents the month.

dd
is an integer from 01 through 31 that represents the day of the month.

hh
is the number of hours ranging from 00 through 23.

mm
is the number of minutes ranging from 00 through 59.

ss.*ss*
is the number of seconds ranging from 00 through 59 with the fraction of a second following the decimal point.

Requirement: If a fraction of a second is specified, the decimal point can be represented only by a period and is required.

- or :
represents one of several special characters, such as the slash (/), dash (-), colon (:), or a blank character that can be used to separate date and time components. Special characters can be used as separators between any date or time component and between the date and the time.

Comparisons

The YMDDTTM*w.d* informat reads datetime values with required separators in the form *<yy>yy-mm-dd*|*hh:mm:ss.ss*.

The MDYAMPM*w.d* in format reads datetime values with optional separators in the form *mm-dd-yy<yy> hh:mm:ss.ss* AM | PM, and requires a space between the date and the time.

The DATETIME*w.d* informat reads datetime values with optional separators in the form *dd-mmm-yy<yy> hh:mm:ss.ss* AM|PM, and the date and time can be separated by a special character.

Examples

```
input @1 dt ymddttm24.;
```

Data Line	Results
2008-03-16 11:23:07.4	1521285787.4
2008 03 16 11 23 07.4	1521285787.4
08.3.16/11:23	1521285780

See Also

Informats:

"DATETIME*w*. Informat" on page 1324

"MDYAMPM*w.d* Informat" on page 1347

YYMMDD*w*. Informat

Reads date values in the form *yymmdd* or *yyyymmdd*.

Category: Date and Time

Syntax

YYMMDD*w*.

Syntax Description

w

specifies the width of the input field.

Default: 6

Range: 6–32

Details

The date values must be in the form *yymmdd* or *yyyymmdd*, where

yy or *yyyy*

is a two-digit or four-digit integer that represents the year.

mm

is an integer from 01 through 12 that represents the month of the year.

dd
is an integer from 01 through 31 that represents the day of the month.

You can separate the year, month, and day values by blanks or by special characters. However, if delimiters are used, place them between all the values. You can also place blanks before and after the date. Make sure the width of the input field allows space for blanks and special characters.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input calendar_date yymmdd10.;
```

Data Line	Results
`----+----1----+`	
`050316`	`16511`
`05/03/16`	`16511`
`05 03 16`	`16511`
`2005-03-16`	`16511`

See Also

Formats:
- "DATE*w*. Format" on page 151
- "DDMMYY*w*. Format" on page 157
- "MMDDYY*w*. Format" on page 196
- "YYMMDD*w*. Format" on page 273

Functions:
- "DAY Function" on page 637
- "MDY Function" on page 924
- "MONTH Function" on page 936
- "YEAR Function" on page 1233

Informats:
- "DATE*w*. Informat" on page 1322
- "DDMMYY*w*. Informat" on page 1326
- "MMDDYY*w*. Informat" on page 1349

System Option:
- "YEARCUTOFF= System Option" on page 2058

YYMMN*w*. Informat

Reads date values in the form *yyyymm* or *yymm*.

Category: Date and Time

Syntax

YYMMN*w*.

Syntax Description

w
specifies the width of the input field.

Default: 4

Range: 4–6

Details

The date values must be in the form *yyyymm* or *yymm*, where

yy or *yyyy*
is a two-digit or four-digit integer that represents the year.

mm
is a two-digit integer that represents the month.

The *N* in the informat name must be used and indicates that you cannot separate the year and month values by blanks or by special characters. SAS automatically adds a day value of 01 to the value to make a valid SAS date variable.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input date1 yymmn6.;
```

Data Line	Results
`----+----1----+`	
`200508`	`16649`

See Also

Formats:

"DATE*w*. Format" on page 151
"DDMMYY*w*. Format" on page 157
"YYMMDD*w*. Format" on page 273
"YYMM*w*. Format" on page 270
"YYMON*w*. Format" on page 276

Functions:

"DAY Function" on page 637
"MONTH Function" on page 936
"MDY Function" on page 924
"YEAR Function" on page 1233
Informats:
"DATE*w*. Informat" on page 1322
"DDMMYY*w*. Informat" on page 1326
"MMDDYY*w*. Informat" on page 1349
"YYMMDD*w*. Informat" on page 1410
System Option:
"YEARCUTOFF= System Option" on page 2058

YYQ*w*. Informat

Reads quarters of the year in the form *yy*Q*q* or *yyyy*Q*q*.

Category: Date and Time

Syntax

YYQ*w*.

Syntax Description

w
specifies the width of the input field.

Default: 6 (For SAS version 6, the default is 4.)

Range: 4–32 (For SAS version 6, the range is 4–6.)

Details

The quarter must be in the form *yy*Q*q* or *yyyy*Q*q*, where

yy or *yyyy*
is an integer that represents the two-digit or four-digit year.

q
is an integer (1, 2, 3, or 4) that represents the quarter of the year. You can also represent the quarter as 01, 02, 03, or 04.

The letter Q must separate the year value and the quarter value. The year value, the letter Q, and the quarter value cannot be separated by blanks. A value that is read with YYQ*w*. produces a SAS date value that corresponds to the first day of the specified quarter.

Note: SAS interprets a two-digit year as belonging to the 100-year span that is defined by the YEARCUTOFF= system option. △

Examples

```
input quarter yyq9.;
```

Data Line	Results
----+----1----+	
05Q2	16527
05Q02	16527
2005Q02	16527

See Also

Functions:

"QTR Function" on page 1063

"YEAR Function" on page 1233

"YYQ Function" on page 1237

System Option:

"YEARCUTOFF= System Option" on page 2058

ZD*w.d* Informat

Reads zoned decimal data.

Category: Numeric

See: ZD*w.d* Informat in the documentation for your operating environment.

Syntax

ZD*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 1

Range: 1–32

d

specifies the power of 10 by which to divide the value. This argument is optional.

Range: 1–31

Details

The ZD*w.d* informat reads zoned decimal data in which every digit requires one byte and in which the last byte contains the value's sign along with the last digit.

Note: Different operating environments store zoned decimal values in different ways. However, ZD*w.d* reads zoned decimal values with consistent results if the values are created in the same type of operating environment that you use to run SAS. △

You can enter positive values in zoned decimal format from a personal computer. Some keying devices enable you to enter negative values by overstriking the last digit with a minus sign.

Comparisons

- Like the *w.d* informat, the ZD*w.d* informat reads data in which every digit requires one byte. Use ZDV*w.d* or ZD*w.d* to read zoned decimal data in which the last byte contains the last digit and the sign.
- The ZD*w.d* informat functions like the ZDV*w.d* informat with one exception: ZDV*w.d* validates the input string and disallows invalid data.
- The following table compares the zoned decimal informat with notation in several programming languages:

Language	Zoned Decimal Notation
SAS	ZD3.
PL/I	PICTURE'99T'
COBOL	DISPLAY PIC S 999
IBM assembler	ZL3

Examples

```
input @1 x zd4.;
```

Data Line*	Results
`----+----1`	
F0F1F2C8	**128**

* The data line contains a hexadecimal representation of a binary number that is stored in zoned decimal format on an IBM mainframe computer system. Each byte occupies one column of the input field.

See Also

Informats:

"*w.d* Informat" on page 1407

"ZDV*w.d* Informat" on page 1416

ZDB*w.d* Informat

Reads zoned decimal data in which zeros have been left blank.

Category: Numeric

See: ZDB*w.d* Informat in the documentation for your operating environment.

Syntax

ZDB*w.d*

Syntax Description

w

specifies the width of the input field.

Default: 1

Range: 1–32

d

specifies the power of 10 by which to divide the value. This argument is optional.

Range: 0–31

Details

The ZDB*w.d* informat reads zoned decimal data that are produced in IBM 1410, 1401, and 1620 form, where 0s are left blank rather than being punched.

Examples

```
input @1 x zdb3.;
```

Data Line*	Results
`----+----1`	
F140C2	**102**

* The data line contains a hexadecimal representation of a binary number that is stored in zoned decimal form, including the codes for spaces, on an IBM mainframe operating environment. Each byte occupies one column of the input field.

ZDV*w.d* Informat

Reads and validates zoned decimal data.

Category: Numeric

Syntax

ZDV*w*.*d*

Syntax Description

w
specifies the width of the input field.

Default: 1

Range: 1–32

d
specifies the power of 10 by which to divide the value. This argument is optional.

Range: 1–31

Details

The ZDV*w*.*d* informat reads data in which every digit requires one byte and in which the last byte contains the value's sign along with the last digit. It also validates the input string and disallows invalid data.

ZDV*w*.*d* is dependent on the operating environment. For example, on IBM mainframes, ZDV*w*.*d* requires an F for all high-order nibbles except the last. (In contrast, the ZD*w*.*d* informat ignores the high-order nibbles for all bytes except for the nibbles that are associated with the sign.) The last high-order nibble accepts values ranging from A-F, where A, C, E, and F are positive values and B and D are negative values. The low-order nibble on IBM mainframes must be a numeric digit that ranges from 0-9, as with ZD.

Note: Different operating environments store zoned decimal values in different ways. However, the ZDV*w*.*d* informat reads zoned decimal values with consistent results if the values are created in the same type of operating environment that you use to run SAS. △

Comparisons

The ZDV*w*.*d* informat functions like the ZD*w*.*d* informat with one exception: ZDV*w*.*d* validates the input string and disallows invalid data.

Examples

```
input @1 test zdv4.;
```

Data Line*	Results
`----+----1`	
`F0F1F2C8`	**`128`**

* The data line contains a hexadecimal representation of a binary number stored in zoned decimal form. The example was run on an IBM mainframe. The results might vary depending on your operating environment.

See Also

Informats:

"*w.d* Informat" on page 1407

"ZD*w.d* Informat" on page 1414

Informats Documented in Other Base SAS Publications

The main references for SAS formats are *SAS Language Reference: Dictionary* and the *SAS National Language Support (NLS): Reference Guide*. See the documentation for your operating environment for host-specific information about formats.

SAS National Language Support: Reference Guide

Table 5.5 Summary of NLS Formats by Category

Category	Informats for NLS	Description
BIDI text handling	$LOGVSw. Informat	Reads a character string that is in left-to-right logical order, and then converts the character string to visual order.
	$LOGVSRw. Informat	Reads a character string that is in right-to-left logical order, and then converts the character string to visual order.
	$VSLOGw. Informat	Reads a character string that is in visual order, and then converts the character string to left-to-right logical order.
	$VSLOGRw. Informat	Reads a character string that is in visual order, and then converts the character string to right-to-left logical order.
Character	$REVERJw. Informat	Reads character data from right to left and preserves blanks.
	$REVERSw. Informat	Reads character data from right to left, and then left aligns the text.
	$UCS2Bw. Informat	Reads a character string that is encoded in big-endian, 16-bit, UCS2, Unicode encoding, and then converts the character string to the encoding of the current SAS session.
	$UCS2BEw. Informat	Reads a character string that is in the encoding of the current SAS session and then converts the character string to big-endian, 16-bit, UCS2, Unicode encoding.
	$UCS2Lw. Informat	Reads a character string that is encoded in little-endian, 16-bit, UCS2, Unicode encoding, and then converts the character string to the encoding of the current SAS session.
	$UCS2LEw. Informat	Reads a character string that is in the encoding of the current SAS session and then converts the character string to little-endian, 16-bit, UCS2, Unicode encoding.

Category	Informats for NLS	Description
	$UCS2Xw. Informat	Reads a character string that is encoded in 16-bit, UCS2, Unicode encoding, and then converts the character string to the encoding of the current SAS session.
	$UCS2XEw. Informat	Reads a character string that is in the encoding of the current SAS session and then converts the character string to 16-bit, UCS2, Unicode encoding.
	$UCS4Bw. Informat	Reads a character string that is encoded in big-endian, 32-bit, UCS4, Unicode encoding, and then converts the character string to the encoding of the current SAS session.
	$UCS4Lw. Informat	Reads a character string that is encoded in little-endian, 32-bit, UCS4, Unicode encoding, and then converts the character string to the encoding of the current SAS session.
	$UCS4Xw. Informat	Reads a character string that is encoded in 32-bit, UCS4, Unicode encoding, and then converts the character string to the encoding of the current SAS session.
	$UCS4XEw. Informat	Reads a character string that is in the encoding of the current SAS session, and then converts the character string to 32-bit, UCS4, Unicode encoding.
	$UESCw. Informat	Reads a character string that is encoded in UESC representation, and then converts the character string to the encoding of the current SAS session.
	$UESCEw. Informat	Reads a character string that is in the encoding of the current SAS session, and then converts the character string to UESC representation.
	$UNCRw. Informat	Reads an NCR character string, and then converts the character string to the encoding of the current SAS session.
	UNCREw. Informat	Reads a character string in the encoding of the current SAS session, and then converts the character string to NCR.
	$UPARENw. Informat	Reads a character string that is encoded in UPAREN representation, and then converts the character string to the encoding of the current SAS session.
	$UPARENEw. Informat	Reads a character string that is in the encoding of the current SAS session, and then converts the character string to UPAREN representation.
	$UPARENPw. Informat	Reads a character string that is encoded in UPAREN representation, and then converts the character string to the encoding of the current SAS session, with national characters remaining in the encoding of the UPAREN representation.
	$UTF8Xw. informat	Reads a character string that is encoded in UTF-8, and then converts the character string to the encoding of the current SAS session.

Category	Informats for NLS	Description
DBCS	$KANJIw. Informat	Removes shift code data from DBCS data.
	$KANJIXw. Informat	Adds shift-code data to DBCS data.
Date and Time	JDATEYMDw. Informat	Reads Japanese kanji date values in the format *yymmmdd* or *yyyymmmdd*.
	JNENGOw. Informat	Reads Japanese kanji date values in the form *yymmdd*.
	MINGUOw. Informat	Reads dates in Taiwanese format.
	NENGOw. Informat	Reads Japanese date values in the form *eyymmdd*.
	NLDATEw. Informat	Reads the date value in the specified locale, and then converts the date value to the local SAS date value.
	NLDATMw. Informat	Reads the datetime value of the specified locale, and then converts the datetime value to the local SAS datetime value.
	NLTIMAPw. Informat	Reads the time value and uses a.m. and p.m. in the specified locale, and then converts the time value to the local SAS time value.
	NLTIMEw. Informat	Reads the time value in the specified locale, and then converts the time value to the local SAS time value.
Hebrew text handling	$CPTDWw. Informat	Reads a character string that is in Hebrew DOS (cp862) encoding, and then converts the character string to Windows (cp1255) encoding.
	$CPTWDw. Informat	Reads a character string that is in Windows (cp1255) encoding, and then converts the character string to Hebrew DOS (cp862) encoding.
Numeric	EUROw.d Informat	Reads numeric values, removes embedded characters in European currency, and reverses the comma and decimal point.
	EUROXw.d Informat	Reads numeric values and removes embedded characters in European currency.
	NLMNIAEDw.d Informat	Reads the monetary format of the international expression for the United Arab Emirates.
	NLMNIAUDw.d Informat	Reads the monetary format of the international expression for Australia.
	NLMNIBGNw.d Informat	Reads the monetary format of the international expression for Bulgaria.
	NLMNIBRLw.d Informat	Reads the monetary format of the international expression for Brazil.
	NLMNICADw.d Informat	Reads the monetary format of the international expression for Canada.
	NLMNICHFw.d Informat	Reads the monetary format of the international expression for Liechtenstein and Switzerland.
	NLMNICNYw.d Informat	Reads the monetary format of the international expression for China.

Category	Informats for NLS	Description
	NLMNICZKw.d Informat	Reads the monetary format of the international expression for the Czech Republic.
	NLMNIDKKw.d Informat	Reads the monetary format of the international expression for Denmark, Faroe Island, and Greenland.
	NLMNIEEKw.d Informat	Reads the monetary format of the international expression for Estonia.
	NLMNIEGPw.d Informat	Reads the monetary format of the international expression for Egypt.
	NLMNIEURw.d Informat	Reads the monetary format of the international expression for Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Slovenia, and Spain.
	NLMNIGBPw.d Informat	Reads the monetary format of the international expression for the United Kingdom.
	NLMNIHKDw.d Informat	Reads the monetary format of the international expression for Hong Kong.
	NLMNIHRKw.d Informat	Reads the monetary format of the international expression for Croatia.
	NLMNIHUFw.d Informat	Reads the monetary format of the international expression for Hungary.
	NLMNIIDRw.d Informat	Reads the monetary format of the international expression for Indonesia.
	NLMNIILSw.d Informat	Reads the monetary format of the international expression for Israel.
	NLMNIINRw.d Informat	Reads the monetary format of the international expression for India.
	NLMNIJPYw.d Informat	Reads the monetary format of the international expression for Japan.
	NLMNIKRWw.d Informat	Reads the monetary format of the international expression for South Korea.
	NLMNILTLw.d Informat	Reads the monetary format of the international expression for Lithuania.
	NLMNILVLw.d Informat	Reads the monetary format of the international expression for Latvia.
	NLMNIMOPw.d Informat	Reads the monetary format of the international expression for Macau.
	NLMNIMXNw.d Informat	Reads the monetary format of the international expression for Mexico.
	NLMNIMYRw.d Informat	Reads the monetary format of the international expression for Malaysia.
	NLMNINOKw.d Informat	Reads the monetary format of the international expression for Norway.
	NLMNINZDw.d Informat	Reads the monetary format of the international expression for New Zealand.

Category	Informats for NLS	Description
	NLMNIPLNw.d Informat	Reads the monetary format of the international expression for Poland.
	NLMNIRUBw.d Informat	Reads the monetary format of the international expression for Russia.
	NLMNISEKw.d Informat	Reads the monetary format of the international expression for Sweden.
	NLMNISGDw.d Informat	Reads the monetary format of the international expression for Singapore.
	NLMNITHBw.d Informat	Reads the monetary format of the international expression for Thailand.
	NLMNITRYw.d Informat	Reads the monetary format of the international expression for Turkey.
	NLMNITWDw.d Informat	Reads the monetary format of the international expression for Taiwan.
	NLMNIUSDw.d Informat	Reads the monetary format of the international expression for the Caribbean, Puerto Rico and the United States.
	NLMNIZARw.d Informat	Reads the monetary format of the international expression for South Africa.
	NLMNLAEDw.d Informat	Reads the monetary format of the local expression for the United Arab Emirates.
	NLMNLAUDw.d Informat	Reads the monetary format of the local expression for Australia.
	NLMNLBGNw.d Informat	Reads the monetary format of the local expression for Bulgaria.
	NLMNLBRLw.d Informat	Reads the monetary format of the local expression for Brazil.
	NLMNLCADw.d Informat	Reads the monetary format of the local expression for Canada.
	NLMNLCHFw.d Informat	Reads the monetary format of the local expression for Liechtenstein and Switzerland.
	NLMNLCNYw.d Informat	Reads the monetary format of the local expression for China.
	NLMNLCZKw.d Informat	Reads the monetary format of the local expression for the Czech Republic.
	NLMNLDKKw.d Informat	Reads the monetary format of the local expression for Denmark, the Faroe Island, and Greenland.
	NLMNLEEKw.d Informat	Reads the monetary format of the local expression for Estonia.
	NLMNLEGPw.d Informat	Reads the monetary format of the local expression for Egypt.

Category	Informats for NLS	Description
	NLMNLEURw.d Informat	Reads the monetary format of the local expression for Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, and Spain.
	NLMNLGBPw.d Informat	Reads the monetary format of the local expression for the United Kingdom.
	NLMNLHKDw.d Informat	Reads the monetary format of the local expression for Hong Kong.
	NLMNLHRKw.d Informat	Reads the monetary format of the local expression for Croatia.
	NLMNLHUFw.d Informat	Reads the monetary format of the local expression for Hungary.
	NLMNLIDRw.d Informat	Reads the monetary format of the local expression for Indonesia.
	NLMNLILSw.d Informat	Reads the monetary format of the local expression for Israel.
	NLMNLINRw.d Informat	Reads the monetary format of the local expression for India.
	NLMNLJPYw.d Informat	Reads the monetary format of the local expression for Japan.
	NLMNLKRWw.d Informat	Reads the monetary format of the local expression for South Korea.
	NLMNLLTLw.d Informat	Reads the monetary format of the local expression for Lithuania.
	NLMNLLVLw.d Informat	Reads the monetary format of the local expression for Latvia.
	NLMNLMOPw.d Informat	Reads the monetary format of the local expression for Macau.
	NLMNLMXNw.d Informat	Reads the monetary format of the local expression for Mexico.
	NLMNLMYRw.d Informat	Reads the monetary format of the local expression for Malaysia.
	NLMNLNOKw.d Informat	Reads the monetary format of the local expression for Norway.
	NLMNLNZDw.d Informat	Reads the monetary format of the local expression for New Zealand.
	NLMNLPLNw.d Informat	Reads the monetary format of the local expression for Poland.
	NLMNLRUBw.d Informat	Reads the monetary format of the local expression for Russia.
	NLMNLSEKw.d Informat	Reads the monetary format of the local expression for Sweden.
	NLMNLSGDw.d Informat	Reads the monetary format of the local expression for Singapore.

Category	Informats for NLS	Description
	NLMNLTHBw.d Informat	Reads the monetary format of the local expression for Thailand.
	NLMNLTRYw.d Informat	Reads the monetary format of the local expression for Turkey.
	NLMNLTWDw.d Informat	Reads the monetary format of the local expression for Taiwan.
	NLMNLUSDw.d Informat	Reads the monetary format of the local expression for the Caribbean, Puerto Rico, and the United States.
	NLMNLZARw.d Informat	Reads the monetary format of the local expression for South Africa.
	NLMNYw.d Informat	Reads monetary data in the specified locale for the local expression, and then converts the data to a numeric value.
	NLMNYIw.d Informat	Reads monetary data in the specified locale for the international expression, and then converts the data to a numeric value.
	NLNUMw.d Informat	Reads numeric data in the specified locale for local expressions, and then converts the data to a numeric value.
	NLNUMIw.d Informat	Reads numeric data in the specified locale for international expressions, and then converts the data to a numeric value.
	NLPCTw.d Informat	Reads percentage data in the specified locale for local expressions, and then converts the data to a numeric value.
	NLPCTIw.d Informat	Reads percentage data in the specified locale for international expressions, and then converts the data to a numeric value.
	YENw.d Informat	Removes embedded yen signs, commas, and decimal points.

CHAPTER 6

Statements

Definition of Statements

A *SAS statement* is a series of items that can include keywords, SAS names, special characters, and operators. All SAS statements end with a semicolon. A SAS statement either requests SAS to perform an operation or gives information to the system.

This documentation covers two types of SAS statements:

- statements that are used in DATA step programming
- statements that are global in scope and can be used anywhere in a SAS program.

The *Base SAS Procedures Guide* gives detailed descriptions of the SAS statements that are specific to each SAS procedure. *SAS Output Delivery System: User's Guide* gives detailed descriptions of the Output Delivery System (ODS) statements.

DATA Step Statements

Executable and Declarative Statements

DATA step statements are executable or declarative statements that can appear in the DATA step. *Executable statements* result in some action during individual iterations of the DATA step; *declarative statements* supply information to SAS and take effect when the system compiles program statements.

The following tables show the SAS executable and declarative statements that you can use in the DATA step.

Executable Statements

ABORT	IF, Subsetting	PUT, Column
Array Reference	IF-THEN/ELSE	PUT, Formatted
Assignment	INFILE	PUT, List
CALL	INPUT	PUT, Named
CONTINUE	GO TO	PUT
DECLARE	INPUT, Column	PUT, ODS
DELETE	INPUT, Formatted	PUTLOG
DESCRIBE	INPUT, List	REDIRECT
DISPLAY	INPUT, Named	REMOVE
DO	LEAVE	REPLACE
DO, Iterative	LINK	RETURN
DO UNTIL	LIST	SELECT
DO WHILE	LOSTCARD	SET
ERROR	MERGE	STOP
EXECUTE	MODIFY	Sum
FILE	Null	UPDATE
FILE, ODS	OUTPUT	

Declarative Statements

ARRAY	DATALINES4	Labels, Statement
ATTRIB	DROP	LENGTH
BY	END	RENAME
CARDS	FORMAT	RETAIN
CARDS4	INFORMAT	WHERE
DATA	KEEP	WINDOW
DATALINES	LABEL	

DATA Step Statements by Category

In addition to being either executable or declarative, SAS DATA step statements can be grouped into five functional categories:

Table 6.1 Categories of DATA Step Statements

Statements Category	Functionality
Action	□ create and modify variables □ select only certain observations to process in the DATA step □ look for errors in the input data □ work with observations as they are being created
Control	□ skip statements for certain observations □ change the order that statements are executed □ transfer control from one part of a program to another
File-handling	□ work with files used as input to the data set □ work with files to be written by the DATA step
Information	□ give SAS additional information about the program data vector □ give SAS additional information about the data set or data sets that are being created.
Window Display	□ display and customize windows.

The following table lists and briefly describes the DATA step statements by category.

Table 6.2 Categories and Descriptions of DATA Step Statements

Category	Statement	Description
Action	"ABORT Statement" on page 1436	Stops executing the current DATA step, SAS job, or SAS session.
	"Assignment Statement" on page 1447	Evaluates an expression and stores the result in a variable.
	"CALL Statement" on page 1457	Invokes a SAS CALL routine.
	"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476	Declares a hash or hash iterator object; creates an instance of and initializes data for a hash or hash iterator object.
	"DECLARE Statement, Java Object" on page 1483	Declares a Java object; creates an instance of and initializes data for a Java object.
	"DELETE Statement" on page 1486	Stops processing the current observation.
	"DESCRIBE Statement" on page 1487	Retrieves source code from a stored compiled DATA step program or a DATA step view.
	"ERROR Statement" on page 1502	Sets _ERROR_ to 1. A message written to the SAS log is optional.
	"EXECUTE Statement" on page 1503	Executes a stored compiled DATA step program .
	"IF Statement, Subsetting" on page 1581	Continues processing only those observations that meet the condition of the specified expression.
	"LIST Statement" on page 1670	Writes to the SAS log the input data record for the observation that is being processed.
	"LOSTCARD Statement" on page 1676	Resynchronizes the input data when SAS encounters a missing or invalid record in data that has multiple records per observation.
	"Null Statement" on page 1701	Signals the end of data lines or acts as a placeholder.
	"OUTPUT Statement" on page 1704	Writes the current observation to a SAS data set.
	"PUTLOG Statement" on page 1738	Writes a message to the SAS log.
	"REDIRECT Statement" on page 1740	Points to different input or output SAS data sets when you execute a stored program.
	"REMOVE Statement" on page 1741	Deletes an observation from a SAS data set.
	"REPLACE Statement" on page 1745	Replaces an observation in the same location.
	"STOP Statement" on page 1775	Stops execution of the current DATA step.

Category	Statement	Description
	"Sum Statement" on page 1777	Adds the result of an expression to an accumulator variable.
	"WHERE Statement" on page 1792	Selects observations from SAS data sets that meet a particular condition.
Control	"CONTINUE Statement" on page 1464	Stops processing the current DO-loop iteration and resumes processing the next iteration.
	"DO Statement" on page 1491	Specifies a group of statements to be executed as a unit.
	"DO Statement, Iterative" on page 1492	Executes statements between the DO and END statements repetitively, based on the value of an index variable.
	"DO UNTIL Statement" on page 1496	Executes statements in a DO loop repetitively until a condition is true.
	"DO WHILE Statement" on page 1497	Executes statements in a DO-loop repetitively while a condition is true.
	"END Statement" on page 1500	Ends a DO group or SELECT group processing.
	"GO TO Statement" on page 1579	Directs program execution immediately to the statement label that is specified and, if followed by a RETURN statement, returns execution to the beginning of the DATA step.
	"IF-THEN/ELSE Statement" on page 1582	Executes a SAS statement for observations that meet specific conditions.
	"Labels, Statement" on page 1651	Identifies a statement that is referred to by another statement.
	"LEAVE Statement" on page 1653	Stops processing the current loop and resumes with the next statement in the sequence.
	"LINK Statement" on page 1669	Directs program execution immediately to the statement label that is specified and, if followed by a RETURN statement, returns execution to the statement that follows the LINK statement.
	"RETURN Statement" on page 1752	Stops executing statements at the current point in the DATA step and returns to a predetermined point in the step.
	"SELECT Statement" on page 1760	Executes one of several statements or groups of statements.
File-handling	"BY Statement" on page 1452	Controls the operation of a SET, MERGE, MODIFY, or UPDATE statement in the DATA step and sets up special grouping variables.
	"CARDS Statement" on page 1458	Specifies that data lines follow.
	"CARDS4 Statement" on page 1458	Specifies that data lines that contain semicolons follow.
	"DATA Statement" on page 1465	Begins a DATA step and provides names for any output SAS data sets, views, or programs.

Category	Statement	Description
	"DATALINES Statement" on page 1474	Specifies that data lines follow.
	"DATALINES4 Statement" on page 1475	Indicates that data lines that contain semicolons follow.
	"FILE Statement" on page 1503	Specifies the current output file for PUT statements.
	"INFILE Statement" on page 1591	Specifies an external file to read with an INPUT statement.
	"INPUT Statement" on page 1617	Describes the arrangement of values in the input data record and assigns input values to the corresponding SAS variables.
	"INPUT Statement, Column" on page 1632	Reads input values from specified columns and assigns them to the corresponding SAS variables.
	"INPUT Statement, Formatted" on page 1635	Reads input values with specified informats and assigns them to the corresponding SAS variables.
	"INPUT Statement, List" on page 1639	Scans the input data record for input values and assigns them to the corresponding SAS variables.
	"INPUT Statement, Named" on page 1645	Reads data values that appear after a variable name that is followed by an equal sign and assigns them to corresponding SAS variables.
	"MERGE Statement" on page 1679	Joins observations from two or more SAS data sets into a single observation.
	"MODIFY Statement" on page 1684	Replaces, deletes, and appends observations in an existing SAS data set in place but does not create an additional copy.
	"PUT Statement" on page 1708	Writes lines to the SAS log, to the SAS output window, or to an external location that is specified in the most recent FILE statement.
	"PUT Statement, Column" on page 1724	Writes variable values in the specified columns in the output line.
	"PUT Statement, Formatted" on page 1727	Writes variable values with the specified format in the output line.
	"PUT Statement, List" on page 1731	Writes variable values and the specified character strings in the output line.
	"PUT Statement, Named" on page 1736	Writes variable values after the variable name and an equal sign.
	"SET Statement" on page 1764	Reads an observation from one or more SAS data sets.
	"UPDATE Statement" on page 1787	Updates a master file by applying transactions.
Information	"ARRAY Statement" on page 1440	Defines the elements of an array.
	"Array Reference Statement" on page 1445	Describes the elements in an array to be processed.

Category	Statement	Description
	"ATTRIB Statement" on page 1448	Associates a format, informat, label, and length with one or more variables.
	"DROP Statement" on page 1499	Excludes variables from output SAS data sets.
	"FORMAT Statement" on page 1576	Associates formats with variables.
	"INFORMAT Statement" on page 1614	Associates informats with variables.
	"KEEP Statement" on page 1648	Specifies the variables to include in output SAS data sets.
	"LABEL Statement" on page 1650	Assigns descriptive labels to variables.
	"LENGTH Statement" on page 1654	Specifies the number of bytes for storing variables.
	"MISSING Statement" on page 1682	Assigns characters in your input data to represent special missing values for numeric data.
	"RENAME Statement" on page 1743	Specifies new names for variables in output SAS data sets.
	"RETAIN Statement" on page 1747	Causes a variable that is created by an INPUT or assignment statement to retain its value from one iteration of the DATA step to the next.
Window Display	"DISPLAY Statement" on page 1488	Displays a window that is created with the WINDOW statement.
	"WINDOW Statement" on page 1797	Creates customized windows for your applications.

Global Statements

Definition

Global statements generally provide information to SAS, request information or data, move between different modes of execution, or set values for system options. Other global statements (ODS statements) deliver output in a variety of formats, such as in Hypertext Markup Language (HTML). You can use global statements anywhere in a SAS program. Global statements are not executable; they take effect as soon as SAS compiles program statements.

Other SAS software products have additional global statements that are used with those products. For information, see the SAS documentation for those products.

Global Statements by Category

The following table lists and describes SAS global statements, organized by function into eight categories:

Table 6.3 Global Statements by Category

Statements Category	Functionality
Data Access	associate reference names with SAS libraries, SAS catalogs, external files and output devices, and access remote files.
Log Control	alter the appearance of the SAS log.
ODS: Output Control	choose objects to send to output destinations; edit the output format.
ODS: SAS Formatted	apply default styles to SAS specific entities such as a SAS data set, SAS output listing, or a SAS document.
ODS: Third-Party Formatted	apply styles to the output objects that are used by applications outside of SAS.
Operating Environment	access the operating environment directly.
Output Control	add titles and footnotes to your SAS output; deliver output in a variety of formats.
Program Control	govern the way SAS processes your SAS program.

The following table provides brief descriptions of SAS global statements. For more detailed information, see the individual statements.

Table 6.4 Categories and Descriptions of Global Statements

Category	Statement	Description
Data Access	"CATNAME Statement" on page 1459	Logically combines two or more catalogs into one by associating them with a catref (a shortcut name); clears one or all catrefs; lists the concatenated catalogs in one concatenation or in all concatenations.
	"FILENAME Statement" on page 1520	Associates a SAS fileref with an external file or an output device, disassociates a fileref and external file, or lists attributes of external files.
	"FILENAME Statement, CATALOG Access Method" on page 1526	Enables you to reference a SAS catalog as an external file.
	"FILENAME, CLIPBOARD Access Method" on page 1529	Enables you to read text data from and write text data to the clipboard on the host computer.
	"FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532	Enables you to send electronic mail programmatically from SAS using the SMTP (Simple Mail Transfer Protocol) e-mail interface.
	"FILENAME Statement, FTP Access Method" on page 1542	Enables you to access remote files by using the FTP protocol.
	"FILENAME Statement, SFTP Access Method" on page 1554	Enables you to access remote files by using the SFTP protocol.
	"FILENAME Statement, SOCKET Access Method" on page 1559	Enables you to read from or write to a TCP/IP socket.
	"FILENAME Statement, URL Access Method" on page 1563	Enables you to access remote files by using the URL access method.
	"FILENAME Statement, WebDAV Access Method" on page 1567	Enables you to access remote files by using the WebDAV protocol.
	"LIBNAME Statement" on page 1656	Associates or disassociates a SAS library with a libref (a shortcut name), clears one or all librefs, lists the characteristics of a SAS library, concatenates SAS libraries, or concatenates SAS catalogs.
	"LIBNAME Statement for WebDAV Server Access" on page 1665	Associates a libref with a SAS library and enables access to a WebDAV (Web-based Distributed Authoring And Versioning) server.
Log Control	"Comment Statement" on page 1462	Specifies the purpose of the statement or program.
	"PAGE Statement" on page 1707	Skips to a new page in the SAS log.

Category	Statement	Description
	"SKIP Statement" on page 1775	Creates a blank line in the SAS log.
Operating Environment	"X Statement" on page 1808	Issues an operating-environment command from within a SAS session.
Output Control	"FOOTNOTE Statement" on page 1573	Writes up to 10 lines of text at the bottom of the procedure or DATA step output.
	"TITLE Statement" on page 1779	Specifies title lines for SAS output.
Program Control	"CHECKPOINT EXECUTE_ALWAYS Statement" on page 1462	Indicates to execute the DATA step or PROC step that immediately follows without considering the checkpoint-restart data.
	"DM Statement" on page 1489	Submits SAS Program Editor, Log, Procedure Output or text editor commands as SAS statements.
	"ENDSAS Statement" on page 1501	Terminates a SAS job or session after the current DATA or PROC step executes.
	"%INCLUDE Statement" on page 1584	Brings a SAS programming statement, data lines, or both, into a current SAS program.
	"%LIST Statement" on page 1672	Displays lines that are entered in the current session.
	"LOCK Statement" on page 1673	Acquires and releases an exclusive lock on an existing SAS file.
	"OPTIONS Statement" on page 1703	Specifies or changes the value of one or more SAS system options.
	"RUN Statement" on page 1753	Executes the previously entered SAS statements.
	"%RUN Statement" on page 1754	Ends source statements following a %INCLUDE * statement.
	"SASFILE Statement" on page 1755	Opens a SAS data set and allocates enough buffers to hold the entire file in memory.
	"SYSECHO Statement" on page 1778	Fires a global statement complete event and passes a text string back to the IOM client.

Dictionary

ABORT Statement

Stops executing the current DATA step, SAS job, or SAS session.

Valid: in a DATA step

Category: Action

Type: Executable

See: ABORT Statement in the documentation for your operating environment.

Syntax

ABORT <ABEND | CANCEL <FILE> | RETURN | > <*n*> <NOLIST>;

Without Arguments

If you specify no argument, the ABORT statement produces these results under the following methods of operation:

batch mode and noninteractive mode

- stops processing the current DATA step and writes an error message to the SAS log. Data sets can contain an incomplete number of observations or no observations, depending on when SAS encountered the ABORT statement.
- sets the OBS= system option to 0.
- continues limited processing of the remainder of the SAS job, including executing macro statements, executing system options statements, and syntax checking of program statements.
- creates output data sets for subsequent DATA and PROC steps with no observations.

windowing environment

- stops processing the current DATA step
- creates a data set that contains the observations that are processed before the ABORT statement is encountered
- prints a message to the log that an ABORT statement terminated the DATA step
- continues processing any DATA or PROC steps that follow the ABORT statement.

interactive line mode

stops processing the current DATA step. Any further DATA steps or procedures execute normally.

Arguments

ABEND
causes abnormal termination of the current SAS job or session. Results depend on the method of operation:

- batch mode and noninteractive mode
  - stops processing immediately
  - sends an error message to the SAS log that states that execution was terminated by the ABEND option of the ABORT statement
  - does not execute any subsequent statements or check syntax
  - returns control to the operating environment; further action is based on how your operating environment and your site treat jobs that end abnormally.
- windowing environment and interactive line mode
  - causes your windowing environment and interactive line mode to stop processing immediately and return you to your operating environment.

CANCEL <FILE>
causes the execution of the submitted statements to be canceled. Results depend on the method of operation:

- batch mode and noninteractive mode
  - the entire SAS program and SAS system are terminated
  - an error message is written to the SAS log
- windowing environment and interactive line mode
  - clears only the current submitted program
  - other subsequent submitted programs are not affected
  - an error message is written to the SAS log
- workspace server and stored process server
  - clears only the currently submitted program
  - other subsequent submit calls are not affected
  - an error message is written to the SAS log
- SAS IntrNet application server
  - creates a separate execution for each request and submits the request code. A CANCEL argument in the request code clears the current submitted code but does not terminate the execution or the SAS session.

FILE
when coded as an option to the CANCEL argument in an autoexec file or in a %INCLUDE file, causes only the contents of the autoexec file or %INCLUDE file to be cleared by the ABORT statement. Other submitted source statements will be executed after the autoexec or %INCLUDE file.

Warning: When the ABORT CANCEL FILE option is executed within a %INCLUDE file, all open macros are closed and execution resumes at the next source line of code.

Restriction: The CANCEL argument cannot be submitted using SAS/SHARE, SAS/CONNECT, or SAS/AF.

RETURN
causes the immediate normal termination of the current SAS job or session. Results depend on the method of operation:

- batch mode and noninteractive mode
  - stops processing immediately
  - sends an error message to the SAS log stating that execution was terminated by the RETURN option in the ABORT statement
  - does not execute any subsequent statements or check syntax
  - returns control to your operating environment with a condition code indicating an error
- windowing environment
  - causes your windowing environment and interactive line mode to stop processing immediately and return you to your operating environment.

n
is an integer value that enables you to specify a condition code:

- when used with the CANCEL argument, the value is placed in the SYSINFO automatic macro variable
- when not used with the CANCEL argument, SAS returns the value to the operating environment when the execution stops. The range of values for *n* depends on your operating environment.

NOLIST
suppresses the output of all variables to the SAS log.
Requirement: NOLIST must be the last option in the ABORT statement.

Details

The ABORT statement causes SAS to stop processing the current DATA step. What happens next depends on

- the method you use to submit your SAS statements
- the arguments you use with ABORT
- your operating environment.

The ABORT statement usually appears in a clause of an IF-THEN statement or a SELECT statement that is designed to stop processing when an error condition occurs.

Note: The return code generated by the ABORT statement is ignored by SAS if the system option ERRORABEND is in effect. △

Note: When you execute an ABORT statement in a DATA step, SAS does not use data sets that were created in the step to replace existing data sets with the same name. △

Operating Environment Information: The only difference between the ABEND and RETURN options is that with ABEND further action is based on how your operating environment and site treat jobs that end abnormally. RETURN simply returns a condition code that indicates an error. △

Comparisons

- When you use the SAS windowing environment or interactive line mode, the ABORT statement and the STOP statement both stop processing. The ABORT

statement sets the value of the automatic variable _ERROR_ to 1, and the STOP statement does not.

- In batch or noninteractive mode, the ABORT and STOP statements also have different effects. Both stop processing, but only ABORT sets the value of the automatic variable _ERROR_ to 1. Use the STOP statement, therefore, when you want to stop only the current DATA step and continue processing with the next step.

Examples

This example uses the ABORT statement as part of an IF-THEN statement to stop execution of SAS when it encounters a data value that would otherwise cause a division-by-zero condition.

```
if volume=0 then abort 255;
   density=mass/volume;
```

The *n* value causes SAS to return the condition code 255 to the operating environment when the ABORT statement executes.

See Also

Statement:

"STOP Statement" on page 1775

ARRAY Statement

Defines the elements of an array.

Valid: in a DATA step

Category: Information

Type: Declarative

Syntax

ARRAY *array-name* { *subscript* } <$><*length*>
<*array-elements*> <(*initial-value-list*)>;

Arguments

array-name
specifies the name of the array.

Restriction: *Array-name* must be a SAS name that is not the name of a SAS variable in the same DATA step.

CAUTION:
Using the name of a SAS function as an array name can cause unpredictable results. If you inadvertently use a function name as the name of the array, SAS treats

parenthetical references that involve the name as array references, not function references, for the duration of the DATA step. A warning message is written to the SAS log. △

{subscript}
describes the number and arrangement of elements in the array by using an asterisk, a number, or a range of numbers. *Subscript* has one of these forms:

{*dimension-size(s)*}
specifies the number of elements in each dimension of the array. *Dimension-size* is a numeric representation of either the number of elements in a one-dimensional array or the number of elements in each dimension of a multidimensional array.

Tip: You can enclose the subscript in braces ({}), brackets ([ ])or parentheses (( )).

Example: An array with one dimension can be defined as

```
array simple{3} red green yellow;
```

This ARRAY statement defines an array that is named SIMPLE that groups together three variables that are named RED, GREEN, and YELLOW.

Example: An array with more than one dimension is known as a multidimensional array. You can have any number of dimensions in a multidimensional array. For example, a two-dimensional array provides row and column arrangement of array elements. This statement defines a two-dimensional array with five rows and three columns:

```
array x{5,3} score1-score15;
```

SAS places variables into a two-dimensional array by filling all rows in order, beginning at the upper-left corner of the array (known as row-major order).

{<*lower* :>*upper*<, ...<*lower* :> *upper*>}
are the bounds of each dimension of an array, where *lower* is the lower bound of that dimension and *upper* is the upper bound.

Range: In most explicit arrays, the subscript in each dimension of the array ranges from 1 to *n*, where *n* is the number of elements in that dimension.

Example: In the following example, the value of each dimension is by default the upper bound of that dimension.

```
array x{5,3} score1-score15;
```

As an alternative, the following ARRAY statement is a longhand version of the previous example:

```
array x{1:5,1:3} score1-score15;
```

Tip: For most arrays, 1 is a convenient lower bound. Thus, you do not need to specify the lower and upper bounds. However, specifying both bounds is useful when the array dimensions have a convenient beginning point other than 1.

Tip: To reduce the computational time that is needed for subscript evaluation, specify a lower bound of 0.

{*}
specifies that SAS is to determine the subscript by counting the variables in the array. When you specify the asterisk, also include *array-elements*.

Restriction: You cannot use the asterisk with _TEMPORARY_ arrays or when you define a multidimensional array.

$
specifies that the elements in the array are character elements.

Tip: The dollar sign is not necessary if the elements have been previously defined as character elements.

length
specifies the length of elements in the array that have not been previously assigned a length.

array-elements
specifies the names of the elements that make up the array. *Array-elements* must be either all numeric or all character, and they can be listed in any order. The elements can be

variables
lists variable names.

Range: The names must be either variables that you define in the ARRAY statement or variables that SAS creates by concatenating the array name and a number. For example, when the subscript is a number (not the asterisk), you do not need to name each variable in the array. Instead, SAS creates variable names by concatenating the array name and the numbers 1, 2, 3, …*n*.

Tip: These SAS variable lists enable you to reference variables that have been previously defined in the same DATA step:

NUMERIC
specifies all numeric variables.

CHARACTER
specifies all character variables.

ALL
specifies all variables.

Restriction: If you use _ALL_, all the previously defined variables must be of the same type.

Featured in: Example 1 on page 1444

TEMPORARY
creates a list of temporary data elements.

Range: Temporary data elements can be numeric or character.

Tip: Temporary data elements behave like DATA step variables with these exceptions:

- They do not have names. Refer to temporary data elements by the array name and dimension.
- They do not appear in the output data set.
- You cannot use the special subscript asterisk (*) to refer to all the elements.
- Temporary data element values are always automatically retained, rather than being reset to missing at the beginning of the next iteration of the DATA step.

Tip: Arrays of temporary elements are useful when the only purpose for creating an array is to perform a calculation. To preserve the result of the calculation, assign it to a variable. You can improve performance time by using temporary data elements.

(*initial-value-list*)
gives initial values for the corresponding elements in the array. The values for elements can be numbers or character strings. You must enclose all character strings in quotation marks. To specify one or more initial values directly, use the following format:

(*initial-value(s)*)

To specify an iteration factor and nested sublists for the initial values, use the following format:

<*constant-iter-value**> <(>*constant value* | *constant-sublist*<)>

Restriction: If you specify both an *initial-value-list* and *array-elements*, then *array-elements* must be listed before *initial-value-list* in the ARRAY statement.

Tip: You can assign initial values to both variables and temporary data elements.

Tip: Elements and values are matched by position. If there are more array elements than initial values, the remaining array elements receive missing values and SAS issues a warning.

Featured in: Example 2 on page 1444, and Example 3 on page 1444

Tip: You can separate the values in the initial value list with either a comma or a blank space.

Tip: You can also use a shorthand notation for specifying a range of sequential integers. The increment is always +1.

Tip: If you have not previously specified the attributes of the array elements (such as length or type), the attributes of any initial values that you specify are automatically assigned to the corresponding array element.

Note: Initial values are retained until a new value is assigned to the array element. △

Tip: When any (or all) elements are assigned initial values, all elements behave as if they were named on a RETAIN statement.

Examples: The following examples show how to use the iteration factor and nested sublists. All of these ARRAY statements contain the same initial value list:

- `ARRAY x{10} x1-x10 (10*5);`
- `ARRAY x{10} x1-x10 (5*(5 5));`
- `ARRAY x{10} x1-x10 (5 5 3*(5 5) 5 5);`
- `ARRAY x{10} x1-x10 (2*(5 5) 5 5 2*(5 5));`
- `ARRAY x{10} x1-x10 (2*(5 2*(5 5)));`

Details

The ARRAY statement defines a set of elements that you plan to process as a group. You refer to elements of the array by the array name and subscript. Because you usually want to process more than one element in an array, arrays are often referenced within DO groups.

Comparisons

- Arrays in the SAS language are different from arrays in many other languages. A SAS array is simply a convenient way of temporarily identifying a group of variables. It is not a data structure, and *array-name* is not a variable.
- An ARRAY statement defines an array. An array reference uses an array element in a program statement.

Examples

Example 1: Defining Arrays

```
array rain {5} janr febr marr aprr mayr;
array days{7} d1-d7;
array month{*} jan feb jul oct nov;
array x{*} _NUMERIC_;
array qbx{10};
array meal{3};
```

Example 2: Assigning Initial Numeric Values

```
array test{4} t1 t2 t3 t4 (90 80 70 70);
array test{4} t1-t4 (90 80 2*70);
array test{4} _TEMPORARY_ (90 80 70 70);
```

Example 3: Defining Initial Character Values

```
array test2{*} $ a1 a2 a3 ('a','b','c');
```

Example 4: Defining More Advanced Arrays

```
array new{2:5} green jacobs denato fetzer;
array x{5,3} score1-score15;
array test{3:4,3:7} test1-test10;
array temp{0:999} _TEMPORARY_;
array x{10} (2*1:5);
```

Example 5: Creating a Range of Variable Names That Have Leading Zeros The following example shows that you can create a range of variable names that have leading zeros. Each variable name has a length of three characters, and the names sort correctly (A01, A02, … A10). Without leading zeros, the variable names would sort in the following order: A1, A10, A2, … A9.

```
options pageno=1 nodate ps=64 ls=80;

data test (drop=i);
   array a(10) A01-A10;
   do i=1 to 10;
      a(i)=i;
   end;
run;

proc print noobs data=test;
run;
```

Output 6.1 Array Names That Have Leading Zeros

```
                              The SAS System                              1

      A01    A02    A03    A04    A05    A06    A07    A08    A09    A10

       1      2      3      4      5      6      7      8      9      10
```

See Also

Statement:

"Array Reference Statement" on page 1445

"Array Processing" in *SAS Language Reference: Concepts*

Array Reference Statement

Describes the elements in an array to be processed.

Valid: in a DATA step
Category: Information
Type: Declarative

Syntax

array-name { *subscript* }

Arguments

array-name
is the name of an array that was previously defined with an ARRAY statement in the same DATA step.

{*subscript*}
specifies the subscript. Any of these forms can be used:

{*variable-1*< , *…variable-n*>}
specifies a variable, or variable list that is usually used with DO-loop processing. For each execution of the DO loop, the current value of this variable becomes the subscript of the array element being processed.

Featured in: Example 1 on page 1446

Tip: You can enclose a subscript in braces ({ }), brackets ([ ]), or parentheses (( )).

{*}
forces SAS to treat the elements in the array as a variable list.

Tip: The asterisk can be used with the INPUT and PUT statements, and with some SAS functions.

Tip: This syntax is provided for convenience and is an exception to usual array processing.

Restriction: When you define an array that contains temporary array elements, you cannot reference the array elements with an asterisk.

Featured in: Example 4 on page 1447

expression-1< , . . . *expression-n*>
specifies a SAS expression.

Range: The expression must evaluate to a subscript value when the statement that contains the array reference executes. The expression can also be an integer with a value between the lower and upper bounds of the array, inclusive.

Featured in: Example 3 on page 1447

Details

- To refer to an array in a program statement, use an array reference. The ARRAY statement that defines the array must appear in the DATA step before any references to that array. An array definition is only in effect for the duration of the DATA step. If you want to use the same array in several DATA steps, redefine the array in each step.

  CAUTION:
  Using the name of a SAS function as an array name can cause unpredictable results. If you inadvertently use a function name as the name of the array, SAS treats parenthetical references that involve the name as array references, not function references, for the duration of the DATA step. A warning message is written to the SAS log. △

- You can use an array reference anywhere that you can write a SAS expression, including SAS functions and these SAS statements:
  - assignment statement
  - sum statement
  - DO UNTIL(*expression*)
  - DO WHILE(*expression*)
  - IF
  - INPUT
  - PUT
  - SELECT
  - WINDOW.
- The DIM function is often used with the iterative DO statement to return the number of elements in a dimension of an array, when the lower bound of the dimension is 1. If you use DIM, you can change the number of array elements without changing the upper bound of the DO statement. For example, because DIM(NEW) returns a value of 4, the following statements process all the elements in the array:

```
array new{*} score1-score4;
   do i=1 to dim(new);
      new{i}=new{i}+10;
   end;
```

Comparisons

- An ARRAY statement defines an array, whereas an array reference defines the members of the array to process.

Examples

Example 1: Using Iterative DO-Loop Processing In this example, the statements process each element of the array, using the value of variable I as the subscript on the array references for each iteration of the DO loop. If an array element has a value of 99, the IF-THEN statement changes that value to 100.

```
array days{7} d1-d7;
   do i=1 to 7;
```

```
      if days{i}=99 then days{i}=100;
   end;
```

Example 2: Referencing Many Arrays in One Statement You can refer to more than one array in a single SAS statement. In this example, you create two arrays, DAYS and HOURS. The statements inside the DO loop substitute the current value of variable I to reference each array element in both arrays.

```
array days{7} d1-d7;
   array hours{7} h1-h7;
      do i=1 to 7;
         if days{i}=99 then days{i}=100;
         hours{i}=days{i}*24;
      end;
```

Example 3: Specifying the Subscript In this example, the INPUT statement reads in variables A1, A2, and the third element (A3) of the array named ARR1:

```
array arr1{*} a1-a3;
x=1;
input a1 a2 arr1{x+2};
```

Example 4: Using the Asterisk References as a Variable List

□
```
array cost{10} cost1-cost10;
totcost=sum(of cost {*});
```

□
```
array days{7} d1-d7;
input days {*};
```

□
```
array hours{7} h1-h7;
put hours {*};
```

See Also

Function:

"DIM Function" on page 655

Statements

"ARRAY Statement" on page 1440

"DO Statement, Iterative" on page 1492

"Array Processing" in *SAS Language Reference: Concepts*

Assignment Statement

Evaluates an expression and stores the result in a variable.

Valid: in a DATA step

Category: Action

Type: Executable

Syntax

variable=expression;

Arguments

variable
names a new or existing variable.

Range: *Variable* can be a variable name, array reference, or SUBSTR function.

Tip: Variables that are created by the Assignment statement are not automatically retained.

expression
is any SAS expression.

Tip: *expression* can contain the variable that is used on the left side of the equal sign. When a variable appears on both sides of a statement, the original value on the right side is used to evaluate the expression, and the result is stored in the variable on the left side of the equal sign. For more information, see "Expressions" in *SAS Language Reference: Concepts*.

Details

Assignment statements evaluate the expression on the right side of the equal sign and store the result in the variable that is specified on the left side of the equal sign.

Examples

These assignment statements use different types of expressions:

- `name='Amanda Jones';`
- `WholeName='Ms. '||name;`
- `a=a+b;`

See Also

Statement:

"Sum Statement" on page 1777

ATTRIB Statement

Associates a format, informat, label, and length with one or more variables.

Valid: in a DATA step

Category: Information

Type: Declarative

See: ATTRIB Statement in the documentation for your operating environment.

Syntax

ATTRIB *variable-list(s) attribute-list(s)* ;

Arguments

variable-list(s)
names the variables that you want to associate with the attributes.

Tip: List the variables in any form that SAS allows.

attribute-list(s)
specifies one or more attributes to assign to *variable-list*. Specify one or more of these attributes in the ATTRIB statement:

FORMAT=*format*
associates a format with variables in *variable-list.*

Tip: The format can be either a standard SAS format or a format that is defined with the FORMAT procedure.

INFORMAT=*informat*
associates an informat with variables in *variable-list.*

Tip: The informat can be either a standard SAS informat or an informat that is defined with the FORMAT procedure.

LABEL='*label*'
associates a label with variables in *variable-list.*

LENGTH=<$>*length*
specifies the length of variables in *variable-list.*

Requirement: Put a dollar sign ($) in front of the length of character variables.

Tip: Use the ATTRIB statement before the SET statement to change the length of variables in an output data set when you use an existing data set as input.

Range: For character variables, the range is 1 to 32,767 for all operating environments.

Operating Environment Information: For numeric variables, the minimum length you can specify with the LENGTH= specification is 2 in some operating environments and 3 in others. △

Restriction: You cannot change the length of a variable using LENGTH= from PROC DATASETS.

TRANSCODE=YES | NO
specifies whether character variables can be transcoded. Use TRANSCODE=NO to suppress transcoding. For more information about transcoding, see "Transcoding" in the *SAS National Language Support (NLS): Reference Guide*.

Default: YES

Restriction: The TRANSCODE=NO attribute is not supported by some SAS Workspace Server clients. In SAS 9.2, if the attribute is not supported, variable values with TRANSCODE=NO are replaced (masked) with asterisks (*). Before SAS 9.2, variables with TRANSCODE=NO were transcoded.

Restriction: Prior releases of SAS cannot access a SAS 9.1 data set that contains a variable with a TRANSCODE=NO attribute.

Restriction: Transcode suppression is not supported by the V6TAPE engine.

Interaction: You can use the VTRANSCODE and VTRANSCODEX functions to return a value that indicates whether transcoding is on or off for a character variable.

Interaction: If the TRANSCODE= attribute is set to NO for any character variable in a data set, then PROC CONTENTS prints a transcode column that contains the TRANSCODE= value for each variable in the data set. If all variables in the data set are set to the default TRANSCODE= value (YES), then no transcode column prints.

Details

The Basics Using the ATTRIB statement in the DATA step permanently associates attributes with variables by changing the descriptor information of the SAS data set that contains the variables.

You can use ATTRIB in a PROC step, but the rules are different.

How SAS Treats Variables When You Assign Informats with the INFORMAT= Option in the ATTRIB Statement Informats that are associated with variables by using the INFORMAT= option in the ATTRIB statement behave like informats that are used with modified list input. SAS reads the variables by using the scanning feature of list input, but applies the informat. In modified list input, SAS

- does not use the value of *w* in an informat to specify column positions or input field widths in an external file
- uses the value of *w* in an informat to specify the length of previously undefined character variables
- ignores the value of *w* in numeric informats
- uses the value of *d* in an informat in the same way it usually does for numeric informats
- treats blanks that are embedded as input data as delimiters unless you change their status with the DLM= or DLMSTR= option specification in an INFILE statement.

If you have coded the INPUT statement to use another style of input, such as formatted input or column input, that style of input is not used when you use the INFORMAT= option in the ATTRIB statement.

How SAS Treats Transcoded Variables When You Use the SET and MERGE Statements When you use the SET or MERGE statement to create a data set from several data sets, SAS makes the TRANSCODE= attribute of the variable in the output data set equal to the TRANSCODE= value of the variable in the first data set. See Example 2 on page 1451 and Example 3 on page 1451.

Note: The TRANSCODE= attribute is set when the variable is first seen on an input data set or in an ATTRIB TRANSCODE= statement. If a SET or MERGE statement comes before an ATTRIB TRANSCODE= statement and the TRANSCODE= attribute contradicts the SET statement, a warning will occur. △

Comparisons

You can use either an ATTRIB statement or an individual attribute statement such as FORMAT, INFORMAT, LABEL, and LENGTH to change an attribute that is associated with a variable.

Examples

Example 1: Examples of ATTRIB Statements with Varying Numbers of Variables and Attributes Here are examples of ATTRIB statements that contain different numbers of variables and attributes:

- single variable and single attribute:

```
attrib cost length=4;
```

- single variable with multiple attributes:

```
attrib saleday informat=mmddyy.
format=worddate.;
```

- multiple variables with the same multiple attributes:

```
attrib x y length=$4 label='TEST VARIABLE';
```

- multiple variables with different multiple attributes:

```
attrib x length=$4 label='TEST VARIABLE'
       y length=$2 label='RESPONSE';
```

- variable list with single attribute:

```
attrib month1-month12
       label='MONTHLY SALES';
```

Example 2: Using the SET Statement with Transcoded Variables

In this example, which uses the SET statement, the variable Z's TRANSCODE= attribute in data set A is NO because B is the first data set and Z's TRANSCODE= attribute in data set B is NO.

```
data b;
   length z $4;
   z = 'ice';
   attrib z transcode = no;
data c;
   length z $4;
   z = 'snow';
   attrib z transcode = yes;
data a;
   set b;
   set c;
   /* Check transcode setting for variable Z */
   rc1 = vtranscode(z);
   put rc1=;
run;
```

Example 3: Using the MERGE Statement with Transcoded Variables In this example, which uses the MERGE statement, the variable Z's TRANSCODE= attribute in data set A is YES because C is the first data set and Z's TRANSCODE= attribute in data set C is YES.

```
data b;
   length z $4;
   z = 'ice';
   attrib z transcode = no;
```

```
data c;
   length z $4;
   z = 'snow';
   attrib z transcode = yes;
data a;
   merge c b;
   /* Check transcode setting for variable Z */
   rc1 = vtranscode(z);
   put rc1=;
run;
```

See Also

Statements:

"FORMAT Statement" on page 1576
"INFORMAT Statement" on page 1614
"LABEL Statement" on page 1650
"LENGTH Statement" on page 1654

Functions:

VTRANSCODE in the *SAS National Language Support (NLS): Reference Guide*
VTRANSCODEX in the *SAS National Language Support (NLS): Reference Guide*

BY Statement

Controls the operation of a SET, MERGE, MODIFY, or UPDATE statement in the DATA step and sets up special grouping variables.

Valid: in a DATA step or a PROC step

Category: File-handling

Type: Declarative

Syntax

BY <DESCENDING> *variable-1*
<…<DESCENDING> *variable-n* > <NOTSORTED><GROUPFORMAT>;

Arguments

DESCENDING
specifies that the data sets are sorted in descending order by the variable that is specified. DESCENDING means largest to smallest numerically, or reverse alphabetical for character variables.

Restriction: You cannot use the DESCENDING option with data sets that are indexed because indexes are always stored in ascending order.

Featured in: Example 2 on page 1455

GROUPFORMAT
uses the formatted values, instead of the internal values, of the BY variables to determine where BY groups begin and end, and therefore how FIRST.*variable* and LAST.*variable* are assigned. Although the GROUPFORMAT option can appear anywhere in the BY statement, the option applies to *all* variables in the BY statement.

Restriction: You must sort the observations in a data set based on the value of the BY variables before using the GROUPFORMAT option in the BY statement.

Restriction: You can use the GROUPFORMAT option in a BY statement only in a DATA step.

Interaction: If you also use the NOTSORTED option, you can group the observations in a data set by the formatted value of the BY variables without requiring that the data set be sorted or indexed.

Tip: Using the GROUPFORMAT option is useful when you define your own formats to display data that is grouped.

Tip: Using the GROUPFORMAT option in the DATA step ensures that BY groups that you use to create a data set match the BY groups in PROC steps that report grouped, formatted data.

Comparison: BY-group processing in the DATA step using the GROUPFORMAT option is the same as BY-group processing with formatted values in SAS procedures.

See Also: By-Group Processing in the DATA Step in *SAS Language Reference: Concepts*

Featured in: Example 4 on page 1455

variable
names each variable by which the data set is sorted or indexed. These variables are referred to as BY variables for the current DATA or PROC step.

Tip: The data set can be sorted or indexed by more than one variable.

Featured in: Example 1 on page 1455, Example 2 on page 1455, Example 3 on page 1455, and Example 4 on page 1455

NOTSORTED
specifies that observations with the same BY value are grouped together but are not necessarily sorted in alphabetical or numeric order.

Restriction: You cannot use the NOTSORTED option with the MERGE and UPDATE statements.

Tip: The NOTSORTED option can appear anywhere in the BY statement.

Tip: Using the NOTSORTED option is useful if you have data that falls into other logical groupings such as chronological order or categories.

Featured in: Example 3 on page 1455

Details

How SAS Identifies the Beginning and End of a BY Group

SAS identifies the beginning and end of a BY group by creating two temporary variables for each BY variable: FIRST.*variable* and LAST.*variable*. The value of these variables is either 0 or 1. SAS sets the value of FIRST.*variable* to 1 when it reads the first observation in a BY group, and sets the value of LAST.*variable* to 1 when it reads the last observation in a BY group. These temporary variables are available for DATA step programming but are not added to the output data set.

For a complete explanation of how SAS processes grouped data and of how to prepare your data, see "By-Group Processing in the DATA Step" in *SAS Language Reference: Concepts*.

In a DATA Step The BY statement applies only to the SET, MERGE, MODIFY, or UPDATE statement that precedes it in the DATA step, and only one BY statement can accompany each of these statements in a DATA step.

The data sets that are listed in the SET, MERGE, or UPDATE statements must be sorted by the values of the variables that are listed in the BY statement or have an appropriate index. As a default, SAS expects the data sets to be arranged in ascending numeric order or in alphabetical order. The observations can be arranged by one of the following methods:

- sorting the data set
- creating an index for the variables
- inputting the observations in order.

Note: MODIFY does not require sorted data, but sorting can improve performance. △

Note: The BY statement honors the linguistic collation of data that is sorted by using the SORT procedure with the SORTSEQ=LINGUISTIC option. △

For more information, see "How to Prepare Your Data Sets" in *SAS Language Reference: Concepts*.

In a PROC Step You can specify the BY statement with some SAS procedures to modify their action. Refer to the individual procedure in the *Base SAS Procedures Guide* for a discussion of how the BY statement affects processing for SAS procedures.

With SAS Views If you create a DATA step view by reading from a DBMS and the SET, MERGE, UPDATE, or MODIFY statement is followed by a BY statement, the BY statement might cause the DBMS to sort the data in order to return the data in sorted order. Sorting the data could increase execution time.

Processing BY Groups SAS assigns the following values to FIRST.*variable* and LAST.*variable*:

- FIRST.*variable* has a value of 1 under the following conditions:
  - when the current observation is the first observation that is read from the data set.
  - when you do not use the GROUPFORMAT option and the internal value of the variable in the current observation differs from the internal value in the previous observation.

    If you use the GROUPFORMAT option, FIRST.*variable* has a value of 1 when the formatted value of the variable in the current observation differs from the formatted value in the previous observation.
  - FIRST.*variable* has a value of 1 for any preceding variable in the BY statement.

    In all other cases, FIRST.*variable* has a value of 0.
- LAST.*variable* has a value of 1 under the following conditions:
  - when the current observation is the last observation that is read from the data set.
  - when you use the GROUPFORMAT option and the internal value of the variable in the current observation differs from the internal value in the next observation.

If you use the GROUPFORMAT option, LAST.*variable* has a value of 1 when the formatted value of the variable in the current observation differs from the formatted value in the next observation.

- LAST.*variable* has a value of 1 for any preceding variable in the BY statement.

In all other cases, LAST.*variable* has a value of 0.

Examples

Example 1: Specifying One or More BY Variables

- Observations are in ascending order of the variable DEPT:

```
by dept;
```

- Observations are in alphabetical (ascending) order by CITY and, within each value of CITY, in ascending order by ZIPCODE:

```
by city zipcode;
```

Example 2: Specifying Sort Order

- Observations are in ascending order of SALESREP and, within each SALESREP value, in descending order of the values of JANSALES:

```
by salesrep descending jansales;
```

- Observations are in descending order of BEDROOMS, and, within each value of BEDROOMS, in descending order of PRICE:

```
by descending bedrooms descending price;
```

Example 3: BY-Group Processing with Nonsorted Data Observations are ordered by the name of the month in which the expenses were accrued:

```
by month notsorted;
```

Example 4: Grouping Observations By Using Formatted Values The following example illustrates the use of the GROUPFORMAT option.

```
proc format;
   value range
      low -55 = 'Under 55'
      55-60   = '55 to 60'
      60-65   = '60 to 65'
      65-70   = '65 to 70'
      other   = 'Over 70';
run;

proc sort data=class out=sorted_class;
   by height;
run;

data _null_;
   format height range.;
   set sorted_class;
      by height groupformat;
   if first.height then
```

```
 put 'Shortest in ' height 'measures ' height:best12.;
run;
```

SAS writes the following output to the log:

```
Shortest in Under 55 measures 51.3
Shortest in 55 to 60 measures 56.3
Shortest in 60 to 65 measures 62.5
Shortest in 65 to 70 measures 65.3
Shortest in Over 70 measures 72
```

Example 5: Combining Multiple Observations and Grouping Them Based on One BY Value The following example shows how to use FIRST.*variable* and LAST.*variable* with BY-group processing.

```
options pageno=1 nodate ls=80 ps=64;

data Inventory;
   length RecordID 8 Invoice $ 30 ItemLine $ 50;
   infile datalines;
   input RecordID Invoice ItemLine &;
   drop RecordID;
   datalines;
A74  A5296  Highlighters
A75  A5296  Lot # 7603
A76  A5296  Yellow Blue Green
A77  A5296  24 per box
A78  A5297  Paper Clips
A79  A5297  Lot # 7423
A80  A5297  Small Medium Large
A81  A5298  Gluestick
A82  A5298  Lot # 4422
A83  A5298  New item
A84  A5299  Rubber bands
A85  A5299  Lot # 7892
A86  A5299  Wide width, Narrow width
A87  A5299  1000 per box
;

data combined;
   array Line[4] $ 60 ;
   retain Line1-Line4;
   keep Invoice Line1-Line4;

   set Inventory;
   by Invoice;

   if first.Invoice then do;
      call missing(of Line1-Line4);
      records = 0;
   end;

   records + 1;
   Line[records]=ItemLine;
```

```
   if last.Invoice then output;
run;

proc print data=combined;
   title 'Office Supply Inventory';
run;
```

Output 6.2 Output from Combining Multiple Observations

```
                          Office Supply Inventory                            1

Obs     Line1        Line2     Line3                          Line4      Invoice

 1   Highlighters  Lot # 7603  Yellow Blue Green          24 per box      A5296
 2   Paper Clips   Lot # 7423  Small Medium Large                         A5297
 3   Gluestick     Lot # 4422  New item                                  A5298
 4   Rubber bands  Lot # 7892  Wide width, Narrow width  1000 per box    A5299
```

See Also

Statements:

"MERGE Statement" on page 1679
"MODIFY Statement" on page 1684
"SET Statement" on page 1764
"UPDATE Statement" on page 1787

CALL Statement

Invokes a SAS CALL routine.

Valid: in a DATA step
Category: Action
Type: Executable

Syntax

CALL *routine*(*parameter-1*<, *…parameter-n*>);

Arguments

routine
specifies the name of the SAS CALL routine that you want to invoke. For information about available routines, see Chapter 4, "Functions and CALL Routines," on page 295.

(*parameter*)
is a piece of information to be passed to or returned from the routine.

Requirement: Enclose this information, which depends on the specific routine, in parentheses.

Tip: You can specify additional parameters, separated by commas.

Details

SAS CALL routines can assign variable values and perform other system functions.

See Also

Chapter 4, "Functions and CALL Routines," on page 295

CARDS Statement

Specifies that data lines follow.

Valid: in a DATA step
Category: File-handling
Type: Declarative
Alias: DATALINES, LINES
See: "DATALINES Statement" on page 1474
See Also: CARDS Statement in the *SAS Companion for UNIX Environments*

CARDS4 Statement

Specifies that data lines that contain semicolons follow.

Valid: in a DATA step
Category: File-handling
Type: Declarative
Alias: DATALINES4, LINES4
See: "DATALINES4 Statement" on page 1475

CATNAME Statement

Logically combines two or more catalogs into one by associating them with a catref (a shortcut name); clears one or all catrefs; lists the concatenated catalogs in one concatenation or in all concatenations.

Valid: Anywhere

Category: Data Access

Syntax

CATNAME <*libref.*> *catref*
< *(libref-1.catalog-1* <(ACCESS=READONLY)>
<...*libref-n.catalog–n* <(ACCESS=READONLY)>)> ;

CATNAME <*libref.*> *catref* CLEAR | _ALL_ CLEAR;

CATNAME <*libref.*> *catref* LIST | _ALL_ LIST;

Arguments

libref
is any previously assigned SAS libref. If you do not specify a libref, SAS concatenates the catalog in the Work library, using the catref that you specify.

Restriction: The libref must have been previously assigned.

catref
is a unique catalog reference name for a catalog or a catalog concatenation that is specified in the statement. Separate the catref from the libref with a period, as in *libref.catref*. Any SAS name can be used for this catref.

catalog
is the name of a catalog that is available for use in the catalog concatenation.

Options

CLEAR
disassociates a currently assigned *catref* or *libref.catref*.

Tip: Specify a specific *catref* or *libref.catref* to disassociate it from a single concatenation. Specify _ALL_ CLEAR to disassociate all currently assigned *catref* or *libref.catref* concatenations.

ALL CLEAR
disassociates all currently assigned *catref* or *libref.catref* concatenations.

LIST
writes the catalog names that are included in the specified concatenation to the SAS log.

Tip: Specify *catref* or *libref.catref* to list the attributes of a single concatenation. Specify _ALL_ to list the attributes of all catalog concatenations in your current session.

ALL LIST
writes all catalog names that are included in any current catalog concatenation to the SAS log.

ACCESS=READONLY
assigns a read-only attribute to the catalog. SAS, therefore, will allow users to read from the catalog entries but not to update information or to write new information.

Details

Why Use CATNAME? CATNAME is useful because it enables you to access entries in multiple catalogs by specifying a single catalog reference name (*libref.catref* or *catref*). After you create a catalog concatenation, you can specify the catref in any context that accepts a simple (non-concatenated) catref.

Rules for Catalog Concatenation To use catalog concatenation effectively, you must understand the rules that determine how catalog entries are located among the concatenated catalogs:

1. When a catalog entry is opened for input or update, the concatenated catalogs are searched and the first occurrence of the specified entry is used.
2. When a catalog entry is opened for output, it will be created in the first catalog that is listed in the concatenation.

   Note: A new catalog entry is created in the first catalog even if there is an entry with the same name in another part of the concatenation. △

   Note: If the first catalog in a concatenation that is opened for update does not exist, the item will be written to the next catalog that exists in the concatenation. △
3. When you want to delete or rename a catalog entry, only the first occurrence of the entry is affected.
4. Any time a list of catalog entries is displayed, only one occurrence of a catalog entry name is shown.

   Note: Even if the name occurs multiple times in the concatenation, only the first occurrence is shown. △

Comparisons

- The CATNAME statement is like a LIBNAME statement for catalogs. The LIBNAME statement allows you to assign a shortcut name to a SAS library so that you can use the shortcut name to find the files and use the data they contain. CATNAME allows you to assign a short name <*libref.*>*catref* (libref is optional) to one or more catalogs so that SAS can find the catalogs and use all or some of the entries in each catalog.
- The CATNAME statement *explicitly* concatenates SAS catalogs. You can use the LIBNAME statement to *implicitly* concatenate SAS catalogs.

Examples

Example 1: Assigning and Using a Catalog Concatenation You might need to access entries in several SAS catalogs. The most efficient way to access the information is to logically concatenate the catalogs. Catalog concatenation enables access to the information without actually creating a new, separate, and possibly very large catalog.

Assign librefs to the SAS libraries that contain the catalogs that you want to concatenate:

```
libname mylib1 'data-library-1';
libname mylib2 'data-library-2';
```

Assign a catref, which can be any valid SAS name, to the list of catalogs that you want to logically concatenate:

```
catname allcats (mylib1.catalog1 mylib2.catalog2);
```

The SAS log displays this message:

Output 6.3 Log Output from CATNAME Statement

```
NOTE: Catalog concatenation WORK.ALLCATS has been created.
```

Because no libref is specified, the libref is WORK by default. When you want to access a catalog entry in either of these catalogs, use the libref WORK and the catalog reference name ALLCATS instead of the original librefs and catalog names. For example, to access a catalog entry named APPKEYS.KEYS in the catalog MYLIB1.CATALOG1, specify

```
work.allcats.appkeys.keys
```

Example 2: Creating a Nested Catalog Concatenation After you create a concatenated catalog, you can use CATNAME to combine your concatenation with other single catalogs or other concatenated catalogs. Nested catalog concatenation is useful, because you can use a single catref to access many different catalog combinations.

```
libname local 'my_dir';
libname main  'public_dir';

catname private_catalog (local.my_application_code
                         local.my_frames
                         local.my_formats);

catname combined_catalogs (private_catalog
                         main.public_catalog);
```

In the above example, you could work on private copies of your application entries by using PRIVATE_CATALOG. If you want to see how your entries function when they are combined with the public version of the application, you can use COMBINED_CATALOGS.

See Also

Statements:

"FILENAME Statement" on page 1520

"FILENAME Statement, CATALOG Access Method" on page 1526

"LIBNAME Statement" on page 1656 for a discussion of *implicitly* concatenating SAS catalogs.

CHECKPOINT EXECUTE_ALWAYS Statement

Indicates to execute the DATA step or PROC step that immediately follows without considering the checkpoint-restart data.

Valid: Anywhere

Category: Program Control

Syntax

CHECKPOINT EXECUTE_ALWAYS;

Without Arguments

The CHECKPOINT EXECUTE_ALWAYS statement indicates to SAS that the DATA step or PROC step that immediately follows is to be executed without considering the checkpoint data.

Details

If checkpoint-restart mode is enabled and a batch program terminates without completing, the program can be rerun beginning with the DATA step or PROC step that was executing when it terminated. DATA or PROC steps that completed before the batch program terminated are not reexecuted. If a DATA step or a PROC step must be reexecuted, you can add the CHECKPOINT EXECUTE_ALWAYS statement before the step. Using the CHECKPOINT EXECUTE_ALWAYS statement ensures that SAS always executes the step without regard to the checkpoint-restart data.

See Also

System Options:

"STEPCHKPT System Option" on page 2013

"STEPCHKPTLIB= System Option" on page 2014

"STEPRESTART System Option" on page 2016

"Restarting Batch Programs" in *SAS Language Reference: Concepts*

Comment Statement

Specifies the purpose of the statement or program.

Valid: anywhere

Category: Log Control

Syntax

**message*;

or

/**message**/

Arguments

****message*;**
specifies the text that explains or documents the statement or program.

Range: These comments can be any length and are terminated with a semicolon.

Restriction: These comments must be written as separate statements.

Restriction: These comments cannot contain internal semicolons or unmatched quotation marks.

Restriction: A macro statement or macro variable reference that is contained inside this form of comment is processed by the SAS macro facility. This form of comment cannot be used to hide text from the SAS macro facility.

Tip: When using comments within a macro definition or to hide text from the SAS macro facility, use this style comment:

```
/* message */
```

/*message*/
specifies the text that explains or documents the statement or program.

Range: These comments can be any length.

Restriction: This type of comment cannot be nested.

Tip: These comments can contain semicolons and unmatched quotation marks.

Tip: You can write these comments within statements or anywhere a single blank can appear in your SAS code.

Tip: In the Microsoft Windows operating environment, if you use the Enhanced Editor, you can comment out a block of code by highlighting the block and then pressing CTRL-/ (forward slash). To uncomment a block of code, highlight the block and press CTRL-SHIFT-/ (forward slash).

Details

You can use the comment statement anywhere in a SAS program to document the purpose of the program, explain unusual segments of the program, or describe steps in a complex program or calculation. SAS ignores text in comment statements during processing.

CAUTION:
Avoid placing the /* comment symbols in columns 1 and 2. In some operating environments, SAS might interpret a /* in columns 1 and 2 as a request to end the SAS program or session. △

Note: You can add these lines to your code to fix unmatched comment tags, unmatched quotation marks, and missing semicolons.

```
/* '; * "; */;
quit;
run;
```

△

Examples

These examples illustrate the two types of comments:

- This example uses the *message; format:

```
*This code finds the number in the BY group;
```

- This example uses the *message; format:

```
*----------------------------------------*
|  This uses one comment statement      |
|           to draw a box.              |
*----------------------------------------*;
```

- This example uses the /*message*/ format:

```
input @1 name $20.  /* last name    */
      @200 test 8.  /* score test   */
      @50 age 3.;   /* customer age */
```

- This example uses the /*message*/ format:

```
/* For example 1 use: x=abc;
   for example 2 use: y=ghi; */
```

CONTINUE Statement

Stops processing the current DO-loop iteration and resumes processing the next iteration.

Valid: in a DATA step
Category: Control
Type: Executable
Restriction: Can be used only in a DO loop

Syntax

CONTINUE;

Without Arguments

The CONTINUE statement has no arguments. It stops processing statements within the current DO-loop iteration based on a condition. Processing resumes with the next iteration of the DO loop.

Comparisons

- The CONTINUE statement stops the processing of the current iteration of a loop and resumes with the next iteration; the LEAVE statement causes processing of the current loop to end.

- You can use the CONTINUE statement only in a DO loop; you can use the LEAVE statement in a DO loop or a SELECT group.

Examples

This DATA step creates a report of benefits for new full-time employees. If an employee's status is PT (part-time), the CONTINUE statement prevents the second INPUT statement and the OUTPUT statement from executing.

```
data new_emp;
   drop i;
   do i=1 to 5;
      input name $ idno status $;
         /* return to top of loop  */
         /* when condition is true */
      if status='PT' then continue;
         input benefits $10.;
         output;
   end;
   datalines;
Jones 9011 PT
Thomas 876 PT
Richards 1002 FT
Eye/Dental
Kelly 85111 PT
Smith 433 FT
HMO
;
```

See Also

Statements:

"DO Statement, Iterative" on page 1492
"LEAVE Statement" on page 1653

DATA Statement

Begins a DATA step and provides names for any output SAS data sets, views, or programs.

Valid: in a DATA step
Category: File-handling
Type: Declarative

Syntax

❶**DATA** <*data-set-name-1* <(*data-set-options-1*)>>
<… *data-set-name-n* <(*data-set-options-n*)>> </ <DEBUG> <NESTING> <STACK = *stack-size*>> <NOLIST>;

❷**DATA** _NULL_ </ <DEBUG> <NESTING> <STACK = *stack-size*>> <NOLIST>;

❸**DATA** *view-name* <*data-set-name-1* <(*data-set-options-1*)>>
<… *data-set-name-n* <(*data-set-options-n*)>> /
VIEW=*view-name* <(<*password-option*><SOURCE=*source-option*>)>
<NESTING> <NOLIST>;

❹**DATA** *data-set-name* / PGM=*program-name*
<(<*password-option*><SOURCE=*source-option*>)> <NESTING> <NOLIST>;

❺**DATA** VIEW=*view-name* <(*password-option*)> <NOLIST>;
DESCRIBE;

❻**DATA** PGM=*program-name* <(*password-option*)> <NOLIST>;
<**DESCRIBE**;>
<**REDIRECT** INPUT | OUTPUT *old-name-1* = *new-name-1*<… *old-name-n* = *new-name-n*>;>
<**EXECUTE**;>

Without Arguments

If you omit the arguments, the DATA step automatically names each successive data set that you create as DATA*n*, where *n* is the smallest integer that makes the name unique.

Arguments

data-set-name
names the SAS data file or DATA step view that the DATA step creates. To create a DATA step view, you must specify at least one *data-set-name* and that *data-set-name* must match *view-name*.

Restriction: *data-set-name* must conform to the rules for SAS names, and additional restrictions might be imposed by your operating environment.

Tip: You can execute a DATA step without creating a SAS data set. See Example 5 on page 1472 for an example. For more information, see "❷When Not Creating a Data Set" on page 1469.

See also: For details about the types of SAS data set names and when to use each type, see "Names in the SAS Language" in *SAS Language Reference: Concepts*.

(*data-set-options*)
specifies optional arguments that the DATA step applies when it writes observations to the output data set.

See also: "Definition of Data Set Options" on page 10 for more information and Chapter 2, "SAS Data Set Options," on page 9 for a list of data set options .

Featured in: Example 1 on page 1470

/ DEBUG
enables you to debug your program interactively by helping to identify logic errors, and sometimes data errors.

/ NESTING
specifies that a note will be printed to the SAS log for the beginning and end of each DO-END and SELECT-END nesting level. This option enables you to debug mismatched DO-END and SELECT-END statements and is particularly useful in large programs where the nesting level is not obvious.

/ STACK=*stack-size*
specifies the maximum number of nested LINK statements.

NULL
specifies that SAS does not create a data set when it executes the DATA step.

VIEW=*view-name*
names a view that the DATA step uses to store the input DATA step view.

Restriction: *view-name* must match one of the data set names.

Restriction: SAS creates only one view in a DATA step.

Tip: If you specify additional data sets in the DATA statement, SAS creates these data sets when the view is processed in a subsequent DATA or PROC step. Views have the capability of generating other data sets at the time the view is executed.

Tip: SAS macro variables resolve when the view is created. Use the SYMGET function to delay macro variable resolution until the view is processed.

Featured in: Example 2 on page 1471 and Example 3 on page 1471

password-option
assigns a password to a stored compiled DATA step program or a DATA step view. The following password options are available:

ALTER=*alter-password*
assigns an *alter* password to a SAS data file. The password allows you to protect or replace a stored compiled DATA step program or a DATA step view.

Requirement: If you use an ALTER password in creating a stored compiled DATA step program or a DATA step view, an ALTER password is required to replace the program or view.

Requirement: If you use an ALTER password in creating a stored compiled DATA step program or a DATA step view, an ALTER password is required to execute a DESCRIBE statement.

Alias: PROTECT=

READ=*read-password*
assigns a *read* password to a SAS data file. The password allows you to read or execute a stored compiled DATA step program or a DATA step view.

Requirement: If you use a READ password in creating a stored compiled DATA step program or a DATA step view, a READ password is required to execute the program or view.

Requirement: If you use a READ password in creating a stored compiled DATA step program or a DATA step view, a READ password is required to execute DESCRIBE and EXECUTE statements. If you use an invalid password, SAS will execute the DESCRIBE statement.

Tip: If you use a READ password in creating a stored compiled DATA step program or a DATA step view, no password is required to replace the program or view.

Alias: EXECUTE=

PW=*password*
assigns a READ and ALTER password, both having the same value.

SOURCE=*source-option*
specifies one of the following source options:

SAVE
saves the source code that created a stored compiled DATA step program or a DATA step view.

ENCRYPT
encrypts and saves the source code that created a stored compiled DATA step program or a DATA step view.

Tip: If you encrypt source code, use the ALTER password option as well. SAS issues a warning message if you do not use ALTER.

NOSAVE
does not save the source code.

CAUTION:
If you use the NOSAVE option for a DATA step view, the view cannot be migrated or copied from one version of SAS to another version. △

Default: SAVE

PGM=*program-name*
names the stored compiled program that SAS creates or executes in the DATA step. To *create* a stored compiled program, specify a slash (/) before the PGM= option. To *execute* a stored compiled program, specify the PGM= option without a slash (/).

Tip: SAS macro variables resolve when the stored program is created. Use the SYMGET function to delay macro variable resolution until the view is processed.

Featured in: Example 4 on page 1471

NOLIST
suppresses the output of all variables to the SAS log when the value of _ERROR_ is 1.

Restriction: NOLIST must be the last option in the DATA statement.

Details

Using the DATA Statement The DATA step begins with the DATA statement. You use the DATA statement to create the following types of output: SAS data sets, data views, and stored programs. You can specify more than one output in a DATA statement. However, only one of the outputs can be a data view. You create a view by specifying the ❸VIEW= option and a stored program by specifying the ❹PGM=option.

Using Both a READ and an ALTER Password If you use both a READ and an ALTER password in creating a stored compiled DATA step program or a DATA step view, the following items apply:

- A READ or ALTER password is required to execute the stored compiled DATA step program or DATA step view.
- A READ or ALTER password is required if the stored compiled DATA step program or DATA step view contains both DESCRIBE and EXECUTE statements.
  - If you use an ALTER password with the DESCRIBE and EXECUTE statements, the following items apply:
    - SAS executes both the DESCRIBE and the EXECUTE statements.
    - If you execute a stored compiled DATA step program or DATA step view with an invalid ALTER password:
      - The DESCRIBE statement does not execute.
      - In batch mode, the EXECUTE statement has no effect.
      - In interactive mode, SAS prompts you for a READ password. If the READ password is valid, SAS processes the EXECUTE statement. If it is invalid, SAS does not process the EXECUTE statement.

- If you use a READ password with the DESCRIBE and EXECUTE statements, the following items apply:
  - In interactive mode, SAS prompts you for the ALTER password:
    - If you enter a valid ALTER password, SAS executes both the DESCRIBE and the EXECUTE statements.
    - If you enter an invalid ALTER password, SAS processes the EXECUTE statement but not the DESCRIBE statement.
  - In batch mode, SAS processes the EXECUTE statement but not the DESCRIBE statement.
  - In both interactive and batch modes, if you specify an invalid READ password SAS does not process the EXECUTE statement.

- An ALTER password is required if the stored compiled DATA step program or DATA step view contains a DESCRIBE statement.
- An ALTER password is required to replace the stored compiled DATA step program or DATA step view.

❶Creating an Output Data Set Use the DATA statement to create one or more output data sets. You can use data set options to customize the output data set. The following DATA step creates two output data sets, example1 and example2. It uses the data set option DROP to prevent the variable IDnumber from being written to the example2 data set.

```
data example1 example2 (drop=IDnumber);
   set sample;
   . . .more SAS statements. . .
run;
```

❷When Not Creating a Data Set Usually, the DATA statement specifies at least one data set name that SAS uses to create an output data set. However, when the purpose of a DATA step is to write a report or to write data to an external file, you might not want to create an output data set. Using the keyword _NULL_ as the data set name causes SAS to execute the DATA step without writing observations to a data set. This example writes to the SAS log the value of Name for each observation. SAS does not create an output data set.

```
data _NULL_;
   set sample;
   put Name ID;
run;
```

❸Creating a DATA Step View You can create DATA step views and execute them at a later time. The following DATA step example creates a DATA step view. It uses the SOURCE=ENCRYPT option to both save and encrypt the source code.

```
data phone_list / view=phone_list (source=encrypt);
   set customer_list;
   . . .more SAS statements. . .
run;
```

For more information about DATA step views, see “SAS Data Views” in *SAS Language Reference: Concepts*.

❹Creating a Stored Compiled DATA Step Program The ability to compile and store DATA step programs allows you to execute the stored programs later. Stored compiled DATA step programs can reduce processing costs by eliminating the need to compile

DATA step programs repeatedly. The following DATA step example compiles and stores a DATA step program. It uses the ALTER password option, which allows the user to replace an existing stored program, and to protect the stored compiled program from being replaced.

```
data testfile / pgm=stored.test_program (alter=sales);
   set sales_data;
   . . .more SAS statements. . .
run;
```

For more information about stored compiled DATA step programs, see "Stored Compiled DATA Step Programs" in *SAS Language Reference: Concepts*.

❺Describing a DATA Step View The following example uses the DESCRIBE statement in a DATA step view to write a copy of the source code to the SAS log.

```
data view=inventory;
   describe;
run;
```

For information about the DESCRIBE statement, see the "DESCRIBE Statement" on page 1487.

❻Executing a Stored Compiled DATA Step Program The following example executes a stored compiled DATA step program. It uses the DESCRIBE statement to write a copy of the source code to the SAS log.

```
libname stored 'SAS library';

data pgm=stored.employee_list;
   describe;
   execute;
run;
```

For information about the DESCRIBE statement, see the "DESCRIBE Statement" on page 1487. For information about the EXECUTE statement, see the "EXECUTE Statement" on page 1503.

Examples

Example 1: Creating Multiple Data Files and Using Data Set Options This DATA statement creates more than one data set, and it changes the contents of the output data sets:

```
data error (keep=subject date weight)
     fitness(label='Exercise Study'
             rename=(weight=pounds));
```

The ERROR data set contains three variables. SAS assigns a label to the FITNESS data set and renames the variable *weight* to *pounds*.

Example 2: Creating Input DATA Step Views This DATA step creates an input DATA step view instead of a SAS data file:

```
libname ourlib 'SAS-library';

data ourlib.test / view=ourlib.test;
   set ourlib.fittest;
   tot=sum(of score1-score10);
run;
```

Example 3: Creating a View and a Data File This DATA step creates an input DATA step view named THEIRLIB.TEST and an additional temporary SAS data set named SCORETOT:

```
libname ourlib 'SAS-library-1';
libname theirlib 'SAS-library-2';

data theirlib.test scoretot
   / view=theirlib.test;
   set ourlib.fittest;
   tot=sum(of score1-score10);
run;
```

SAS does not create the data file SCORETOT until a subsequent DATA or PROC step processes the view THEIRLIB.TEST.

Example 4: Storing and Executing a Compiled Program The first DATA step produces a stored compiled program named STORED.SALESFIG:

```
libname in 'SAS-library-1 ';
libname stored 'SAS-library-2 ';

data salesdata / pgm=stored.salesfig;
   set in.sales;
   qtr1tot=jan+feb+mar;
run;
```

SAS creates the data set SALESDATA when it executes the stored compiled program STORED.SALESFIG.

```
data pgm=stored.salesfig;
run;
```

Example 5: Creating a Custom Report The second DATA step in this program produces a custom report and uses the _NULL_ keyword to execute the DATA step without creating a SAS data set:

```
data sales;
   input dept : $10. jan feb mar;
   datalines;
shoes 4344 3555 2666
housewares 3777 4888 7999
appliances 53111 7122 41333
;

data _null_;
   set sales;
   qtr1tot=jan+feb+mar;
   put 'Total Quarterly Sales: '
       qtr1tot dollar12.;
run;
```

Example 6: Using a Password with a Stored Compiled DATA Step Program The first DATA step creates a stored compiled DATA step program called STORED.ITEMS. This program includes the ALTER password, which limits access to the program.

```
libname stored 'SAS-library';

data employees / pgm=stored.items (alter=klondike);
   set sample;
   if TotalItems > 200 then output;
   run;
```

This DATA step executes the stored compiled DATA step program STORED.ITEMS. It uses the DESCRIBE statement to print the source code to the SAS log. Because the program was created with the ALTER password, you must use the password if you use the DESCRIBE statement. If you do not enter the password, SAS will prompt you for it.

```
data pgm=stored.items (alter=klondike);
   describe;
   execute;
run;
```

Example 7: Displaying Nesting Levels The following program has two nesting levels. SAS will generate four log messages, one begin and end message for each nesting level.

```
data _null_ /nesting;
   do i = 1 to 10;
      do j = 1 to 5;
         put i= j=;
      end;
   end;
run;
```

Output 6.4 Nesting Level Debug (partial SAS log)

```
6    data _null_ /nesting;
7       do i = 1 to 10;
            -
            719
NOTE 719-185: *** DO begin level 1 ***.

8           do j = 1 to 5;
               -
               719
NOTE 719-185: *** DO begin level 2 ***.

9              put i= j=;
10          end;
            ---
            720
NOTE 720-185: *** DO end level 2 ***.

11      end;
        ---
        720
NOTE 720-185: *** DO end level 1 ***.

12   run;
```

See Also

Statements:

"DESCRIBE Statement" on page 1487
"EXECUTE Statement" on page 1503
"LINK Statement" on page 1669
"Definition of Data Set Options" on page 10

DATALINES Statement

Specifies that data lines follow.

Valid: in a DATA step

Category: File-handling

Type: Declarative

Aliases: CARDS, LINES

Restriction: Data lines cannot contain semicolons. Use "DATALINES4 Statement" on page 1475 when your data contain semicolons.

Syntax

DATALINES;

Without Arguments

Use the DATALINES statement with an INPUT statement to read data that you enter directly in the program, rather than data stored in an external file.

Details

Using the DATALINES Statement The DATALINES statement is the last statement in the DATA step and immediately precedes the first data line. Use a null statement (a single semicolon) to indicate the end of the input data.

You can use only one DATALINES statement in a DATA step. Use separate DATA steps to enter multiple sets of data.

Reading Long Data Lines SAS handles data line length with the CARDIMAGE system option. If you use CARDIMAGE, SAS processes data lines exactly like 80–byte punched card images padded with blanks. If you use NOCARDIMAGE, SAS processes data lines longer than 80 columns in their entirety. Refer to "CARDIMAGE System Option" on page 1859 for details.

Using Input Options with In-stream Data The DATALINES statement does not provide input options for reading data. However, you can access some options by using the DATALINES statement in conjunction with an INFILE statement. Specify DATALINES in the INFILE statement to indicate the source of the data and then use the options you need. See Example 2 on page 1475.

Comparisons

- Use the DATALINES statement whenever data do not contain semicolons. If your data contain semicolons, use the DATALINES4 statement.
- The following SAS statements also read data or point to a location where data are stored:
  - The INFILE statement points to raw data lines stored in another file. The INPUT statement reads those data lines.
  - The %INCLUDE statement brings SAS program statements or data lines stored in SAS files or external files into the current program.
  - The SET, MERGE, MODIFY, and UPDATE statements read observations from existing SAS data sets.

Examples

Example 1: Using the DATALINES Statement In this example, SAS reads a data line and assigns values to two character variables, NAME and DEPT, for each observation in the DATA step:

```
data person;
   input name $ dept $;
   datalines;
John Sales
Mary Acctng
;
```

Example 2: Reading In-stream Data with Options This example takes advantage of options available with the INFILE statement to read in-stream data lines. With the DELIMITER= option, you can use list input to read data values that are delimited by commas instead of blanks.

```
data person;
   infile datalines delimiter=',';
   input name $ dept $;
   datalines;
John,Sales
Mary,Acctng
;
```

See Also

Statements:

"DATALINES4 Statement" on page 1475
"INFILE Statement" on page 1591

System Option:

"CARDIMAGE System Option" on page 1859

DATALINES4 Statement

Indicates that data lines that contain semicolons follow.

Valid: in a DATA step
Category: File-handling
Type: Declarative
Aliases: CARDS4, LINES4

Syntax

DATALINES4;

Without Arguments

Use the DATALINES4 statement together with an INPUT statement to read data that contain semicolons that you enter directly in the program.

Details

The DATALINES4 statement is the last statement in the DATA step and immediately precedes the first data line. Follow the data lines with four consecutive semicolons that are located in columns 1 through 4.

Comparisons

Use the DATALINES4 statement when data contain semicolons. If your data do not contain semicolons, use the DATALINES statement.

Examples

In this example, SAS reads data lines that contain internal semicolons until it encounters a line of four semicolons. Execution continues with the rest of the program.

```
data biblio;
   input number citation $50.;
   datalines4;
  KIRK, 1988
2  LIN ET AL., 1995; BRADY, 1993
3  BERG, 1990; ROA, 1994; WILLIAMS, 1992
;;;;
```

See Also

Statements:

"DATALINES Statement" on page 1474

DECLARE Statement, Hash and Hash Iterator Objects

Declares a hash or hash iterator object; creates an instance of and initializes data for a hash or hash iterator object.

Valid: in a DATA step

Category: Action
Type: Executable
Alias: DCL

Syntax

❶**DECLARE** *object object-reference*;

❷**DECLARE** *object object-reference*<(<*argument_tag-1*: *value-1*<, …*argument_tag-n*: *value-n*>>)>;

Arguments

object
specifies the component object. It can be one of the following values:

hash
specifies a hash object. The hash object provides a mechanism for quick data storage and retrieval. The hash object stores and retrieves data based on lookup keys.

See Also: "Using the Hash Object" in *SAS Language Reference: Concepts*

hiter
specifies a hash iterator object. The hash iterator object enables you to retrieve the hash object's data in forward or reverse key order.

See Also: "Using the Hash Iterator Object" in *SAS Language Reference: Concepts*

object-reference
specifies the object reference name for the hash or hash iterator object.

argument_tag
specifies the information that is used to create an instance of the hash object.
There are five valid hash object argument tags:

dataset: '*dataset_name* <(*datasetoption*)>'
Specifies the name of a SAS data set to load into the hash object.
The name of the SAS data set can be a literal or character variable. The data set name must be enclosed in single or double quotation marks. Macro variables must be enclosed in double quotation marks.
You can use SAS data set options when declaring a hash object in the DATASET argument tag. Data set options specify actions that apply only to the SAS data set with which they appear. They enable you to perform the following operations:

- renaming variables
- selecting a subset of observations based on observation number for processing
- selecting observations using the WHERE option
- dropping or keeping variables from a data set loaded into a hash object, or for an output data set that is specified in an OUTPUT method call
- specifying a password for a data set.

The following syntax is used:

```
dcl hash h (dataset: 'x (where = (i > 10))');
```

For a list of SAS data set options, see "Data Set Options by Category" on page 12.

Note: If the data set contains duplicate keys, the default is to keep the first instance in the hash object; subsequent instances are ignored. To store the last instance in the hash object or an error message written to the SAS log if there is a duplicate key, use the DUPLICATE argument tag. △

duplicate: '*option*'
determines whether to ignore duplicate keys when loading a data set into the hash object. The default is to store the first key and ignore all subsequent duplicates. Option can be one of the following values:

'replace' | 'r'
stores the last duplicate key record.

'error' | 'e'
reports an error to the log if a duplicate key is found.

The following example that uses the REPLACE option stores **brown** for the key 620 and **blue** for the key 531. If you use the default, **green** would be stored for 620 and **yellow** would be stored for 531.

```
data table;
  input key data $;
  datalines;
  531 yellow
  620 green
  531 blue
  908 orange
  620 brown
  143 purple
 run;

data _null_;
length key 8 data $ 8;
if (_n_ = 1) then do;
    declare hash myhash(dataset: "table", duplicate: "r");
    rc = myhash.definekey('key');
    rc = myhash.definedata('data');
    myhash.definedone();
 end;

rc = myhash.output(dataset:"otable");
run;
```

hashexp: *n*
The hash object's internal table size, where the size of the hash table is 2^n.

The value of HASHEXP is used as a power-of-two exponent to create the hash table size. For example, a value of 4 for HASHEXP equates to a hash table size of 2^4, or 16. The maximum value for HASHEXP is 20.

The hash table size is not equal to the number of items that can be stored. Imagine the hash table as an array of 'buckets.' A hash table size of 16 would have 16 'buckets.' Each bucket can hold an infinite number of items. The efficiency of the hash table lies in the ability of the hashing function to map items to and retrieve items from the buckets.

You should specify the hash table size relative to the amount of data in the hash object in order to maximize the efficiency of the hash object lookup routines. Try different HASHEXP values until you get the best result. For example, if the hash object contains one million items, a hash table size of 16 (HASHEXP = 4) would

work, but not very efficiently. A hash table size of 512 or 1024 (HASHEXP = 9 or 10) would result in the best performance.

Default: 8, which equates to a hash table size of 2^8 or 256

ordered: '*option*'
Specifies whether or how the data is returned in key-value order if you use the hash object with a hash iterator object or if you use the hash object OUTPUT method.

option can be one of the following values:

'ascending' | 'a'	Data is returned in ascending key-value order. Specifying '**ascending**' is the same as specifying '**yes**'.
'descending' | 'd'	Data is returned in descending key-value order.
'YES' | 'Y'	Data is returned in ascending key-value order. Specifying '**yes**' is the same as specifying '**ascending**'.
'NO' | 'N'	Data is returned in some undefined order.

Default: NO

The argument can also be enclosed in double quotation marks.

multidata: '*option*'
specifies whether multiple data items are allowed for each key.

option can be one of the following values:

'YES' | 'Y'	Multiple data items are allowed for each key.
'NO' | 'N'	Only one data item is allowed for each key.

Default: NO

See Also: "Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

The argument value can also be enclosed in double quotation marks.

suminc: '*variable-name*'
maintains a summary count of hash object keys. The SUMINC argument tag is given a DATA step variable, which holds the sum increment—that is, how much to add to the key summary for each reference to the key. The SUMINC value treats a missing value as zero, like the SUM function. For example, a key summary changes using the current value of the DATA step variable.

```
dcl hash myhash(suminc: 'count');
```

See Also: "Maintaining Key Summaries" in *SAS Language Reference: Concepts*.

See Also: "Initializing Hash Object Data Using a Constructor" and "Declaring and Instantiating a Hash Iterator Object" in *SAS Language Reference: Concepts*.

Details

The Basics

To use a DATA step component object in your SAS program, you must declare and create (instantiate) the object. The DATA step component interface provides a mechanism for accessing predefined component objects from within the DATA step.

For more information about the predefined DATA step component objects, see "Using DATA Step Component Objects" in *SAS Language Reference: Concepts*.

❶Declaring a Hash or Hash Iterator Object

You use the DECLARE statement to declare a hash or hash iterator object.

```
declare hash h;
```

The DECLARE statement tells SAS that the object reference H is a hash object.

After you declare the new hash or hash iterator object, use the _NEW_ operator to instantiate the object. For example, in the following line of code, the _NEW_ operator creates the hash object and assigns it to the object reference H:

```
h = _new_ hash( );
```

❷Using the DECLARE Statement to Instantiate a Hash or Hash Iterator Object As an alternative to the two-step process of using the DECLARE statement and the _NEW_ operator to declare and instantiate a hash or hash iterator object, you can use the DECLARE statement to declare and instantiate the hash or hash iterator object in one step. For example, in the following line of code, the DECLARE statement declares and instantiates a hash object and assigns it to the object reference H:

```
declare hash h( );
```

The previous line of code is equivalent to using the following code:

```
declare hash h;
h = _new_ hash( );
```

A *constructor* is a method that you can use to instantiate a hash object and initialize the hash object data. For example, in the following line of code, the DECLARE statement declares and instantiates a hash object and assigns it to the object reference H. In addition, the hash table size is initialized to a value of 16 (2^4) using the argument tag, HASHEXP.

```
declare hash h(hashexp: 4);
```

Using SAS Data Set Options When Loading a Hash Object SAS data set options can be used when declaring a hash object that uses the DATASET argument tag. Data set options specify actions that apply only to the SAS data set with which they appear. They enable you to perform the following operations:

- renaming variables
- selecting a subset of observations based on observation number for processing
- selecting observations using the WHERE option
- dropping or keeping variables from a data set loaded into a hash object, or for an output data set that is specified in an OUTPUT method call
- specifying a password for a data set.

The following syntax is used:

```
dcl hash h(dataset: 'x (where = (i > 10))');
```

For more examples of using data set options, see Example 4 on page 1482. For a list of data set options, see "Data Set Options by Category" on page 12.

Comparisons

You can use the DECLARE statement and the _NEW_ operator, or the DECLARE statement alone to declare and instantiate an instance of a hash or hash iterator object.

Examples

Example 1: Declaring and Instantiating a Hash Object by Using the DECLARE Statement and _NEW_ Operator This example uses the DECLARE statement to declare a hash object. The _NEW_ operator is used to instantiate the hash object.

```
data _null_;
   length k $15;
   length d $15;
   if _N_ = 1 then do;
      /* Declare and instantiate hash object "myhash" */
      declare hash myhash;
      myhash = _new_ hash( );
      /* Define key and data variables */
      rc = myhash.defineKey('k');
      rc = myhash.defineData('d');
      rc = myhash.defineDone( );
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;
   /* Create constant key and data values */
   rc = myhash.add(key: 'Labrador', data: 'Retriever');
   rc = myhash.add(key: 'Airedale', data: 'Terrier');
   rc = myhash.add(key: 'Standard', data: 'Poodle');
   /* Find data associated with key and write data to log */
   rc = myhash.find(key: 'Airedale');
   if (rc = 0) then
      put d=;
   else
      put 'Key Airedale not found';
run;
```

Example 2: Declaring and Instantiating a Hash Object by Using the DECLARE Statement This example uses the DECLARE statement to declare and instantiate a hash object in one step.

```
data _null_;
   length k $15;
   length d $15;
   if _N_ = 1 then do;
      /* Declare and instantiate hash object "myhash" */
      declare hash myhash( );
      rc = myhash.defineKey('k');
      rc = myhash.defineData('d');
      rc = myhash.defineDone( );
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;
   /* Create constant key and data values */
   rc = myhash.add(key: 'Labrador', data: 'Retriever');
   rc = myhash.add(key: 'Airedale', data: 'Terrier');
   rc = myhash.add(key: 'Standard', data: 'Poodle');
   /* Find data associated with key and write data to log*/
   rc = myhash.find(key: 'Airedale');
   if (rc = 0) then
      put d=;
   else
      put 'Key Airedale not found';
run;
```

Example 3: Instantiating and Sizing a Hash Object This example uses the DECLARE statement to declare and instantiate a hash object. The hash table size is set to 16 (2^4).

```
data _null_;
   length k $15;
   length d $15;
   if _N_ = 1 then do;
      /* Declare and instantiate hash object "myhash". */
      /* Set hash table size to 16. */
      declare hash myhash(hashexp: 4);
      rc = myhash.defineKey('k');
      rc = myhash.defineData('d');
      rc = myhash.defineDone( );
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;
   /* Create constant key and data values */
   rc = myhash.add(key: 'Labrador', data: 'Retriever');
   rc = myhash.add(key: 'Airedale', data: 'Terrier');
   rc = myhash.add(key: 'Standard', data: 'Poodle');
   rc = myhash.find(key: 'Airedale');
   /* Find data associated with key and write data to log*/
   if (rc = 0) then
      put d=;
   else
      put 'Key Airedale not found';
run;
```

Example 4: Using SAS Data Set Options When Loading a Hash Object The following examples use various SAS data set options when declaring a hash object:

```
data x;
retain j 999;
do i = 1 to 20;
   output;
 end;
run;

/* Using the WHERE option. */
data _null_;
  length i 8;
  dcl hash h(dataset: 'x (where =(i > 10))', ordered: 'a');
  h.definekey('i');
  h.definedone();
  h.output(dataset: 'out');
  run;

/* Using the DROP option. */
data _null_;
  length i 8;
  dcl hash h(dataset: 'x (drop = j)', ordered: 'a');
  h.definekey(all: 'y');
  h.definedone();
  h.output(dataset: 'out (where =( i < 8))');
  run;
```

```
/* Using the FIRSTOBS option. */
data _null_;
  length i j 8;
  dcl hash h(dataset: 'x (firstobs=5)', ordered: 'a');
  h.definekey(all: 'y');
  h.definedone();
  h.output(dataset: 'out');
  run;

/* Using the OBS option. */
data _null_;
  length i j 8;
  dcl hash h(dataset: 'x (obs=5)', ordered: 'd');
  h.definekey(all: 'y');
  h.definedone();
  h.output(dataset: 'out (rename =(j=k))');
  run;
```

For a list of SAS data set options, see "Data Set Options by Category" on page 12.

See Also

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Chapter 9, "Hash and Hash Iterator Object Language Elements," on page 2087

"Using DATA Step Component Objects" in *SAS Language Reference: Concepts*

DECLARE Statement, Java Object

Declares a Java object; creates an instance of and initializes data for a Java object.

Valid: in a DATA step
Category: Action
Type: Executable
Alias: DCL

Syntax

❶**DECLARE JAVAOBJ** *object-reference*;

❷**DECLARE JAVAOBJ** *object-reference* ("*java-class*", <*argument-1* , … *argument-n*>);

Arguments

object-reference
specifies the object reference name for the Java object.

java-class
specifies the name of the Java class to be instantiated.

Requirement: The Java class name must be enclosed in either double or single quotation marks.

Requirement: If you specify a Java package path, you must use forward slashes (/) and not periods (.) in the path. For example, an incorrect classname is **"java.util.Hashtable"**. The correct classname is **"java/util/Hashtable"**.

argument
specifies the information that is used to create an instance of the Java object. Valid values for *argument* depend on the Java object.

See also: "❷Using the DECLARE Statement to Instantiate a Java Object" on page 1484

Details

The Basics To use a DATA step component object in your SAS program, you must declare and create (instantiate) the object. The DATA step component interface provides a mechanism for accessing predefined component objects from within the DATA step.

For more information, see "Using DATA Step Component Objects" in *SAS Language Reference: Concepts*.

❶Declaring a Java Object

You use the DECLARE statement to declare a Java object.

```
declare javaobj j;
```

The DECLARE statement tells SAS that the object reference J is a Java object.

After you declare the new Java object, use the _NEW_ operator to instantiate the object. For example, in the following line of code, the _NEW_ operator creates the Java object and assigns it to the object reference J:

```
j = _new_ javaobj("somejavaclass");
```

❷Using the DECLARE Statement to Instantiate a Java Object

Instead of the two-step process of using the DECLARE statement and the _NEW_ operator to declare and instantiate a Java object, you can use the DECLARE statement to declare and instantiate the Java object in one step. For example, in the following line of code, the DECLARE statement declares and instantiates a Java object and assigns the Java object to the object reference J:

```
declare javaobj j("somejavaclass");
```

The preceding line of code is equivalent to using the following code:

```
declare javaobj j;
j = _new_ javaobj("somejavaclass");
```

A *constructor* is a method that you can use to instantiate a component object and initialize the component object data. For example, in the following line of code, the DECLARE statement declares and instantiates a Java object and assigns the Java object to the object reference J. Note that the only required argument for a Java object constructor is the name of the Java class to be instantiated. All other arguments are constructor arguments for the Java class itself. In the following example, the Java class name, **testjavaclass**, is the constructor, and the values **100** and **.8** are constructor arguments.

```
declare javaobj j("testjavaclass", 100, .8);
```

Comparisons

You can use the DECLARE statement and the _NEW_ operator, or the DECLARE statement alone to declare and instantiate an instance of a Java object.

Examples

Example 1: Declaring and Instantiating a Java Object by Using the DECLARE Statement and the _NEW_ Operator In the following example, a simple Java class is created. The DECLARE statement and the _NEW_ operator are used to create an instance of this class.

```
/* Java code */
import java.util.*;
import java.lang.*;

public class simpleclass
{
   public int i;
   public double d;
}


/* DATA step code
data _null_;
   declare javaobj myjo;
   myjo = _new_ javaobj("simpleclass");
run;
```

Example 2: Using the DECLARE Statement to Create and Instantiate a Java Object In the following example, a Java class is created for a hash table. The DECLARE statement is used to create and instantiate an instance of this class by specifying the capacity and load factor. In this example, a wrapper class, **mhash**, is necessary because the DATA step's only numeric type is equivalent to the Java type DOUBLE.

```
/* Java code */
import java.util.*;

public class mhash extends Hashtable;
{
   mhash (double size, double load)
      {
         super ((int)size, (float)load);
      }
}


/* DATA step code */
data _null_;
   declare javaobj h("mhash", 100, .8);
run;
```

See Also

Operator:

"_NEW_ Operator, Java Object" on page 2163

Chapter 9, "Hash and Hash Iterator Object Language Elements," on page 2087

"Using DATA Step Component Objects" in *SAS Language Reference: Concepts*

DELETE Statement

Stops processing the current observation.

Valid: in a DATA step
Category: Action
Type: Executable

Syntax

DELETE;

Without Arguments

When DELETE executes, the current observation is not written to a data set, and SAS returns immediately to the beginning of the DATA step for the next iteration.

Details

The DELETE statement is often used in a THEN clause of an IF-THEN statement or as part of a conditionally executed DO group.

Comparisons

- Use the DELETE statement when it is easier to specify a condition that excludes observations from the data set or when there is no need to continue processing the DATA step statements for the current observation.
- Use the subsetting IF statement when it is easier to specify a condition for including observations.
- Do not confuse the DROP statement with the DELETE statement. The DROP statement excludes variables from an output data set; the DELETE statement excludes observations.

Examples

Example 1: Using the DELETE Statement as Part of an IF-THEN Statement When the value of LEAFWT is missing, the current observation is deleted:

```
if leafwt=. then delete;
```

Example 2: Using the DELETE Statement to Subset Raw Data

```
data topsales;
   infile file-specification;
```

```
   input region office product yrsales;
   if yrsales<100000 then delete;
run;
```

See Also

Statements:

"DO Statement" on page 1491
"DROP Statement" on page 1499
"IF Statement, Subsetting" on page 1581
"IF-THEN/ELSE Statement" on page 1582

DESCRIBE Statement

Retrieves source code from a stored compiled DATA step program or a DATA step view.

Valid: in a DATA step

Category: Action

Type: Executable

Restriction: Use DESCRIBE only with stored compiled DATA step programs and DATA step views.

Requirement: You must specify the PGM= or the VIEW= option in the DATA statement.

Syntax

DESCRIBE;

Without Arguments

Use the DESCRIBE statement to retrieve program source code from a stored compiled DATA step program or a DATA step view. SAS writes the source statements to the SAS log.

Details

Use the DESCRIBE statement without the EXECUTE statement to retrieve source code from a stored compiled DATA step program or a DATA step view. Use the DESCRIBE statement with the EXECUTE statement to retrieve source code and execute a stored compiled DATA step program. For information about how to use these statements with the DATA statement, see "DATA Statement" on page 1465.

See Also

Statements:

"DATA Statement" on page 1465
"EXECUTE Statement" on page 1503

DISPLAY Statement

Displays a window that is created with the WINDOW statement.

Valid: in a DATA step

Category: Window Display

Type: Executable

Syntax

DISPLAY *window*<*.group*> <NOINPUT > <BLANK> <BELL > <DELETE>;

Arguments

***window*< *.group*>**
names the window and group of fields to be displayed. This field is preceded by a period (.).

Tip: If the window has more than one group of fields, give the complete *window.group* specification. If a window contains a single unnamed group, use only *window*.

NOINPUT
specifies that you cannot input values into fields that are displayed in the window.

Default: If you omit NOINPUT, you can input values into unprotected fields that are displayed in the window.

Restriction: If you use NOINPUT in all DISPLAY statements in a DATA step, you *must* include a STOP statement to stop processing the DATA step.

Tip: The NOINPUT option is useful when you want to allow values to be entered into a window at some times but not others. For example, you can display a window once for entering values and a second time for verifying them.

BLANK
clears the window.

Tip: Use the BLANK option when you want to display different groups of fields in a window and you do not want text from the previous group to appear in the current display.

BELL
produces an audible alarm, beep, or bell sound when the window is displayed if your personal computer is equipped with a speaker device that provides sound.

DELETE
deletes the display of the window after processing passes from the DISPLAY statement on which the option appears.

Details

You must create a window in the same DATA step that you use to display it. Once you display a window, the window remains visible until you display another window over it or until the end of the DATA step. When you display a window that contains fields

where you enter values, either enter a value or press ENTER at *each* unprotected field to cause SAS to proceed to the next display. You cannot skip any fields.

While a window is being displayed, use commands and function keys to view other windows, to change the size of the current window, and so on.

A DATA step that contains a DISPLAY statement continues execution until the last observation that is read by a SET, MERGE, UPDATE, MODIFY, or INPUT statement has been processed or until a STOP or ABORT statement is executed. You can also issue the END command on the command line of the window to stop the execution of the DATA step.

You must create a window before you can display it. See the "WINDOW Statement" on page 1797 for a description of how to create windows. A window that is displayed with the DISPLAY statement does not become part of the SAS log or output file.

Examples

This DATA step creates and displays a window named START. The START window fills the entire screen. Both lines of text are centered.

```
data _null_;
   window start
          #5 @28 'WELCOME TO THE SAS SYSTEM'
          #12 @30 'PRESS ENTER TO CONTINUE';
   display start;
   stop;
run;
```

Although the START window in this example does not require you to input any values, you must press ENTER to cause the execution to proceed to the STOP statement. If you omit the STOP statement, the DATA step executes endlessly unless you enter END on the command line of the window.

Note: Because this DATA step does not read any observations, SAS cannot detect an end-of-file to cause DATA step execution to cease. If you add the NOINPUT option to the DISPLAY statement, the window displays quickly and is removed. △

See Also

Statement:

"WINDOW Statement" on page 1797

DM Statement

Submits SAS Program Editor, Log, Procedure Output or text editor commands as SAS statements.

Valid: anywhere

Category: Program Control

Syntax

DM <*window*> *'command(s)'* <*window*> <CONTINUE>;

Arguments

window
specifies the active window. For more information, see “Details” on page 1490.

Default: If you omit the window name, SAS uses the Program Editor window as the default.

'command(s)'
can be any windowing command or text editor command and must be enclosed in single quotation marks. If you want to issue several commands, separate them with semicolons.

CONTINUE
causes SAS to execute any SAS statements that follow the DM statement in the Program Editor window and, if a windowing command in the DM statement called a window, makes that window active.

Tip: Any windows that are activated by the SAS statements (such as the Output window) appear before the window that is to be made active.

Note: For example, if you specify Log as the active window and have other SAS statements that follow the DM statement (for example, in an autoexec file), those statements are not submitted to SAS until control returns to the SAS interface.

Details

Execution occurs when the DM statement is submitted to SAS. You can use this statement to modify the windowing environment:

- Change SAS interface features during a SAS session.
- Change SAS interface features at the beginning of each SAS session by placing the DM statement in an autoexec file.
- Perform utility functions in windowing applications, such as saving a file with the FILE command or clearing a window with the CLEAR command.

Window placement affects the outcome of the statement:

- If you name a window before the commands, those commands apply to that window.
- If you name a window after the commands, SAS executes the commands and then makes that window the active window. The active window is opened and contains the cursor.

Examples

Example 1: Using the DM Statement

- ```
  dm 'color text cyan; color command red';
  ```
- ```
  dm log 'clear; pgm; color numbers green'
     output;
  ```
- ```
  dm 'caps on';
  ```
- ```
  dm log 'clear' output;
  ```

Example 2: Using the CONTINUE Option with SAS Statements That Do Not Activate a Window This example causes SAS to display the first window of the SAS/AF application, executes the DATA step, moves the cursor to the first field of the SAS/AF application window, and makes that window active.

```
dm 'af c=your-program' continue;

data temp;
   . . . more SAS statements . . .
run;
```

Example 3: Using the CONTINUE Option with SAS Statements That Activate a Window This example displays the first window of the SAS/AF application and executes the PROC PRINT step, which activates the OUTPUT window. Closing the OUTPUT window moves the cursor to the last active window..

```
dm 'af c=your-program' continue;

proc print data=temp;
run;
```

DO Statement

Specifies a group of statements to be executed as a unit.

Valid: in a DATA step
Category: Control
Type: Executable

Syntax

DO;
...more SAS statements...

END;

Without Arguments

Use the DO statement for simple DO group processing.

Details

The DO statement is the simplest form of DO group processing. The statements between the DO and END statements are called a *DO group*. You can nest DO statements within DO groups.

Note: The memory capabilities of your system can limit the number of nested DO statements you can use. For details, see the SAS documentation about how many levels of nested DO statements your system's memory can support. △

A simple DO statement is often used within IF-THEN/ELSE statements to designate a group of statements to be executed depending on whether the IF condition is true or false.

Comparisons

There are three other forms of the DO statement:

- The iterative DO statement executes statements between DO and END statements repetitively based on the value of an index variable. The iterative DO statement can contain a WHILE or UNTIL clause.
- The DO UNTIL statement executes statements in a DO loop repetitively until a condition is true, checking the condition after each iteration of the DO loop.
- The DO WHILE statement executes statements in a DO loop repetitively while a condition is true, checking the condition before each iteration of the DO loop.

Examples

In this simple DO group, the statements between DO and END are performed only when YEARS is greater than 5. If YEARS is less than or equal to 5, statements in the DO group do not execute, and the program continues with the assignment statement that follows the ELSE statement.

```
if years>5 then
   do;
      months=years*12;
      put years= months=;
   end;
   else yrsleft=5-years;
```

See Also

Statements:

"DO Statement, Iterative" on page 1492
"DO UNTIL Statement" on page 1496
"DO WHILE Statement" on page 1497

DO Statement, Iterative

Executes statements between the DO and END statements repetitively, based on the value of an index variable.

Valid: in a DATA step

Category: Control

Type: Executable

Syntax

DO *index-variable=specification-1* <, . . . *specification-n*>;
. . . *more SAS statements* . . .

END;

Arguments

index-variable
names a variable whose value governs execution of the DO group. The *index -variable* argument is required.

Tip: Unless you specify to drop it, the index variable is included in the data set that is being created.

CAUTION:
Avoid changing the index variable within the DO group. If you modify the index variable within the iterative DO group, you might cause infinite looping. △

specification
denotes an expression or a series of expressions in this form
start <TO *stop*> <BY *increment*>
<WHILE(*expression*) | UNTIL(*expression*)>

Requirement: The iterative DO statement requires at least one *specification* argument.

Tip: The order of the optional TO and BY clauses can be reversed.

Tip: When you use more than one *specification*, each one is evaluated before its execution.

start
specifies the initial value of the index variable.

Restriction: When it is used with TO *stop* or BY *increment*, *start* must be a number or an expression that yields a number.

Explanation: When it is used without TO *stop* or BY *increment*, the value of *start* can be a series of items expressed in this form:

item-1 <, . . . item-n >;

The items can be either all numeric or all character constants, or they can be variables. Enclose character constants in quotation marks. The DO group is executed once for each value in the list. If a WHILE condition is added, it applies only to the item that it immediately follows.

The DO group is executed first with *index-variable* equal to *start*. The value of *start* is evaluated before the first execution of the loop.

Featured in: Example 1 on page 1494

TO *stop*
specifies the ending value of the index variable. This argument is optional.

Restriction: *Stop* must be a number or an expression that yields a number.

Explanation: When both *start* and *stop* are present, execution continues (based on the value of *increment*) until the value of *index-variable* passes the value of *stop*. When only *start* and *increment* are present, execution continues (based on the value of *increment*) until a statement directs execution out of the loop, or until a WHILE or UNTIL expression that is specified in the DO statement is satisfied. If neither *stop* nor *increment* is specified, the group executes according to the value of *start*. The value of *stop* is evaluated before the first execution of the loop.

Tip: Any changes to *stop* made within the DO group do not affect the number of iterations. To stop iteration of a loop before it finishes processing, change the

value of *index-variable* so that it passes the value of *stop*, or use a LEAVE statement to go to a statement outside the loop.

Featured in: Example 1 on page 1494

BY *increment*
specifies a positive or negative number (or an expression that yields a number) to control the incrementing of *index-variable*. This argument is optional.

Explanation: The value of *increment* is evaluated before the execution of the loop. Any changes to the increment that are made within the DO group do not affect the number of iterations. If no increment is specified, the index variable is increased by 1. When *increment* is positive, *start* must be the lower bound and *stop*, if present, must be the upper bound for the loop. If *increment* is negative, *start* must be the upper bound and *stop*, if present, must be the lower bound for the loop.

Featured in: Example 1 on page 1494

WHILE(*expression*) | UNTIL(*expression*)
evaluates, either before or after execution of the DO group, any SAS expression that you specify. Enclose the expression in parentheses. This argument is optional.

Restriction: A WHILE or UNTIL specification affects only the last item in the clause in which it is located.

Explanation: A WHILE expression is evaluated before each execution of the loop, so that the statements inside the group are executed repetitively while the expression is true. An UNTIL expression is evaluated after each execution of the loop, so that the statements inside the group are executed repetitively until the expression is true.

Featured in: Example 1 on page 1494

See Also: "DO WHILE Statement" on page 1497 and "DO UNTIL Statement" on page 1496 for more information.

Comparisons

There are three other forms of the DO statement:

- The DO statement, the simplest form of DO-group processing, designates a group of statements to be executed as a unit, usually as a part of IF-THEN/ELSE statements.
- The DO UNTIL statement executes statements in a DO loop repetitively until a condition is true, checking the condition after each iteration of the DO loop.
- The DO WHILE statement executes statements in a DO loop repetitively while a condition is true, checking the condition before each iteration of the DO loop.

Examples

Example 1: Using Various Forms of the Iterative DO Statement

- These iterative DO statements use a list of items for the value of *start*:
  - `do month='JAN','FEB','MAR';`
  - `do count=2,3,5,7,11,13,17;`
  - `do i=5;`
  - `do i=var1, var2, var3;`
  - `do i='01JAN2001'd,'25FEB2001'd,'18APR2001'd;`

- These iterative DO statements use the *start* TO *stop* syntax:
  - `do i=1 to 10;`
  - `do i=1 to exit;`
  - `do i=1 to x-5;`
  - `do i=1 to k-1, k+1 to n;`
  - `do i=k+1 to n-1;`
- These iterative DO statements use the BY *increment* syntax:
  - `do i=n to 1 by -1;`
  - ```
    do i=.1 to .9 by .1, 1 to 10 by 1,
         20 to 100 by 10;
    ```
  - `do count=2 to 8 by 2;`
- These iterative DO statements use WHILE and UNTIL clauses:
  - `do i=1 to 10 while(x<y);`
  - `do i=2 to 20 by 2 until((x/3)>y);`
  - `do i=10 to 0 by -1 while(month='JAN');`
- In this example, the DO loop is executed when I=1 and I=2; the WHILE condition is evaluated when I=3, and the DO loop is executed if the WHILE condition is true.

```
DO I=1,2,3 WHILE (condition);
```

Example 2: Using the Iterative DO Statement without Infinite Looping In each of the following examples, the DO group executes ten times. The first example demonstrates the preferred approach.

```
   /* correct coding */
do i=1 to 10;
   ...more SAS statements...
end;
```

The next example uses the TO and BY arguments.

```
do i=1 to n by m;
   ...more SAS statements...
   if i=10 then leave;
end;
if i=10 then put 'EXITED LOOP';
```

Example 3: Stopping the Execution of the DO Loop In this example, setting the value of the index variable to the current value of EXIT causes the loop to terminate.

```
data iterate1;
   input x;
   exit=10;
   do i=1 to exit;
      y=x*normal(0);
         /* if y>25,           */
         /* changing i's value */
         /* stops execution    */
      if y>25 then i=exit;
      output;
   end;
```
```

```
   datalines;
5
000
2500
;
```

See Also

Statements:

"ARRAY Statement" on page 1440
"Array Reference Statement" on page 1445
"DO Statement" on page 1491
"DO UNTIL Statement" on page 1496
"DO WHILE Statement" on page 1497
"GO TO Statement" on page 1579

DO UNTIL Statement

Executes statements in a DO loop repetitively until a condition is true.

Valid: in a DATA step

Category: Control

Type: Executable

Syntax

DO UNTIL (*expression*);
...more SAS statements...

END;

Arguments

(expression)
is any SAS expression, enclosed in parentheses. You must specify at least one *expression*.

Details

The expression is evaluated at the bottom of the loop after the statements in the DO loop have been executed. If the expression is true, the DO loop does not iterate again.

Note: The DO loop always iterates at least once. △

Comparisons

There are three other forms of the DO statement:

- The DO statement, the simplest form of DO-group processing, designates a group of statements to be executed as a unit, usually as a part of IF-THEN/ELSE statements.
- The iterative DO statement executes statements between DO and END statements repetitively based on the value of an index variable.
- The DO WHILE statement executes statements in a DO loop repetitively while a condition is true, checking the condition before each iteration of the DO loop. The DO UNTIL statement evaluates the condition at the bottom of the loop; the DO WHILE statement evaluates the condition at the top of the loop.

  Note: The statements in a DO UNTIL loop always execute at least one time, whereas the statements in a DO WHILE loop do not iterate even once if the condition is false. △

Examples

These statements repeat the loop until N is greater than or equal to 5. The expression N>=5 is evaluated at the bottom of the loop. There are five iterations in all (0, 1, 2, 3, 4).

```
n=0;
   do until(n>=5);
      put n=;
      n+1;
   end;
```

See Also

Statements:

"DO Statement" on page 1491
"DO Statement, Iterative" on page 1492
"DO WHILE Statement" on page 1497

DO WHILE Statement

Executes statements in a DO-loop repetitively while a condition is true.

Valid: in a DATA step
Category: Control
Type: Executable

Syntax

DO WHILE (*expression*);
...more SAS statements...

END;

Arguments

(*expression*)
is any SAS expression, enclosed in parentheses. You must specify at least one *expression*.

Details

The expression is evaluated at the top of the loop before the statements in the DO loop are executed. If the expression is true, the DO loop iterates. If the expression is false the first time it is evaluated, the DO loop does not iterate even once.

Comparisons

There are three other forms of the DO statement:

- The DO statement, the simplest form of DO-group processing, designates a group of statements to be executed as a unit, usually as a part of IF-THEN/ELSE statements.
- The iterative DO statement executes statements between DO and END statements repetitively based on the value of an index variable.
- The DO UNTIL statement executes statements in a DO loop repetitively until a condition is true, checking the condition after each iteration of the DO loop. The DO WHILE statement evaluates the condition at the top of the loop; the DO UNTIL statement evaluates the condition at the bottom of the loop.

  Note: If the expression is false, the statements in a DO WHILE loop do not execute. However, because the DO UNTIL expression is evaluated at the bottom of the loop, the statements in the DO UNTIL loop always execute at least once. △

Examples

These statements repeat the loop while N is less than 5. The expression N<5 is evaluated at the top of the loop. There are five iterations in all (0, 1, 2, 3, 4).

```
n=0;
   do while(n<5);
      put n=;
      n+1;
   end;
```

See Also

Statements:

"DO Statement" on page 1491
"DO Statement, Iterative" on page 1492
"DO UNTIL Statement" on page 1496

DROP Statement

Excludes variables from output SAS data sets.

Valid: in a DATA step
Category: Information
Type: Declarative

Syntax

DROP *variable-list*;

Arguments

variable-list
specifies the names of the variables to omit from the output data set.

Tip: You can list the variables in any form that SAS allows.

Details

The DROP statement applies to all the SAS data sets that are created within the same DATA step and can appear anywhere in the step. The variables in the DROP statement are available for processing in the DATA step. If no DROP or KEEP statement appears, all data sets that are created in the DATA step contain all variables. Do not use both DROP and KEEP statements within the same DATA step.

Comparisons

- The DROP statement differs from the DROP= data set option in the following ways:
  - You cannot use the DROP statement in SAS procedure steps.
  - The DROP statement applies to all output data sets that are named in the DATA statement. To exclude variables from some data sets but not from others, use the DROP= data set option in the DATA statement.
- The KEEP statement is a parallel statement that specifies a list of variables to write to output data sets. Use the KEEP statement instead of the DROP statement if the number of variables to include is significantly smaller than the number to omit.
- Do not confuse the DROP statement with the DELETE statement. The DROP statement excludes variables from output data sets; the DELETE statement excludes observations.

Examples

- These examples show the correct syntax for listing variables with the DROP statement:
  - `drop time shift batchnum;`
  - `drop grade1-grade20;`

□ In this example, the variables PURCHASE and REPAIR are used in processing but are not written to the output data set INVENTRY:

```
data inventry;
   drop purchase repair;
   infile file-specification;
   input unit part purchase repair;
   totcost=sum(purchase,repair);
run;
```

See Also

Data Set Option:

"DROP= Data Set Option" on page 22

Statements:

"DELETE Statement" on page 1486

"KEEP Statement" on page 1648

END Statement

Ends a DO group or SELECT group processing.

Valid: in a DATA step

Category: Control

Type: Declarative

Syntax

END;

Without Arguments

Use the END statement to end DO group or SELECT group processing.

Details

The END statement must be the last statement in a DO group or a SELECT group.

Examples

This example shows a simple DO group and a simple SELECT group:

□
```
do;
      . . .more SAS statements. . .
end;
```

□
```
select(expression);
   when(expression) SAS statement;
   otherwise SAS statement;
end;
```

See Also

Statements:

"DO Statement" on page 1491

"SELECT Statement" on page 1760

ENDSAS Statement

Terminates a SAS job or session after the current DATA or PROC step executes.

Valid: anywhere

Category: Program Control

Syntax

ENDSAS;

Without Arguments

The ENDSAS statement terminates a SAS job or session.

Details

ENDSAS is most useful in interactive or windowing sessions.

Note: ENDSAS statements are always executed at the point that they are encountered in a DATA step. Use the ABORT RETURN statement to stop processing when an error condition occurs—for example, in the clause of an IF-THEN statement or a SELECT statement. △

Comparisons

You can also terminate a SAS job or session by using the BYE or the ENDSAS command from any SAS window command line. For details, refer to the online Help for SAS windows.

See Also

"SYSSTARTID Automatic Macro Variable" in *SAS Macro Language: Reference*

ERROR Statement

Sets _ERROR_ to 1. A message written to the SAS log is optional.

Valid: in a DATA step
Category: Action
Type: Executable

Syntax

ERROR <*message*>;

Without Arguments

Using ERROR without an argument sets the automatic variable _ERROR_ to 1 writes a blank message to the log.

Arguments

message
writes a message to the log.

Tip: *Message* can include character literals (enclosed in quotation marks), variable names, formats, and pointer controls.

Details

The ERROR statement sets the automatic variable _ERROR_ to 1. Writing a message that you specify to the SAS log is optional. When _ERROR_ = 1, SAS writes the data lines that correspond to the current observation in the SAS log.

Using ERROR is equivalent to using these statements in combination:

- an assignment statement setting _ERROR_ to 1
- a FILE LOG statement
- a PUT statement (if you specify a message)
- a PUT; statement (if you do not specify a message)
- another FILE statement resetting FILE to any previously specified setting.

Examples

In the following examples, SAS writes the error message and the variable name and value to the log for each observation that satisfies the condition in the IF-THEN statement.

- In this example, the ERROR statement automatically resets the FILE statement specification to the previously specified setting.

```
file file-specification;
   if type='teen' & age > 19 then
      error 'type and age don"t match ' age=;
```

- This example uses a series of statements to produce the same results.`file file-specification;`

```
 if type='teen' & age > 19 then
```

```
do;
   file log;
   put 'type and age don"t match ' age=;
   _error_=1;
   file file-specification;
end;
```

See Also

Statement:
"PUT Statement" on page 1708

EXECUTE Statement

Executes a stored compiled DATA step program .

Valid: in a DATA step
Category: Action
Type: Executable
Restriction: Use EXECUTE with stored compiled DATA step programs only.
Requirement: You must specify the PGM= option in the DATA step.

Syntax

EXECUTE;

Without Arguments

The EXECUTE statement executes a stored compiled DATA step program.

Details

Use the DESCRIBE statement with the EXECUTE statement in the same DATA step to retrieve the source code and execute a stored compiled DATA step program. If you do not specify either statement, EXECUTE is assumed. The order in which you use the statements is interchangeable. The DATA step program executes when it reaches a step boundary. For information about how to use these statements with the DATA statement, see "DATA Statement" on page 1465.

See Also

Statements:
"DATA Statement" on page 1465
"DESCRIBE Statement" on page 1487

FILE Statement

Specifies the current output file for PUT statements.

Valid: in a DATA step
Category: File-handling
Type: Executable
See: FILE Statement in the documentation for your operating environment.

Syntax

FILE *file-specification* <*device-type*> <*options*> <*operating-environment-options*>;

Arguments

file-specification
identifies an external file that the DATA step uses to write output from a PUT statement. *File-specification* can have these forms:

'external-file'
specifies the physical name of an external file, which is enclosed in quotation marks. The physical name is the name by which the operating environment recognizes the file.

fileref
specifies the fileref of an external file.

Requirement: You must have previously associated *fileref* with an external file in a FILENAME statement or function, or in an appropriate operating environment command. There is only one exception to this rule: when you use the FILEVAR= option, the fileref is simply a placeholder.

See Also: "FILENAME Statement" on page 1520

fileref(file)
specifies a fileref that is previously assigned to an external file that is an aggregate grouping of files. Follow the fileref with the name of a file or member, which is enclosed in parentheses.

Note: A file that is located in an aggregate storage location and has a name that is not a valid SAS name must have its name enclosed in quotation marks. △

Requirement: You must previously associate *fileref* with an external file in a FILENAME statement or function, or in an appropriate operating environment command.

See Also: "FILENAME Statement" on page 1520

Operating Environment Information: Different operating environments call an aggregate grouping of files by different names, such as a directory, a MACLIB, or a partitioned data set. For details, see the SAS documentation for your operating environment. △

LOG
is a reserved fileref that directs the output that is produced by any PUT statements to the SAS log.

At the beginning of each execution of a DATA step, the fileref that indicates where the PUT statements write is automatically set to LOG. Therefore, the first PUT statement in a DATA step always writes to the SAS log, unless it is preceded by a FILE statement that specifies otherwise.

Tip: Because output lines are by default written to the SAS log, use a FILE LOG statement to restore the default action or to specify additional FILE statement options.

PRINT
is a reserved fileref that directs the output that is produced by any PUT statements to the same file as the output that is produced by SAS procedures.

Interaction: When you write to a file, the value of the N= option must be either 1 or PAGESIZE.

Tip: When PRINT is the fileref, SAS uses carriage-control characters and writes the output with the characteristics of a print file.

See Also: A complete discussion of print files in *SAS Language Reference: Concepts*

Operating Environment Information: The carriage-control characters that are written to a file can be specific to the operating environment. For details, see the SAS documentation for your operating environment. △

Tip: If the file does not exist in the directory that you specify for file-specification, SAS creates the file. If the directory specified in *file-specification* does not exist, SAS sets the SYSERR macro variable, which can be checked if the ERRORCHECK option is set to STRICT.

device-type
specifies the type of device or the access method that is used if the fileref points to an input or output device or a location that is not a physical file:

DISK
specifies that the device is a disk drive.

Tip: When you assign a fileref to a file on disk, you are not required to specify DISK.

DUMMY
specifies that the output to the file is discarded.

Tip: Specifying DUMMY can be useful for testing.

GTERM
indicates that the output device type is a graphics device that will receive graphics data.

PIPE
specifies an unnamed pipe.

Note: Some operating environments do not support pipes. △

PLOTTER
specifies an unbuffered graphics output device.

PRINTER
specifies a printer or printer spool file.

TAPE
specifies a tape drive.

TEMP
creates a temporary file that exists only as long as the filename is assigned. The temporary file can be accessed only through the logical name and is available only while the logical name exists.

Restriction: Do not specify a physical pathname. If you do, SAS returns an error.

Tip: Files manipulated by the TEMP device can have the same attributes and behave identically to DISK files.

TERMINAL
specifies the user's terminal.

UPRINTER
specifies a Universal Printing printer definition name.

Tip: If you do not specify the printer name in the FILENAME statement, the PRINTERPATH options control which Universal Printer is used and the destination of the output.

Alias: DEVICE=

Requirement: *device-type* must appear right after the physical path. DEVICE=*device-type* can appear anywhere in the statement.

Operating Environment Information: Additional specifications might be required when you specify some devices. See the SAS documentation for your operating environment before specifying a value other than DISK. Values in addition to the ones listed here might be available in some operating environments. △

Options

BLKSIZE=*block-size*
specifies the block size of the output file.

Default: Dependent on your operating environment.

Operating Environment Information: For details, see the FILE Statement in the SAS documentation for your operating environment. △

COLUMN=*variable*
specifies a variable that SAS automatically sets to the current column location of the pointer. This variable, like automatic variables, is not written to the data set.

Alias: COL=

See Also: LINE= on page 1509

DELIMITER= *delimiter(s)*
specifies an alternate delimiter (other than blank) to be used for LIST output where *delimiter* is

'*list-of-delimiting-characters*'
specifies one or more characters to write as delimiters.

Requirement: Enclose the list of characters in quotation marks.

character-variable
specifies a character variable whose value becomes the delimiter.

Alias: DLM=

Default: blank space

Restriction: Even though a character string or character variable is accepted, only the first character of the string or variable is used as the output delimiter. The FILE DLM= processing differs from INFILE DELIMITER= processing.

Interaction: Output that contains embedded delimiters requires the delimiter sensitive data (DSD) option.

Tip: DELIMITER= can be used with the colon (:) modifier (modified LIST output).

Tip: The delimiter is case sensitive.

See Also: DLMSTR= on page 1506, DSD (delimiter sensitive data) on page 1507

DLMSTR= *delimiter*
specifies a character string as an alternate delimiter (other than a blank) to be used for LIST output, where *delimiter* is

'*delimiting-string*'
specifies a character string to write as a delimiter.

Requirement: Enclose the string in quotation marks.

character-variable
specifies a character variable whose value becomes the delimiter.

Default: blank space

Interaction: If you specify more than one DLMSTR= option in the FILE statement, the DLMSTR= option that is specified last will be used. If you specify both the DELIMITER= and DLMSTR= options, the option that is specified last will be used.

Interaction: If you specify RECFM=N, make sure that the LRECL is large enough to hold the largest input item. Otherwise, it might be possible for the delimiter to be split across the record boundary.

See Also: DELIMITER= on page 1506, DLMSOPT= on page 1507, DSD (delimiter sensitive data) on page 1507

DLMSOPT= 'T' | 't'
specifies a parsing option for the DLMSTR= T option that removes trailing blanks of the string delimiter.

Requirement: The DLMSOPT=T option has an effect only when used with the DLMSTR= option.

Tip: The DLMSOPT=T option is useful when you use a variable as the delimiter string

See Also: DLMSTR= on page 1506

DROPOVER
discards data items that exceed the output line length (as specified by the LINESIZE= or LRECL= options in the FILE statement).

Default: FLOWOVER

Explanation: By default, data that exceeds the current line length is written on a new line. When you specify DROPOVER, SAS drops (or ignores) an entire item when there is not enough space in the current line to write it. When an entire item is dropped, the column pointer remains positioned after the last value that is written in the current line. Thus, the PUT statement might write other items in the current output line if they fit in the space that remains or if the column pointer is repositioned. When a data item is dropped, the DATA step continues normal execution (_ERROR_=0). At the end of the DATA step, a message is printed for each file from which data was lost.

Tip: Use DROPOVER when you want the DATA step to continue executing if the PUT statement attempts to write past the current line length, but you do not want the data item that exceeds the line length to be written on a new line.

See Also: FLOWOVER on page 1509 and STOPOVER on page 1513

DSD (delimiter sensitive data)
specifies that data values that contain embedded delimiters, such as tabs or commas, be enclosed in quotation marks. The DSD option enables you to write data values that contain embedded delimiters to LIST output. This option is ignored for other types of output (for example, formatted, column, and named). Any double quotation marks that are included in the data value are repeated. When a variable value contains the delimiter and DSD is used in the FILE statement, the variable value will be enclosed in double quotation marks when the output is generated. For example, the following code

```
DATA _NULL_;
  FILE log dsd;
  x='"lions, tigers, and bears"';
```

```
  put x ' "Oh, my!"';
run;
```

will result in the following output:

```
"""lions, tigers, and bears""", "Oh, my!"
```

If a quoted (text) string contains the delimiter and DSD is used in the FILE statement, then the quoted string will not be enclosed in double quotation marks when used in a PUT statement. For example, the following code

```
DATA _NULL_;
  FILE log dsd;
  PUT 'lions, tigers, and bears';
run;
```

will result in the following output:

```
lions, tigers, and bears
```

Interaction: If you specify DSD, the default delimiter is assumed to be the comma (,). Specify the DELIMITER= or DLMSTR= option if you want to use a different delimiter.

Tip: By default, data values that do not contain the delimiter that you specify are not enclosed in quotation marks. However, you can use the tilde (~) modifier to force any data value, including missing values, to be enclosed in quotation marks, even if it contains no embedded delimiter.

See Also: DELIMITER= on page 1506, DLMSTR= on page 1506

ENCODING= '*encoding-value*'
specifies the encoding to use when writing to the output file. The value for ENCODING= indicates that the output file has a different encoding from the current session encoding.

When you write data to the output file, SAS transcodes the data from the session encoding to the specified encoding.

Default: SAS uses the current session encoding.

See Also: "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*

Featured in: Example 8 on page 1519

FILENAME=*variable*
defines a character variable, whose name you supply, that SAS sets to the value of the physical name of the file currently open for PUT statement output. The physical name is the name by which the operating environment recognizes the file.

Tip: This variable, like automatic variables, is not written to the data set.

Tip: Use a LENGTH statement to make the variable length long enough to contain the value of the physical filename if it is longer than eight characters (the default length of a character variable).

See Also: FILEVAR= on page 1508

Featured in: Example 4 on page 1517

FILEVAR=*variable*
defines a variable whose change in value causes the FILE statement to close the current output file and open a new one the next time the FILE statement executes. The next PUT statement that executes writes to the new file that is specified as the value of the FILEVAR= variable.

Restriction: The value of a FILEVAR= variable is expressed as a character string that contains a physical filename.

Interaction: When you use the FILEVAR= option, the *file-specification* is just a placeholder, not an actual filename or a fileref that has been previously assigned to a file. SAS uses this placeholder for reporting processing information to the SAS log. It must conform to the same rules as a fileref.

Tip: This variable, like automatic variables, is not written to the data set.

Tip: If any of the physical filenames is longer than eight characters (the default length of a character variable), assign the FILEVAR= variable a longer length with another statement, such as a LENGTH statement or an INPUT statement.

See Also: FILENAME= on page 1508

Featured in: Example 5 on page 1517

FLOWOVER

causes data that exceeds the current line length to be written on a new line. When a PUT statement attempts to write beyond the maximum allowed line length (as specified by the LINESIZE= option in the FILE statement), the current output line is written to the file and the data item that exceeds the current line length is written to a new line.

Default: FLOWOVER

Interaction: If the PUT statement contains a trailing @, the pointer is positioned after the data item on the new line, and the next PUT statement writes to that line. This process continues until the end of the input data is reached or until a PUT statement without a trailing @ causes the current line to be written to the file.

See Also: DROPOVER on page 1507 and STOPOVER on page 1513

FOOTNOTES | NOFOOTNOTES

controls whether currently defined footnotes are printed.

Alias: FOOTNOTE | NOFOOTNOTE

Requirement: In order to print footnotes in a DATA step report, you must set the FOOTNOTE option in the FILE statement.

Default: NOFOOTNOTES

HEADER=*label*

defines a statement label that identifies a group of SAS statements that you want to execute each time SAS begins a new output page.

Restriction: The first statement after the label must be an executable statement. Thereafter you can use any SAS statement.

Restriction: Use the HEADER= option only when you write to print files.

Tip: To prevent the statements in this group from executing with each iteration of the DATA step, use two RETURN statements: one precedes the label and the other appears as the last statement in the group.

Featured in: Example 1 on page 1516

LINE=*variable*

defines a variable whose value is the current relative line number within the group of lines available to the output pointer. You supply the variable name; SAS automatically assigns the value.

Range: 1 to the value that is specified by the N= option or with the #*n* line pointer control. If neither is specified, the LINE= variable has a value of 1.

Tip: This variable, like automatic variables, is not written to the data set.

Tip: The value of the LINE= variable is set at the end of PUT statement execution to the number of the next available line.

LINESIZE=*line-size*
sets the maximum number of columns per line for reports and the maximum record length for data files.

Alias: LS=

Default: The default LINESIZE= value is determined by one of two options:

- the LINESIZE= system option when you write to a file that contains carriage-control characters or to the SAS log.
- the LRECL= option in the FILE statement when you write to a file.

Range: From 64 to the maximum logical record length that is allowed in your operating environment.

Operating Environment Information: The highest value allowed for LINESIZE= is dependent on your operating environment. For details, see the SAS documentation for your operating environment. △

Interaction: If a PUT statement tries to write a line that is longer than the value that is specified by the LINESIZE= option, the action that is taken is determined by whether FLOWOVER, DROPOVER, or STOPOVER is in effect. By default (FLOWOVER), SAS writes the line as two or more separate records.

Comparisons: LINESIZE= tells SAS how much of the line to use. LRECL= specifies the physical record length of the file.

See Also: LRECL= on page 1510, DROPOVER on page 1507, FLOWOVER on page 1509, and STOPOVER on page 1513

Featured in: Example 6 on page 1518

LINESLEFT=*variable*
defines a variable whose value is the number of lines left on the current page. You supply the variable name; SAS assigns the value of the number of lines left on the current page to that variable. The value of the LINESLEFT= variable is set at the end of PUT statement execution.

Alias: LL=

Tip: This variable, like automatic variables, is not written to the data set.

Featured in: Example 2 on page 1516

LRECL=*logical-record-length*
specifies the logical record length of the output file.

Operating Environment Information: Values for *logical-record-length* are dependent on the operating environment. For details, see the SAS documentation for your operating environment. △

Default: If you omit the LRECL= option, SAS chooses a value based on the operating environment's file characteristics.

Comparisons: LINESIZE= tells SAS how much of the line to use; LRECL= specifies the physical line length of the file.

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

See Also: LINESIZE= on page 1510, PAD on page 1512, and PAGESIZE= on page 1512

MOD
writes the output lines after any existing lines in the file.

Default: OLD

Restriction: MOD is not accepted under all operating environments.

Operating Environment Information: For more information, see the SAS documentation for your operating environment. △

Restriction: Do not use the MOD option with any ODS destination other than the Listing destination. Otherwise, you might receive unexpected output.

See Also: OLD on page 1512

N=*available-lines*
specifies the number of lines that you want available to the output pointer in the current iteration of the DATA step. *Available-lines* can be expressed as a number (*n*) or as the keyword PAGESIZE or PS.

n
specifies the number of lines that are available to the output pointer. The system can move back and forth between the number of lines that are specified while composing them before moving on to the next set.

PAGESIZE
specifies that the entire page is available to the output pointer.

Alias: PS

Restriction: N=PAGESIZE is valid only when output is printed.

Restriction: If the current output file is a file that is to be printed, *available-lines* must have a value of either 1 or PAGESIZE.

Interactions: There are two ways to control the number of lines available to the output pointer:

- the N= option
- the #*n* line pointer control in a PUT statement.

Interaction: If you omit the N= option and no # pointer controls are used, one line is available; that is, by default, N=1. If N= is not used but there are # pointer controls, N= is assigned the highest value that is specified for a # pointer control in any PUT statement in the current DATA step.

Tip: Setting N=PAGESIZE enables you to compose a page of multiple columns one column at a time.

Featured in: Example 3 on page 1517

ODS < = (*ODS-suboptions*) >
specifies to use the Output Delivery System to format the output from a DATA step. It defines the structure of the data component and holds the results of the DATA step and binds that component to a table definition to produce an output object. ODS sends this object to all open ODS destinations, each of which formats the output appropriately. For information about the *ODS-suboptions*, see the "FILE Statement for ODS". For general information about the Output Delivery System, see *SAS Output Delivery System: User's Guide*.

Default: If you omit the ODS suboptions, the DATA step uses a default table definition (base.datastep.table) that is stored in the SASHELP.TMPLMST template store. This definition defines two generic columns: one for character variables, and one for numeric variables. ODS associates each variable in the DATA step with one of these columns and displays the variables in the order in which they are defined in the DATA step.

Without suboptions, the default table definition uses the variable's label as its column heading. If no label exists, the definition uses the variable's name as the column heading.

Requirement: The ODS option is valid only when you use the fileref PRINT in the FILE statement.

Restriction: You cannot use _FILE_=, FILEVAR=, HEADER=, and PAD with the ODS option.

Interaction: The DELIMITER= and DSD options have no effect on the ODS option. The FOOTNOTES|NOFOOTNOTES, LINESIZE, PAGESIZE, and TITLES | NOTITLES options have an effect only on the LISTING destination.

OLD
replaces the previous contents of the file.

Default: OLD

Restriction: OLD is not accepted under all operating environments.

Operating Environment Information: For details, see the SAS documentation for your operating environment. △

See Also: MOD on page 1510

PAD | NOPAD
controls whether records written to an external file are padded with blanks to the length that is specified in the LRECL= option.

Default: NOPAD is the default when writing to a variable-length file; PAD is the default when writing to a fixed-length file.

Tip: PAD provides a quick way to create fixed-length records in a variable-length file.

See Also: LRECL= on page 1510

PAGESIZE=*value*
sets the number of lines per page for your reports.

Alias: PS=

Default: the value of the PAGESIZE= system option.

Range: The value can range from 15 to 32767.

Interaction: If any TITLE statements are currently defined, the lines they occupy are included in counting the number of lines for each page.

Tip: After the value of the PAGESIZE= option is reached, the output pointer advances to line 1 of a new page.

Tip: If you specify FILE LOG, the number of lines that are output on the first page is reduced by the number of lines in the SAS startup notes. For example, if PAGESIZE=20 and there are nine lines of SAS startup notes, only 11 lines are available for output on the first page.

See Also: "PAGESIZE= System Option" on page 1958

PRINT | NOPRINT
controls whether carriage-control characters are placed in the output lines.

Operating Environment Information: The carriage-control characters that are written to a file can be specific to the operating environment. For details, see the SAS documentation for your operating environment. △

Restriction: When you write to a file, the value of the N= option must be either 1 or PAGESIZE.

Tip: The PRINT option is not necessary if you are using fileref PRINT.

Tip: If you specify FILE PRINT in an interactive SAS session, then the Output window interprets the form-feed control characters as page breaks, and blank lines that are output before the form feed are removed from the output. Writing the results from the Output window to a flat file produces a file without page break characters. If a file needs to contain the form-feed characters, then the FILE statement should include a physical file location and the PRINT option.

RECFM=*record-format*
specifies the record format of the output file.

Range: Values are dependent on the operating environment.

Operating Environment Information: For details, see the SAS documentation for your operating environment. △

STOPOVER
stops processing the DATA step immediately if a PUT statement attempts to write a data item that exceeds the current line length. In such a case, SAS discards the data item that exceeds the current line length, writes the portion of the line that was built before the error occurred, and issues an error message.

Default: FLOWOVER

See Also: FLOWOVER on page 1509 and DROPOVER on page 1507

TITLES | NOTITLES
controls the printing of the current title lines on the pages of files. When NOTITLES is omitted, or when TITLES is specified, SAS prints any titles that are currently defined.

Alias: TITLE | NOTITLE

Default: TITLES

FILE=*variable*
names a character variable that references the current output buffer of this FILE statement. You can use the variable in the same way as any other variable, even as the target of an assignment. The variable is automatically retained and initialized to blanks. Like automatic variables, the _FILE_= variable is not written to the data set.

Restriction: *variable* cannot be a previously defined variable. Make sure that the _FILE_= specification is the first occurrence of this variable in the DATA step. Do not set or change the length of _FILE_= variable with the LENGTH or ATTRIB statements. However, you can attach a format to this variable with the ATTRIB or FORMAT statement.

Interaction: The maximum length of this character variable is the logical record length (LRECL) for the specified FILE statement. However, SAS does not open the file to know the LRECL until before the execution phase. Therefore, the designated size for this variable during the compilation phase is 32,767.

Tip: Modification of this variable directly modifies the FILE statement's current output buffer. Any subsequent PUT statement for this FILE statement outputs the contents of the modified buffer. The _FILE_= variable accesses only the current output buffer of the specified FILE statement even if you use the N= option to specify multiple output buffers.

Tip: To access the contents of the output buffer in another statement without using the _FILE_= option, use the automatic variable _FILE_.

Main Discussion: "Updating the _FILE_ Variable" on page 1514

Operating Environment Options

Operating Environment Information: For descriptions of operating-environment-specific options in the FILE statement, see the SAS documentation for your operating environment. △

Details

Overview By default, PUT statement output is written to the SAS log. Use the FILE statement to route this output to either the same external file to which procedure output is written or to a different external file. You can indicate whether carriage-control characters should be added to the file. See the PRINT | NOPRINT option on page 1512.

You can use the FILE statement in conditional (IF-THEN) processing because it is executable. You can also use multiple FILE statements to write to more than one external file in a single DATA step.

Operating Environment Information: Using the FILE statement requires operating-environment-specific information. See the SAS documentation for your operating environment before you use this statement. △

You can now use the Output Delivery System with the FILE statement to write DATA step results. This functionality is briefly discussed here. For details, see the "FILE Statement for ODS" in *SAS Output Delivery System: User's Guide*.

Updating an External File in Place You can use the FILE statement with the INFILE and PUT statements to update an external file in place, updating either an entire record or only selected fields within a record. Follow these guidelines:

- Always place the INFILE statement first.
- Specify the same fileref or physical filename in the INFILE and FILE statements.
- Use options that are common to both the INFILE and FILE statements in the INFILE statement. (Any such options that are used in the FILE statement are ignored.)
- Use the SHAREBUFFERS option in the INFILE statement to allow the INFILE and FILE statements to use the same buffer, which saves CPU time and enables you to update individual fields instead of entire records.

Accessing the Contents of the Output Buffer In addition to the _FILE_= variable, you can use the automatic _FILE_ variable to reference the contents of the current output buffer for the most recent execution of the FILE statement. This character variable is automatically retained and initialized to blanks. Like other automatic variables, _FILE_ is not written to the data set.

When you specify the _FILE_= option in a FILE statement, this variable is also indirectly referenced by the automatic _FILE_ variable. If the automatic _FILE_ variable is present and you omit _FILE_= in a particular FILE statement, then SAS creates an internal _FILE_= variable for that FILE statement. Otherwise, SAS does not create the _FILE_= variable for a particular FILE.

During execution and at the point of reference, the maximum length of this character variable is the maximum length of the current _FILE_= variable. However, because _FILE_ merely references other variables whose lengths are not known until before the execution phase, the designated length is 32,767 during the compilation phase. For example, if you assign _FILE_ to a new variable whose length is undefined, the default length of the new variable is 32,767. You cannot use the LENGTH statement and the ATTRIB statement to set or override the length of _FILE_. You can use the FORMAT statement and the ATTRIB statement to assign a format to _FILE_.

Updating the _FILE_ Variable Like other SAS variables, you can update the _FILE_ variable. The following two methods are available:

- Use _FILE_ in an assignment statement.
- Use a PUT statement.

You can update the _FILE_ variable by using an assignment statement that has the following form.

FILE = <*'string-in-quotation-marks'* | *character-expression*>

The assignment statement updates the contents of the current output buffer and sets the buffer length to the length of *'string-in-quotation-marks'* or *character-expression*. However, using an assignment statement does not affect the current column pointer of the PUT statement. The next PUT statement for this FILE statement begins to update the buffer at column 1 or at the last known location when you use the trailing @ in the PUT statement.

In the following example, the assignment statement updates the contents of the current output buffer. The column pointer of the PUT statement is not affected:

```
file print;
_file_ = '_FILE_';
put 'This is PUT';
```

SAS creates the following output: **`This is PUT`**

In this example,

```
file print;
_file_ = 'This is from FILE, sir.';
put @14 'both';
```

SAS creates the following output: **`This is from both, sir.`**

You can also update the _FILE_ variable by using a PUT statement. The PUT statement updates the _FILE_ variable because the PUT statement formats data in the output buffer and _FILE_ points to that buffer. However, by default SAS clears the output buffers after a PUT statement executes and outputs the current record (or N= block of records). Therefore, if you want to examine or further modify the contents of _FILE_ before it is output, include a trailing @ or @@ in any PUT statement (when N=1). For other values of N=, use a trailing @ or @@ in any PUT statement where the last line pointer location is on the last record of the record block. In the following example, when N=1

```
file ABC;
put 'Something' @;
Y = _file_||' is here';
file ABC;
put 'Nothing' ;
Y = _file_||' is here';
```

Y is first assigned **`Something is here`** then Y is assigned **`is here`**.

Any modification of _FILE_ directly modifies the current output buffer for the current FILE statement. The execution of any subsequent PUT statements for this FILE statement will output the contents of the modified buffer.

FILE only accesses the contents of the current output buffer for a FILE statement, even when you use the N= option to specify multiple buffers. You can access all the N= buffers, but you must use a PUT statement with the # line pointer control to make the desired buffer the current output buffer.

Comparisons

- The FILE statement specifies the *output* file for PUT statements. The INFILE statement specifies the *input* file for INPUT statements.
- Both the FILE and INFILE statements allow you to use options that provide SAS with additional information about the external file being used.

- In the Program Editor, Log, and Output windows, the FILE command specifies an external file and writes the contents of the window to the file.

Examples

Example 1: Executing Statements When Beginning a New Page This DATA step illustrates how to use the HEADER= option:

- *Write a report.* Use DATA _NULL_ to write a report rather than create a data set.

```
data _null_;
   set sprint;
   by dept;
```

- *Route output to the SAS output window. Point to the header information.* The PRINT fileref routes output to the same location as procedure output. HEADER= points to the label that precedes the statements that create the header for each page:

```
file print header=newpage;
```

- *Start a new page for each department:*

```
if first.dept then put _page_;
   put @22 salesrep @34 salesamt;
```

- *Write a header on each page.* These statements execute each time a new page is begun. RETURN is necessary before the label and as the final statement in a labeled group:

```
return;
   newpage:
      put @20 'Sales for 1989' /
          @20 dept=;
      return;
run;
```

Example 2: Determining New Page by Lines Left on the Current Page This DATA step demonstrates using the LINESLEFT= option to determine where the page break should occur, according to the number of lines left on the current page.

- *Write a report.* Use DATA _NULL_ to write a report rather than create a data set:

```
data _null_;
   set info;
```

- *Route output to the standard SAS output window.* The PRINT fileref routes output to the same location as procedure output. LINESLEFT indicates that the variable REMAIN contains the number of lines left on the current page:

```
file print linesleft=remain pagesize=20;
   put @5 name @30 phone
       @35 bldg @37 room;
```

- *Begin a new page when there are fewer than seven lines left on the current page.* Under this condition, PUT _PAGE_ begins a new page and positions the pointer at line 1:

```
if remain<7 then put _page_ ;
run;
```

Example 3: Arranging the Contents of an Entire Page This example shows how to use N=PAGESIZE in a DATA step to produce a two-column telephone book listing, each column containing a name and a phone number:

- *Create a report and write it to a SAS output window.* Use DATA _NULL_ to write a report rather than create a data set. PRINT is the fileref. SAS uses carriage-control characters to write the output with the characteristics of a print file. N=PAGESIZE makes the entire page available to the output pointer:

```
data _null_;
   file 'external-file' print n=pagesize;
```

- *Specify the columns for the report.* This DO loop iterates twice on each DATA step iteration. The COL value is 1 on the first iteration and 40 on the second:

```
do col=1, 40;
```

- *Write 20 lines of data.* This DO loop iterates 20 times to write 20 lines in column 1. When finished, the outer loop sets COL equal to 40, and this DO loop iterates 20 times again, writing 20 lines of data in the second column. The values of LINE and COL, which are set and incremented by the DO statements, control where the PUT statement writes the values of NAME and PHONE on the page:

```
do line=1 to 20;
   set info;
   put #line @col name $20. +1 phone 4.;
end;
```

- *After composing two columns of data, write the page.* This END statement ends the outer DO loop. The PUT _PAGE_ writes the current page and moves the pointer to the top of a new page:

```
end;
   put _page_;
run;
```

Example 4: Identifying the Current Output File This DATA step causes a file identification message to print in the log and assigns the value of the current output file to the variable MYOUT. The PUT statement, demonstrating the assignment of the proper value to MYOUT, writes the value of that variable to the output file:

```
data _null_;
   length myout $ 200;
   file file-specification filename=myout;
   put myout=;
   stop;
run;
```

The PUT statement writes a line to the current output file that contains the physical name of the file:

```
MYOUT=your-output-file
```

Example 5: Dynamically Changing the Current Output File This DATA step uses the FILEVAR= option to dynamically change the currently opened output file to a new physical file.

- *Write a report. Create a long character variable.* Use DATA _NULL_ to write a report rather than create a data set. The LENGTH statement creates a variable with length long enough to contain the name of an external file:

```
data _null_;
   length name $ 200;
```

- *Read an in-stream data line and assign a value to the NAME variable:*

```
   input name $;
```

- *Close the current output file and open a new one when the NAME variable changes.* The *file-specification* is just a place holder; it can be any valid SAS name:

```
   file file-specification filevar=name mod;
      date = date();
```

- *Append a log record to currently open output file:*

```
   put 'records updated ' date date.;
```

- *Supply the names of the external files:*

```
datalines;
external-file-1
external-file-2
external-file-3
;
```

Example 6: When the Output Line Exceeds the Line Length of the Output File Because the combined lengths of the variables are longer than the output line (80 characters), this PUT statement automatically writes three separate records:

```
file file-specification linesize=80;
   put name $ 1-50 city $ 71-90 state $ 91-104;
```

The value of NAME appears in the first record, CITY begins in the first column of the second record, and STATE in the first column of the third record.

Example 7: Reading Data and Writing Text through a TCP/IP Socket This example shows reading raw data from a file through a TCP/IP socket. The NBYTE= option is used in the INFILE statement:

```
/* Start this first as the server */

   filename serve socket ':5205' server
            recfm=s
            lrecl=25 blocksize=2500;

data _null_;
   nb=25;
   infile serve nbyte=nb;
   input text $char25.;
   put _all_;
run;
```

This example shows writing text to a file through a TCP/IP socket:

```
/* While the server test is running,*/
/*continue with this as the client. */
```

```
filename client socket "&hstname:5205"
         recfm=s
         lrecl=25 blocksize=2500;

data _null_;
   file client;
   put 'Some text to length 25...';
run;
```

Example 8: Specifying an Encoding When Writing to an Output File

This example creates an external file from a SAS data set. The current session encoding is Wlatin1, but the external file's encoding needs to be UTF-8. By default, SAS writes the external file using the current session encoding.

To tell SAS what encoding to use when writing data to the external file, specify the ENCODING= option. When you tell SAS that the external file is to be in UTF-8 encoding, SAS then transcodes the data from Wlatin1 to the specified UTF-8 encoding when writing to the external file.

```
libname myfiles 'SAS-library';

filename outfile 'external-file';

data _null_;
   set myfiles.cars;
   file outfile encoding="utf-8";
   put Make Model Year;
run;
```

See Also

Statements:

"FILE Statement for ODS" in *SAS Output Delivery System: User's Guide*
"FILENAME Statement" on page 1520
"INFILE Statement" on page 1591
"LABEL Statement" on page 1650
"PUT Statement" on page 1708
"RETURN Statement" on page 1752
"TITLE Statement" on page 1779

FILENAME Statement

Associates a SAS fileref with an external file or an output device, disassociates a fileref and external file, or lists attributes of external files.

Valid: anywhere

Category: Data Access

See: FILENAME Statement in the documentation for your operating environment

Syntax

❶ **FILENAME** *fileref* <*device-type*> '*external-file*' <ENCODING='*encoding-value*'> <*options*><*operating-environment-options*>;

❷ **FILENAME** *fileref* <*device-type*><*options*> <*operating-environment-options*>;

❸ **FILENAME** *fileref* CLEAR | _ALL_ CLEAR;

❹ **FILENAME** *fileref* LIST | _ALL_ LIST;

Arguments

fileref
is any SAS name that you use when you assign a new fileref. When you disassociate a currently assigned fileref or when you list file attributes with the FILENAME statement, specify a fileref that was previously assigned with a FILENAME statement or an operating environment-level command.

Tip: The association between a fileref and an external file lasts only for the duration of the SAS session or until you change it or discontinue it by using another FILENAME statement. Change the fileref for a file as often as you want.

'*external-file*'
is the physical name of an external file. The physical name is the name that is recognized by the operating environment.

Operating Environment Information: For details about specifying the physical names of external files, see the SAS documentation for your operating environment. △

Tip: Specify *external-file* when you assign a fileref to an external file.

Tip: You can associate a fileref with a single file or with an aggregate file storage location.

ENCODING= '*encoding-value*'
specifies the encoding to use when SAS is reading from or writing to an external file. The value for ENCODING= indicates that the external file has a different encoding from the current session encoding.

When you read data from an external file, SAS transcodes the data from the specified encoding to the session encoding. When you write data to an external file, SAS transcodes the data from the session encoding to the specified encoding.

For valid encoding values, see "Encoding Values in SAS Language Elements" in *SAS National Language Support (NLS): Reference Guide*.

Default: SAS assumes that an external file is in the same encoding as the session encoding.

Featured in: Example 5 on page 1525 and Example 6 on page 1525

device-type
specifies the type of device or the access method that is used if the fileref points to an input or output device or location that is not a physical file:

DISK
specifies that the device is a disk drive.

Tip: When you assign a fileref to a file on disk, you are not required to specify DISK.

DUMMY
specifies that the output to the file is discarded.

Tip: Specifying DUMMY can be useful for testing.

GTERM
indicates that the output device type is a graphics device that will receive graphics data.

PIPE
specifies an unnamed pipe.

Note: Some operating environments do not support pipes. △

PLOTTER
specifies an unbuffered graphics output device.

PRINTER
specifies a printer or printer spool file.

TAPE
specifies a tape drive.

TEMP
creates a temporary file that exists only as long as the filename is assigned. The temporary file can be accessed only through the logical name and is available only while the logical name exists.

Restriction: Do not specify a physical pathname. If you do, SAS returns an error.

Tip: Files manipulated by the TEMP device can have the same attributes and behave identically to DISK files.

TERMINAL
specifies the user's terminal.

UPRINTER
specifies a Universal Printing printer definition name.

Tip: If you do not specify the printer name in the FILENAME statement, the PRINTERPATH options control which Universal Printer is used and the destination of the output.

Operating Environment Information: Additional specifications might be required when you specify some devices. See the SAS documentation for your operating environment before specifying a value other than DISK. Values in addition to the ones listed here might be available in some operating environments. △

CLEAR
disassociates one or more currently assigned filerefs.

Tip: Specify *fileref* to disassociate a single fileref. Specify _ALL_ to disassociate all currently assigned filerefs.

ALL
specifies that the CLEAR or LIST argument applies to all currently assigned filerefs.

LIST
writes the attributes of one or more files to the SAS log.

Interaction: Specify *fileref* to list the attributes of a single file. Specify _ALL_ to list the attributes of all files that have filerefs in your current session.

Options

RECFM=*record-format*
specifies the record format of the external file.

Operating Environment Information: Values for *record-format* are dependent on the operating environment. For details, see the SAS documentation for your operating environment. △

Operating Environment Options

Operating environment options specify details, such as file attributes and processing attributes, that are specific to your operating environment.

Operating Environment Information: For a list of valid specifications, see the SAS documentation for your operating environment. △

Details

Operating Environment Information

Operating Environment Information: Using the FILENAME statement requires operating environment-specific information. See the SAS documentation for your operating environment before using this statement. Note also that commands are available in some operating environments that associate a fileref with a file and that break that association. △

Definitions

external file
is a file that is created and maintained in the operating environment from which you need to read data, SAS programming statements, or autocall macros, or to which you want to write output. An external file can be a single file or an aggregate storage location that contains many individual external files. See Example 3 on page 1524.

Operating Environment Information: Different operating environments call an aggregate grouping of files by different names, such as a directory, a MACLIB, or a partitioned data set. For details about specifying external files, see the SAS documentation for your operating environment. △

fileref
(a file reference name) is a shorthand reference to an external file. After you associate a fileref with an external file, you can use it as a shorthand reference for that file in SAS programming statements (such as INFILE, FILE, and %INCLUDE) and in other commands and statements in SAS software that access external files.

Reading Delimited Data from an External File Any time a text file originates from anywhere other than the local encoding environment, it might be necessary to specify the ENCODING= option in either EBCDIC or ASCII environments.

For example, when you read an EBCDIC text file on an ASCII platform, it is recommended that you specify the ENCODING= option in the FILENAME statement. However, if you use the DSD and DLM options in the FILENAME statement, the

ENCODING= option is a requirement because these options require certain characters in the session encoding (such as quotation marks, commas, and blanks).

The use of encoding-specific informats should be reserved for use with true binary files. That is, they contain both character and non-character fields.

❶Associating a Fileref with an External File Use this form of the FILENAME statement to associate a fileref with an external file on disk:

FILENAME *fileref 'external-file'* <*operating-environment-options*>;

To associate a fileref with a file other than a disk file, you might need to specify a device type, depending on your operating environment, as shown in this form:

FILENAME *fileref* <*device-type*> <*operating-environment-options*>;

The association between a fileref and an external file lasts only for the duration of the SAS session or until you change it or discontinue it with another FILENAME statement. Change the fileref for a file as often as you want.

To specify a character-set encoding, use the following form:

FILENAME *fileref* <*device-type*> <*operating-environment-options*>;

❷Associating a Fileref with a Terminal, Printer, Universal Printer, or Plotter To associate a fileref with an output device, use this form:

FILENAME *fileref device-type* <*operating-environment-options*>;

❸Disassociating a Fileref from an External File To disassociate a fileref from a file, use a FILENAME statement, specifying the fileref and the CLEAR option.

❹Writing File Attributes to the SAS Log Use a FILENAME statement to write the attributes of one or more external files to the SAS log. Specify *fileref* to list the attributes of one file; use _ALL_ to list the attributes of all the files that have been assigned filerefs in your current SAS session.

FILENAME *fileref* LIST | _ALL_ LIST;

Comparisons

The FILENAME statement assigns a fileref to an external file. The LIBNAME statement assigns a libref to a SAS data set or to a DBMS file that can be accessed like a SAS data set.

Examples

Example 1: Specifying a Fileref or a Physical Filename You can specify an external file either by associating a fileref with the file and then specifying the fileref or by specifying the physical filename in quotation marks:

```
filename sales 'your-input-file';

data jansales;
      /* specifying a fileref */
   infile sales;
   input salesrep $20. +6 jansales febsales
         marsales;
run;
```

```
data jansales;
      /* physical filename in quotation marks */
   infile 'your-input-file';
   input salesrep $20. +6 jansales febsales
         marsales;
run;
```

Example 2: Using a FILENAME and a LIBNAME Statement

This example reads data from a file that has been associated with the fileref GREEN and creates a permanent SAS data set stored in a SAS library that has been associated with the libref SAVE.

```
filename green 'your-input-file';
libname save 'SAS-library';

data save.vegetable;
   infile green;
   input lettuce cabbage broccoli;
run;
```

Example 3: Associating a Fileref with an Aggregate Storage Location If you associate a fileref with an aggregate storage location, use the fileref, followed in parentheses by an individual filename, to read from or write to any of the individual external files that are stored there.

Operating Environment Information: Some operating environments allow you to read from but not write to members of aggregate storage locations. For details, see the SAS documentation for your operating environment. △

In this example, each DATA step reads from an external file (REGION1 and REGION2, respectively) that is stored in the same aggregate storage location and that is referenced by the fileref SALES.

```
filename sales 'aggregate-storage-location';

data total1;
   infile sales(region1);
   input machine $ jansales febsales marsales;
   totsale=jansales+febsales+marsales;
run;

data total2;
   infile sales(region2);
   input machine $ jansales febsales marsales;
   totsale=jansales+febsales+marsales;
run;
```

Example 4: Routing PUT Statement Output In this example, the FILENAME statement associates the fileref OUT with a printer that is specified with an operating environment-dependent option. The FILE statement directs PUT statement output to that printer.

```
filename out printer operating-environment-option;

data sales;
```

```
   file out print;
   input salesrep $20. +6 jansales
         febsales marsales;
   put _infile_;
   datalines;
Jones, E. A.              124357 155321 167895
Lee, C. R.                111245 127564 143255
Desmond, R. T.             97631 101345 117865
;
```

You can use the FILENAME and FILE statements to route PUT statement output to several devices during the same session. To route PUT statement output to your display monitor, use the TERMINAL option in the FILENAME statement, as shown here:

```
filename show terminal;

data sales;
   file show;
   input salesrep $20. +6 jansales
         febsales marsales;
   put _infile_;
   datalines;
Jones, E. A.              124357 155321 167895
Lee, C. R.                111245 127564 143255
Desmond, R. T.             97631 101345 117865
;
```

Example 5: Specifying an Encoding When Reading an External File This example creates a SAS data set from an external file. The external file is in UTF-8 character-set encoding, and the current SAS session is in the Wlatin1 encoding. By default, SAS assumes that an external file is in the same encoding as the session encoding, which causes the character data to be written to the new SAS data set incorrectly.

To tell SAS what encoding to use when reading the external file, specify the ENCODING= option. When you tell SAS that the external file is in UTF-8, SAS then transcodes the external file from UTF-8 to the current session encoding when writing to the new SAS data set. Therefore, the data is written to the new data set correctly in Wlatin1.

```
libname myfiles  'SAS-library';

filename extfile 'external-file' encoding="utf-8";

data myfiles.unicode;
   infile extfile;
   input Make $ Model $ Year;
run;
```

Example 6: Specifying an Encoding When Writing to an External File This example creates an external file from a SAS data set. The current session encoding is Wlatin1, but the external file's encoding needs to be UTF-8. By default, SAS writes the external file using the current session encoding.

To tell SAS what encoding to use when writing data to the external file, specify the ENCODING= option. When you tell SAS that the external file is to be in UTF-8 encoding, SAS then transcodes the data from Wlatin1 to the specified UTF-8 encoding when writing to the external file.

```
libname myfiles  'SAS-library';

filename outfile 'external-file' encoding="utf-8";

data _null_;
   set myfiles.cars;
   file outfile;
   put Make Model Year;
run;
```

See Also

Statements:

- "FILE Statement" on page 1503
- "%INCLUDE Statement" on page 1584
- "INFILE Statement" on page 1591
- "FILENAME Statement, CATALOG Access Method" on page 1526
- "FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532
- "FILENAME Statement, FTP Access Method" on page 1542
- "FILENAME Statement, SOCKET Access Method" on page 1559
- "FILENAME Statement, SFTP Access Method" on page 1554
- "FILENAME Statement, URL Access Method" on page 1563
- "LIBNAME Statement" on page 1656

SAS Windowing Interface Commands:

- FILE and INCLUDE

FILENAME Statement, CATALOG Access Method

Enables you to reference a SAS catalog as an external file.

Valid: anywhere

Category: Data Access

Syntax

FILENAME *fileref* **CATALOG** '*catalog*' <*catalog-options*>;

Arguments

fileref
is a valid fileref.

CATALOG
specifies the access method that enables you to reference a SAS catalog as an external file. You can then use any SAS commands, statements, or procedures that can access external files to access a SAS catalog.

Tip: This access method makes it possible for you to invoke an autocall macro directly from a SAS catalog.

Tip: With this access method you can read any type of catalog entry, but you can write only to entries of type LOG, OUTPUT, SOURCE, and CATAMS.

Tip: If you want to access an entire catalog (instead of a single entry), you must specify its two-level name in the *catalog* parameter.

Alias: LIBRARY

'*catalog*'
is a valid two-, three-, or four-part SAS catalog name, where the parts represent *library.catalog.entry.entrytype*.

Default: The default entry type is CATAMS.

Restriction: The CATAMS entry type is used only by the CATALOG access method. The CPORT and CIMPORT procedures do not support this entry type.

Catalog Options

Catalog-options can be any of the following:

LRECL=*lrecl*
where *lrecl* is the maximum record length for the data in bytes.

Default: For input, the actual LRECL value of the file is the default. For output, the default is 132.

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

RECFM=*recfm*
where *recfm* is one of four record formats:

F — is fixed-record format. Data is transferred in image (binary) mode.

P — is print format.

S — is stream-record format. Data is transferred in image (binary) mode.

Interaction: The amount of data that is read is controlled by the value of the NBYTE= variable in the INFILE statement. The NBYTE= option specifies a variable that is equal to the amount of data to be read. This amount must be less than or equal to LRECL.

See Also: The NBYTE= option on page 1598 in the INFILE statement.

V — is variable-record format (the default). In this format, records have varying lengths, and they are separated by newlines. Data is transferred in image (binary) mode.

Default: V

DESC=*description*
where *description* is a text description of the catalog.

MOD
specifies to append to the file.

Default: If you omit MOD, the file is replaced.

Details

The CATALOG access method in the FILENAME statement enables you to reference a SAS catalog as an external file. You can then use any SAS commands, statements, or procedures that can access external files to access a SAS catalog. As an example, the catalog access method makes it possible for you to invoke an autocall macro directly from a SAS catalog. See Example 5 on page 1529.

With the CATALOG access method you can read any type of catalog entry, but you can write to only entries of type LOG, OUTPUT, SOURCE, and CATAMS. If you want to access an entire catalog (instead of a single entry), you must specify its two-level name in the *catalog* argument.

Examples

Example 1: Using %INCLUDE with a Catalog Entry This example submits the source program that is contained in SASUSER.PROFILE.SASINP.SOURCE:

```
filename fileref1
         catalog 'sasuser.profile.sasinp.source';
%include fileref1;
```

Example 2: Using %INCLUDE with Several Entries in a Single Catalog This example submits the source code from three entries in the catalog MYLIB.INCLUDE. When no entry type is specified, the default is CATAMS.

```
filename dir catalog 'mylib.include';
%include dir(mem1);
%include dir(mem2);
%include dir(mem3);
```

Example 3: Reading and Writing a CATAMS Entry This example uses a DATA step to write data to a CATAMS entry, and another DATA step to read it back in:

```
filename mydata
         catalog 'sasuser.data.update.catams';

   /* write data to catalog entry update.catams */
data _null_;
   file mydata;
   do i=1 to 10;
      put i;
   end;
run;

   /* read data from catalog entry update.catams */
data _null_;
   infile mydata;
   input;
   put _INFILE_;
run;
```

Example 4: Writing to a SOURCE Entry This example writes code to a catalog SOURCE entry and then submits it for processing:

```
filename incit
         catalog 'sasuser.profile.sasinp.source';
```

```
data _null_;
   file incit;
   put 'proc options; run;';
run;

%include incit;
```

Example 5: Executing an Autocall Macro from a SAS Catalog If you store an autocall macro in a SOURCE entry in a SAS catalog, you can point to that entry and invoke the macro in a SAS job. Use these steps:

1. Store the source code for the macro in a SOURCE entry in a SAS catalog. The name of the entry is the macro name.
2. Use a LIBNAME statement to assign a libref to that SAS library.
3. Use a FILENAME statement with the CATALOG specification to assign a fileref to the catalog: *libref.catalog*.
4. Use the SASAUTOS= option and specify the fileref so that the system knows where to locate the macro. Also set MAUTOSOURCE to activate the autocall facility.
5. Invoke the macro as usual: *%macro-name*.

This example points to a SAS catalog named MYSAS.MYCAT. It then invokes a macro named REPORTS, which is stored as a SAS catalog entry named MYSAS.MYCAT.REPORTS.SOURCE:

```
libname mysas 'SAS-library';
filename mymacros catalog 'mysas.mycat';
options sasautos=mymacros mautosource;

%reports
```

See Also

Statements:

"FILENAME Statement" on page 1520
"FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532
"FILENAME Statement, FTP Access Method" on page 1542
"FILENAME Statement, SOCKET Access Method" on page 1559
"FILENAME Statement, SFTP Access Method" on page 1554
"FILENAME Statement, URL Access Method" on page 1563

FILENAME, CLIPBOARD Access Method

Enables you to read text data from and write text data to the clipboard on the host computer.

Valid: anywhere

Category: Data Access

Syntax

FILENAME *fileref* CLIPBRD <BUFFER=*paste-buffer-name*>;

Arguments

fileref
is a valid fileref.

CLIPBRD
specifies the access method that enables you to read data from or write data to the clipboard on the host computer.

BUFFER=*paste-buffer-name*
creates and names the paste buffer. You can create any number of paste buffers by naming them with the BUFFER= argument in the STORE command.

Details

The FILENAME statement, CLIPBOARD Access Method enables you to share data within SAS and between SAS and applications other than SAS.

Comparisons

The STORE command copies marked text in the current window and stores the copy in a paste buffer.

You can also copy data to the clipboard by using the Explorer pop-up menu item **Copy Contents to Clipboard**.

Examples

Example 1: Using ODS to Write a Data Set as HTML to the Clipboard This example uses the Sashelp.Air data set as the input file. The ODS is used to write the data set in HTML format to the clipboard.

```
filename _temp_ clipbrd;
   ods noresults;
   ods listing close;
   ods html file=_temp_ rs=none style=minimal;
   proc print data=Sashelp.'Air'N noobs;
run;
ods html close;
ods results;
ods listing;
filename _temp_;
```

Example 2: Using the DATA Step to Write a Data Set As Comma-separated Values to the Clipboard This example uses the Sashelp.Air data set as the input file. The data is written in the DATA step as comma-separated values to the clipboard.

```
filename _temp1_ temp;
filename _temp2_ clipbrd;
proc contents data=Sashelp."Air"N out=info noprint;
proc sort data=info;
   by npos;
```

```
run;

data _null_;
   set info end=eof;
   ;
   file _temp1_ dsd;
   put name @@;
   if _n_=1 then do;
      call execute("data _null_; set Sashelp.""Air""N; file _temp1_ dsd mod; put");
   end;
   call execute(trim(name));
   if eof then call execute('; run;');
run;

data _null_;
   infile _temp1_;
   file _temp2_;
   input;
   put _infile_;
run;

filename _temp1_ clear;
filename _temp2_ clear;
```

Example 3: Using the DATA Step to Write Text to the Clipboard This example writes three lines to the clipboard.

```
filename clippy clipbrd;

data _null_;
   file clippy;
   put 'Line 1';
   put 'Line 2';
   put 'Line 3';
run;
```

Example 4: Using the DATA Step to Retrieve Text from the Clipboard This example writes three lines to the clipboard and then retrieves them.

```
filename clippy clipbrd;

data _null_;
   file clippy;
   put 'Line 1';
   put 'Line 2';
   put 'Line 3';
run;

data _null_;
   infile clippy;
   input;
   put _infile_;
run;
```

See Also

Command:

The STORE command in the Base SAS Help and Documentation.

FILENAME Statement, EMAIL (SMTP) Access Method

Enables you to send electronic mail programmatically from SAS using the SMTP (Simple Mail Transfer Protocol) e-mail interface.

Valid: Anywhere

Category: Data Access

Syntax

FILENAME *fileref* EMAIL <'*address*' ><*email-options*>;

Arguments

fileref
is a valid file reference. The fileref is a name that is temporarily assigned to an external file or to a device type. Note that the fileref cannot exceed eight characters.

EMAIL
specifies the EMAIL device type, which provides the access method that enables you to send electronic mail programmatically from SAS. In order to use SAS to send a message to an SMTP server, you must enable SMTP e-mail. For more information, see "The SMTP E-Mail Interface" in *SAS Language Reference: Concepts*.

'*address*'
is the e-mail address to which you want to send the message. You must enclose the address in quotation marks. Specifying an address as a FILENAME statement argument is optional if you specify the TO= e-mail option or the PUT statement !EM_TO! directive, which will override an *address* specification.

E-mail Options

You can use any of the following e–mail options in the FILENAME statement to specify attributes for the electronic message.

Note: You can also specify these options in the FILE statement. E-mail options that you specify in the FILE statement override any corresponding e-mail options that you specified in the FILENAME statement. △

ATTACH='*filename.ext*' | ATTACH= ('*filename.ext*' *attachment-options*)
specifies the physical name of the file or files to be attached to the message and any options to modify attachment specifications. The physical name is the name that is recognized by the operating environment. Enclose the physical name in quotation marks. To attach more than one file, enclose the group of files in

parentheses, enclose each file in quotation marks, and separate each with a space. Here are examples:

```
attach="/u/userid/opinion.txt"

attach('C:\Status\June2001.txt' 'C:\Status\July2001.txt')

attach="user.misc.pds(member)"
```

The *attachment-options* include the following:

CONTENT_TYPE='*content/type*'
specifies the content type for the attached file. You must enclose the value in quotation marks. If you do not specify a content type, SAS tries to determine the correct content type based on the filename. For example, if you do not specify a content type, a filename of **home.html** is sent with a content type of text/html.

Aliases: CT= and TYPE=

Default: If SAS cannot determine a content type based on the filename and extension, the default value is text/plain.

ENCODING='*encoding-value*'
specifies the text encoding of the attachment that is read into SAS. You must enclose the value in quotation marks.

See Also: "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*

EXTENSION='*extension*'
specifies a different file extension to be used for the specified attachment. You must enclose the value in quotation marks. This extension is used by the recipient's e-mail program for selecting the appropriate utility to use for displaying the attachment. For example, the following results in the attachment **home.html** being received as **index.htm**:

```
attach=("home.html" name="index" ext="htm")
```

Note: If you specify *extension*="", the specified attachment will have no file extension. △

Alias: EXT=

NAME='*filename*'
specifies a different name to be used for the specified attachment. You must enclose the value in quotation marks. For example, the following results in the attachment **home.html** being received as **index.html**:

```
attach=("home.html" name="index")
```

OUTENCODING='*encoding-value*'
specifies the resulting text encoding for the attachment to be sent. You must enclose the value in quotation marks.

Restriction: Do not specify EBCDIC encoding values, because the SMTP e-mail interface does not support EBCDIC.

See Also: "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*

BCC='*bcc-address*'
specifies the recipient or recipients that you want to receive a blind copy of the electronic mail. Individuals that are listed in the **bcc** field will receive a copy of

the e-mail. The BCC field does not appear in the e-mail header, so that these e-mail addresses cannot be viewed by other recipients.

If a BCC address contains more than one word, then enclose it in quotation marks. To specify more than one address, you must enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify a real name as well as an address, enclose the address in angle brackets (< >). Here are examples:

```
bcc="joe@site.com"

bcc=("joe@site.com" "jane@home.net")

bcc="Joe Smith <joe@site.com>"
```

CC='*cc-address*'
specifies the recipient or recipients to receive a copy of the e-mail message. You must enclose an address in quotation marks. To specify more than one address, enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify a real name as well as an address, enclose the address in angle brackets (< >). Here are examples:

```
cc='joe@site.com'

cc=("joe@site.com" "jane@home.net")

cc="Joe Smith <joe@site.com>"
```

CONTENT_TYPE='*content/type*'
specifies the content type for the message body. If you do not specify a content type, SAS tries to determine the correct content type. You must enclose the value in quotation marks.

Aliases: CT= and TYPE=

Default: text/plain

ENCODING='*encoding-value*'
specifies the text encoding to use for the message body. For valid encoding values, see "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*.

FROM='*from-address*'
specifies the e-mail address of the author of the message that is being sent. The default value for FROM= is the e-mail address of the user who is running SAS. For example, specify this option when the person who is sending the message from SAS is not the author. You must enclose an address in quotation marks. You can specify only one e-mail address. To specify the author's real name along with the address, enclose the address in angle brackets (< >). Here are examples:

```
from='martin@home.com'

from="Brad Martin <martin@home.com>"
```

Requirement: The FROM option is required if the EMAILFROM system option is set.

LRECL=*lrecl*
where *lrecl* is the logical record length of the data.

Default: 256

Interaction: Alternatively, you can specify a global logical record length by using the "LRECL= System Option" on page 1941.

IMPORTANCE='LOW' | 'NORMAL' | 'HIGH'
specifies the priority of the e-mail message. You must enclose the value in quotation marks. You can specify the priority in the language that matches your session encoding. However, SAS will translate the priority into English because the actual message header must contain English in accordance with the RFC-2076 specification (Common Internet Message Headers). Here are examples:

```
filename inventory email 'name@mycompany.com' importance='high';

filename inventory email 'name@mycompany.com' importance='hoch';
```

Default: NORMAL

REPLYTO='*replyto-address*'
specifies the e-mail address(es) for who will receive replies. You must enclose an address in quotation marks. To specify more than one address, enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify a real name along with an address, enclose the address in angle brackets (< >). Here are examples:

```
replyto='hiroshi@home.com'

replyto=('hiroshi@home.com' 'akiko@site.com')

replyto="Hiroshi Mori <mori@site.com>"
```

SUBJECT=*subject*
specifies the subject of the message. If the subject contains special characters or more than one word (that is, it contains at least one blank space), you must enclose the text in quotation marks. Here are examples:

```
subject=Sales

subject="June Sales Report"
```

Note: If you do not enclose a one-word subject in quotation marks, it is converted to uppercase. △

TO='*to-address*'
specifies the primary recipient or recipients of the e-mail message. You must enclose the address in quotation marks. To specify more than one address, enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify a real name as well as an address, enclose the address in angle brackets (< >). Here are examples:

```
to='joe@site.com'

to=("joe@site.com" "jane@home.net")

to="Joe Smith <joe@site.com>"
```

Tip: Specifying TO= overrides the '*address*' argument.

PUT Statement Syntax for EMAIL (SMTP) Access Method

In the DATA step, after using the FILE statement to define your e-mail fileref as the output destination, use PUT statements to define the body of the message. For example,

```
filename mymail email 'martin@site.com' subject='Sending Email';
```

```
data _null_;
   file mymail;
   put 'Hi';
   put 'This message is sent from SAS...';
run;
```

You can also use PUT statements to specify e-mail directives that override the attributes of your message (the e-mail options like TO=, CC=, SUBJECT=, CONTENT_TYPE=, ATTACH=), or to perform actions such as send, end abnormally, or start a new message. Specify only one directive in each PUT statement; each PUT statement can contain only the text that is associated with the directive that it specifies.

The directives that change the attributes of a message are as follows:

'!EM_ATTACH! '*filename.ext*' | ATTACH=('*filename.ext*' *attachment-options*)'
replaces the physical name of the file or files to be attached to the message and any options to modify attachment specifications. The physical name is the name that is recognized by the operating environment. The directive must be enclosed in quotation marks, and the physical name must be enclosed in quotation marks. To attach more than one file, enclose the group of files in parentheses, enclose each file in quotation marks, and separate each with a space. Here are examples:

```
put '!em_attach! /u/userid/opinion.txt';

put '!em_attach! ("C:\Status\June2001.txt" "C:\Status\July2001.txt")';

put '!em_attach! user.misc.pds(member)';
```

The *attachment-options* include the following:

CONTENT_TYPE='*content/type*'
specifies the content type for the attached file. You must enclose the value in quotation marks. If you do not specify a content type, SAS tries to determine the correct content type based on the filename. For example, if you do not specify a content type, a filename of **home.html** is sent with a content type of text/html.

Aliases: CT= and TYPE=

Default: If SAS cannot determine a content type based on the filename and extension, the default value is text/plain.

ENCODING='*encoding-value*'
specifies the text encoding to use for the attachment as it is read into SAS. You must enclose the value in quotation marks. For valid encoding values, see "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*.

EXTENSION='*extension*'
specifies a different file extension to be used for the specified attachment. You must enclose the value in quotation marks. This extension is used by the recipient's e-mail program for selecting the appropriate utility to use for displaying the attachment. For example, the following results in the attachment **home.html** being received as **index.htm**:

```
put '!em_attach! ("home.html" name="index" ext="htm")';
```

Alias: EXT=

Default: TXT

NAME='*filename*'
specifies a different name to be used for the specified attachment. You must enclose the value in quotation marks. For example, the following results in the attachment **home.html** being received as **index.html**:

```
put '!em_attach! ("home.html" name="index")';
```

OUTENCODING='*encoding-value*'
specifies the resulting text encoding for the attachment to be sent. You must enclose the value in quotation marks.

Restriction: Do not specify EBCDIC encoding values, because the SMTP e-mail interface does not support EBCDIC.

See Also: "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*

'!EM_BCC! *bcc-address*'
replaces the current blind copied recipient address(es) with *addresses*. These recipients are not visible to the recipients in the !EM_TO! or !EM_CC! addresses. If you want to specify more than one address, then you must enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify real names along with addresses, enclose the address in angle brackets (< >). Here are examples:

```
put '!em_bcc! joe@site.com';

put '!em_bcc! ("joe@site.com" "jane@home.net")';

put '!em_bcc! Joe Smith <joe@site.com>';
```

'!EM_CC! *cc-address*'
replaces the current copied recipient address(es). The directive must be enclosed in quotation marks. To specify more than one address, enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify real names along with addresses, enclose the address in angle brackets (< >). Here are examples:

```
put '!em_cc! joe@site.com';

put '!em_cc! ("joe@site.com" "jane@home.com")';

put '!em_cc! Joe Smith <joe@site.com>';
```

'!EM_FROM! *from-address*'
replaces the current address of the author of the message being sent, which could be either the default or the one specified by the FROM= e-mail option. The directive must be enclosed in quotation marks. You can specify only one e-mail address. To specify the author's real name along with the address, enclose the address in angle brackets (< >). Here are examples:

```
put '!em_from! martin@home.com';

put '!em_from! Brad Martin <martin@home.com>';
```

'!EM_IMPORTANCE! LOW | NORMAL | HIGH'
specifies the priority of the e-mail message. The directive must be enclosed in quotation marks. You can specify the priority in the language that matches your session encoding. However, SAS will translate the priority into English because the actual message header must contain English in accordance with the RFC-2076 specification (Common Internet Message Headers). Here are examples:

```
put '!em_importance! high';

put '!em_importance! haut';
```

Default: NORMAL

'!EM_REPLYTO! *replyto-address*'
replaces the current address(es) of who will receive replies. The directive must be enclosed in quotation marks. To specify more than one address, enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify a real name along with an address, enclose the address in angle brackets (< >). Here are examples:

```
put '!em_replyto! hiroshi@home.com';

put '!em_replyto! ("hiroshi@home.com" "akiko@site.com")';

put '!em_replyto! Hiroshi Mori <mori@site.com>';
```

'!EM_SUBJECT! *subject*'
replaces the current subject of the message. The directive must be enclosed in quotation marks. If the subject contains special characters or more than one word (that is, it contains at least one blank space), you must enclose the text in quotation marks. Here are examples:

```
put '!em_subject! Sales';

put '!em_subject! "June Sales Report"';
```

'!EM_TO! *to-address*'
replaces the current primary recipient address(es). The directive must be enclosed in quotation marks. To specify more than one address, enclose the group of addresses in parentheses, enclose each address in quotation marks, and separate each address with a space. To specify a real name along with an address, enclose the address in angle brackets (< >). Here are examples:

```
put '!em_to! joe@site.com';

put '!em_to! ("joe@site.com" "jane@home.net")';

put '!em_to! Joe Smith <joe@site.com>';
```

Tip: Specifying !EM_TO! overrides the '*address*' argument and the TO= e-mail option.

Here are the directives that perform actions:

'!EM_SEND!'
sends the message with the current attributes. By default, SAS sends a message when the fileref is closed. The fileref closes when the next FILE statement is encountered or the DATA step ends. If you use this directive, SAS sends the message when it encounters the directive, and again at the end of the DATA step.

This directive is useful for writing DATA step programs that conditionally send messages or use a loop to send multiple messages.

'!EM_ABORT!'
abnormally end the current message. You can use this directive to stop SAS from automatically sending the message at the end of the DATA step. By default, SAS sends a message for each FILE statement.

'!EM_NEWMSG!'
clears all attributes of the current message that were set using PUT statement directives.

Details

You can send electronic mail programmatically from SAS using the EMAIL (SMTP) access method. To send e-mail to an SMTP server, you first specify the SMTP e-mail interface with the EMAILSYS system option, use the FILENAME statement to specify the EMAIL device type, and then submit SAS statements in a DATA step or in SCL code. The e-mail access method has several advantages:

- You can use the logic of the DATA step or SCL to subset e-mail distribution based on a large data set of e-mail addresses.
- You can automatically send e-mail upon completion of a SAS program that you submitted for batch processing.
- You can direct output through e-mail based on the results of processing.

In general, DATA step or SCL code that sends e-mail has the following components:

- a FILENAME statement with the EMAIL device-type keyword
- e-mail options specified in the FILENAME or FILE statement that indicate e-mail recipients, subject, attached file or files, and so on
- PUT statements that define the body of the message
- PUT statements that specify e-mail directives (of the form !EM_*directive*!) that override the e-mail options (for example, TO=, CC=, SUBJECT=, ATTACH=) or perform actions such as send, end abnormally, or start a new message.

You can use encoded e-mail passwords. When a password is encoded with PROC PWENCODE, the output string includes a tag that identifies the string as having been encoded. An example of a tag is {sas001}. The tag indicates the encoding method. Encoding a password enables you to avoid e-mail access authentication with a password in plaintext. Passwords that start with "{sas" trigger an attempt to be decoded. If the decoding succeeds, then that decoded password is used. If the decoding fails, then the password is used as is. For more information, see PROC PWENCODE in the *Base SAS Procedures Guide*.

Examples

Example 1: Sending E-mail with an Attachment Using a DATA Step In order to share a copy of your SAS configuration file with another user, you could send it by submitting the following program. The e-mail options are specified in the FILENAME statement:

```
filename mymail email "JBrown@site.com"
   subject="My SAS Configuration File"
   attach="/u/sas/sasv8.cfg";
```

```
data _null_;
   file mymail;
   put 'Jim,';
   put 'This is my SAS configuration file.';
   put 'I think you might like the';
   put 'new options I added.';
run;
```

The following program sends a message and two file attachments to multiple recipients. For this example, the e-mail options are specified in the FILE statement instead of the FILENAME statement.

```
filename outbox email "ron@acme.com";

data _null_;
   file outbox
      to=("ron@acme.com" "humberto@acme.com")
         /* Overrides value in */
         /* filename statement */
      cc=("miguel@acme.com" "loren@acme.com")
      subject="My SAS Output"
      attach=("C:\sas\results.out" "C:\sas\code.sas")
   ;
   put 'Folks,';
   put 'Attached is my output from the SAS';
   put 'program I ran last night.';
   put 'It worked great!';
run;
```

Example 2: Using Conditional Logic in a DATA Step You can use conditional logic in a DATA step in order to send multiple messages and control which recipients get which message. For example, in order to send customized reports to members of two different departments, the following program produces an e-mail message and attachments that are dependent on the department to which the recipient belongs. In the program, the following occurs:

1. In the first PUT statement, the !EM_TO! directive assigns the TO attribute.
2. The second PUT statement assigns the SUBJECT attribute using the !EM_SUBJECT! directive.
3. The !EM_SEND! directive sends the message.
4. The !EM_NEWMSG! directive clears the message attributes, which must be used to clear message attributes between recipients.
5. The !EM_ABORT! directive abnormally ends the message before the RUN statement causes it to be sent again. The !EM_ABORT! directive prevents the message from being automatically sent at the end of the DATA step.

```
filename reports email "Jim.Smith@work.com";

data _null_;
   file reports;
   length name dept $ 21;
   input name dept;
   put '!EM_TO! ' name;
   put '!EM_SUBJECT! Report for ' dept;
   put name ',';
   put 'Here is the latest report for ' dept '.' ;
```

```
   if dept='marketing' then
      put '!EM_ATTACH! c:\mktrept.txt';
   else  /* ATTACH the appropriate report */
      put '!EM_ATTACH! c:\devrept.txt';
      put '!EM_SEND!';
      put '!EM_NEWMSG!';
      put '!EM_ABORT!';
   datalines;
Susan          marketing
Peter          marketing
Alma           development
Andre          development
;
run;
```

Example 3: Sending Procedure Output in E-mail You can use e-mail to send procedure output. This example illustrates how to send ODS HTML in the body of an e-mail message. Note that ODS HTML procedure output must be sent with the RECORD_SEPARATOR (RS) option set to NONE.

```
filename outbox email
   to='susan@site.com'
   type='text/html'
   subject='Temperature Conversions';

data temperatures;
   do centigrade = -40 to 100 by 10;
      fahrenheit = centigrade*9/5+32;
      output;
   end;
run;

ods html
   body=outbox /* Mail it! */
   rs=none;

title 'Centigrade to Fahrenheit Conversion Table';

proc print;
   id centigrade;
   var fahrenheit;
run;

ods html close;
```

Example 4: Creating and E-mailing an Image The following example illustrates how to create a GIF image and send it from SAS in an e-mail message:

```
filename gsasfile email
   to='Jim@acme.com'
   type='image/gif'
   subject="SAS/GRAPH Output";

goptions dev=gif gsfname=gsasfile;

proc gtestit pic=1;
```

```
run;
```

See Also

Statements:

"FILENAME Statement" on page 1520
"FILENAME Statement, CATALOG Access Method" on page 1526
"FILENAME Statement, FTP Access Method" on page 1542
"FILENAME Statement, SOCKET Access Method" on page 1559
"FILENAME Statement, SFTP Access Method" on page 1554
"FILENAME Statement, URL Access Method" on page 1563

The SMTP E-Mail Interface in *SAS Language Reference: Concepts*

FILENAME Statement, FTP Access Method

Enables you to access remote files by using the FTP protocol.

Valid: anywhere

Category: Data Access

Syntax

FILENAME *fileref* FTP '*external-file*' <*ftp-options*>;

Arguments

fileref
is a valid fileref.

Tip: The association between a fileref and an external file lasts only for the duration of the SAS session or until you change it or discontinue it with another FILENAME statement. You can change the fileref for a file as often as you want.

FTP
specifies the access method that enables you to use File Transfer Protocol (FTP) to read from or write to a file from any host computer that you can connect to on a network with an FTP server running.

Tip: Use FILENAME with FTP when you want to connect to the host computer, to log in to the FTP server, to make records in the specified file available for reading or writing, and to disconnect from the host computer.

'external-file'
specifies the physical name of an external file that you want to read from or write to. The physical name is the name that is recognized by the operating environment.

If the file has an IBM 370 format and a record format of FB or FBA, and if the ENCODING= option is specified, then you must also specify the LRECL= option. If the length of a record is shorter than the value of LRECL, then SAS pads the record with blanks until the record length is equal to the value of LRECL.

Operating Environment Information: For details about specifying the physical names of external files, see the SAS documentation for your operating environment. △

Tip: If you are not transferring a file but performing a task such as retrieving a directory listing, then you do not need to specify a filename. Instead, put empty quotation marks in the statement. See Example 1 on page 1549.

Tip: You can associate a fileref with a single file or with an aggregate file storage location.

Tip: If you use the DIR option, specify the directory in this argument.

ftp-options
specifies details that are specific to your operating environment such as file attributes and processing attributes.

Operating Environment Information: For more information about some of these FTP options, see the SAS documentation for your operating environment. △

FTP Options

AUTHDOMAIN="*auth-domain*"
specifies the name of an authentication domain metadata object in order to connect to the FTP server. The authentication domain references credentials (user ID and password) without your having to explicitly specify the credentials. The *auth-domain* name is case sensitive, and it must be enclosed in double quotation marks.

An administrator creates authentication domain definitions while creating a user definition with the User Manager in SAS Management Console. The authentication domain is associated with one or more login metadata objects that provide access to the FTP server and is resolved by the BASE engine calling the SAS Metadata Server and returning the authentication credentials.

Requirement: The authentication domain and the associated login definition must be stored in a metadata repository, and the metadata server must be running in order to resolve the metadata object specification.

Interaction: If you specify AUTHDOMAIN=, you do not need to specify USER= and PASS=.

See also: For more information about creating and using authentication domains, see the discussion on credential management in the *SAS Intelligence Platform: Security Administration Guide*.

BINARY
is fixed-record format. Thus, all records are of size LRECL with no line delimiters. Data is transferred in image (binary) mode.

The BINARY option overrides the value of RECFM= in the FILENAME FTP statement, if specified, and forces a binary transfer.

Alias: RECFM=F

Interaction: If you specify the BINARY option and the S370V or S370VS option, then SAS ignores the BINARY option.

BLOCKSIZE=*blocksize*
where *blocksize* is the size of the data buffer in bytes.

Default: 32768

CD='*directory*'
issues a command that changes the working directory for the file transfer to the *directory* that you specify.

Interaction: The CD and DIR options are mutually exclusive. If both are specified, FTP ignores the CD option and SAS writes an informational note to the log.

DEBUG
writes to the SAS log informational messages that are sent to and received from the FTP server.

DIR
enables you to access directory files or PDS/PDSE members. Specify the directory name in the *external-file* argument. You must use valid directory syntax for the specified host.

Tip: If you want FTP to append a file extension of DATA to the member name that is specified in the FILE or INFILE statement, then use the FILEEXT option in conjunction with the DIR option. The FILEEXT option is ignored if you specify a file extension in the FILE or INFILE statement.

Tip: If you want FTP to create the directory, then use the NEW option in conjunction with the DIR option. The NEW option will be ignored if the directory exists.

Tip: If the NEW option is omitted and you specify an invalid directory, then a new directory will not be created and you will receive an error message.

Tip: The maximum number of directory or z/OS PDSE members that can be open simultaneously is limited by the number of sockets that can be open simultaneously on an FTP server. The number of sockets that can be open simultaneously is proportional to the number of connections that are set up during the installation of the FTP server. You might want to limit the number of sockets that are open simultaneously to avoid performance degradation.

Interaction: The CD and DIR options are mutually exclusive. If both are specified, FTP ignores the CD option and SAS writes an informational note to the log.

Featured in: Example 10 on page 1552

ENCODING=*encoding-value*
specifies the encoding to use when reading from or writing to the external file. The value for ENCODING= indicates that the external file has a different encoding from the current session encoding.

When you read data from an external file, SAS transcodes the data from the specified encoding to the session encoding. When you write data to an external file, SAS transcodes the data from the session encoding to the specified encoding.

Default: SAS assumes that an external file is in the same encoding as the session encoding.

Tip: The data is transferred in image or binary format and is in local data format. Thus, you must use appropriate SAS informats to read the data correctly.

See Also: "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*

FILEEXT
specifies that the member type of DATA is automatically appended to the member name in the FILE or INFILE statement when you use the DIR option.

Tip: The FILEEXT option is ignored if you specify a file extension in the FILE or INFILE statement.

See Also: LOWCASE_MEMNAME option on page 1545

Featured in: Example 10 on page 1552

HOST='*host*'
where *host* is the network name of the remote host with the FTP server running. You can specify either the name of the host (for example, **`server.pc.mydomain.com`**) or the IP address of the computer (for example, **`2001:db8::`**).

HOSTRESPONSELEN='*size*'
where *size* is the length of the FTP server response message.

Default: 2048 bytes

Range: 2048 to 16384 bytes

Restriction: If you specify a *size* that is less than 2048 or is greater than 16384, the *size* will be set to 2048.

LIST
issues the LIST command to the FTP server. LIST returns the contents of the working directory as records that contain all of the file attributes that are listed for each file.

Tip: The file attributes that are returned will vary, depending on the FTP server that is being accessed.

LOWCASE_MEMNAME
enables autocall macro retrieval of lowercase directory or member names from FTP servers.

Restriction: SAS autocall macro retrieval always searches for uppercase directory member names. Mixed case directory or member names are not supported.

Interaction: If you access files off FTP servers by using the %INCLUDE, FILE, INFILE, or other DATA step I/O statements, case sensitivity will be preserved.

See Also: FILEEXT option on page 1544

LRECL=*lrecl*
where *lrecl* is the logical record length of the data.

Default: 256

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

LS
issues the LS command to the FTP server. LS returns the contents of the working directory as records with no file attributes.

Tip: The file attributes that are returned will vary, depending on the FTP server that is being accessed.

Tip: To return a listing of a subset of files, use the LSFILE= option in addition to LS.

LSFILE='*character-string*'
in combination with the LS option, specifies a character string that enables you to request a listing of a subset of files from the working directory. Enclose the character string in quotation marks.

Restriction: LSFILE= can be used only if LS is specified.

Tip: You can specify a wildcard as part of '*character-string* '.

Tip: The file attributes that are returned will vary, depending on the FTP server that is being accessed.

Example: This statement lists all of the files that start with **sales** and end with **sas**:

```
filename myfile ftp '' ls lsfile='sales*.sas'
        other-ftp-options;
```

MGET
transfers multiple files, similar to the FTP command MGET.

Tip: The whole transfer is treated as one file. However, as the transfer of each new file is started, the EOV= variable is set to 1.

Tip: Specify MPROMPT to prompt the user before each file is sent.

MPROMPT
specifies whether to prompt for confirmation that a file is to be read, if necessary, when the user executes the MGET option.

Restriction: The MPROMPT option is not available on z/OS for batch processing.

NEW
specifies that you want FTP to create the directory when you use the DIR option.

Tip: The NEW option will be ignored if the directory exists.

Restriction: The NEW option is not available under z/OS.

PASS='*password*'
where *password* is the password to use with the user name specified in the USER= option.

Tip: You can specify the PROMPT option instead of the PASS option, which tells the system to prompt you for the password.

Tip: If the user name is **anonymous**, then the remote host might require that you specify your e-mail address as the password.

Tip: To use an encoded password, use the PWENCODE procedure in order to disguise the text string, and then enter the encoded password for the PASS= option. For more information, see the"PWENCODE Procedure" in the *Base SAS Procedures Guide*.

Featured in: Example 6 on page 1551

PORT=*portno*
where *portno* is the port that the FTP daemon monitors on the respective host. The *portno* can be any number between 0 and 65535 that uniquely identifies a service.

Tip: In the Internet community, there is a list of predefined port numbers for specific services. For example, the default port for FTP is 21. A partial list of port numbers is usually available in the **/etc/services** file on any UNIX computer.

PROMPT
specifies to prompt for the user login password, if necessary.

Restriction: The PROMPT option is not available for batch processing under z/OS.

Interaction: If PROMPT is specified without USER=, then the user is prompted for an ID, as well as a password.

Tip: You can use the SAVEUSER on page 1548 option to save the user ID and password after the user ID and password prompt is successfully executed.

RCMD= '*command* '
where *command* is the FTP 'SITE' or 'service' command to send to the FTP server.

FTP servers use SITE commands to provide services that are specific to a system and are essential to file transfer but not common enough to be included in the protocol.

For example, **rcmd='site rdw'** preserves the record descriptor word (RDW) of a z/OS variable blocked data set as a part of the data. See S370V and S370VS below.

Interaction: Some FTP service commands might not run at a particular client site depending on the security permissions and the availability of the commands.

Tip: If you transfer a file with the FTP access method and then cannot read the file, you might need to change the FTP server's UMASK setting.

If the FTP server supports a SITE UMASK setting, you can change the permissions of the file as shown in the following example:

```
filename in ftp '/mydir/accounting/file2.dat'
   host="xxx.fyi.xxx.com"
   user="john"
   rcmd='site umask 022'
   prompt;

data _null;
file in;
put a $80;
run;
```

Tip: You can specify multiple FTP service commands if you separate them by semicolons. Some examples are as follows:

```
rcmd='ascii;site umask 002'


rcmd='stat;site chmod 0400 ~mydir/abc.txt'
```

RECFM=*recfm*

where *recfm* is one of three record formats:

F — is fixed-record format. Thus, all records are of size LRECL with no line delimiters. Data is transferred in image (binary) mode.

Alias: BINARY

The BINARY option overrides the value of RECFM= in the FILENAME FTP statement, if specified, and forces a binary transfer.

S — is stream-record format. Data is transferred in image (binary) mode.

Interaction: The amount of data that is read is controlled by the current LRECL value or by the value of the NBYTE= variable in the INFILE statement. The NBYTE= option specifies a variable that is equal to the amount of data to be read. This amount must be less than or equal to LRECL.

See Also: The NBYTE= option on page 1598 in the INFILE statement.

V — is variable-record format (the default). In this format, records have varying lengths, and they are transferred in text (stream) mode.

Interaction: Any record larger than LRECL is truncated.

Tip: If you are using files with the IBM 370 Variable format or the IBM 370 Spanned Variable format, then you might want

to use the S370V or S370VS options instead of the RECFM= option. See S370V and S370VS below.

Default: V

Interaction: If you specify the RECFM= option and the S370V or S370VS option, then SAS ignores the RECFM= option.

RHELP
issues the HELP command to the FTP server. The results of this command are returned as records.

RSTAT
issues the RSTAT command to the FTP server. The results of this command are returned as records.

SAVEUSER
saves the user ID and password after the user ID and password prompt are successfully executed.

Interaction: The user ID and password are saved only for the duration of the SAS session or until you change the association between the fileref and the external file, or discontinue it with another FILENAME statement.

S370V
indicates that the file being read is in IBM 370 variable format.

Interaction: If you specify this option and the RECFM= option, then SAS ignores the RECFM= option.

Tip: The data is transferred in image or binary format and is in local data format. Thus, you must use appropriate SAS informats to read the data correctly on non-EBCDIC hosts.

Tip: Use the **rcmd='site rdw'** option when you transfer a z/OS data set with a variable-record format to another z/OS data set with a variable-record format to preserve the record descriptor word (rdw) of each record. By default, most FTP servers remove the rdw that exists in each record before it is transferred.

Typically, the 'SITE RDW' command is not necessary when you transfer a data set with a z/OS variable-record format to ASCII, or when you transfer an ASCII file to a z/OS variable-record format.

S370VS
indicates that the file that is being read is in IBM 370 variable-spanned format.

Interaction: If you specify this option and the RECFM= option, then SAS ignores the RECFM= option.

Tip: The data is transferred in image or binary format and is in local data format. Thus, you must use appropriate SAS informats to read the data correctly on non-EBCDIC hosts.

Tip: Use the **rcmd='site rdw'** option when you transfer a z/OS data set with a variable-record format to another z/OS data set with a variable-record format to preserve the record descriptor word (rdw) of each record. By default, most FTP servers remove the rdw that exists in each record before it is transferred.

Typically, the 'SITE RDW' command is not necessary when you transfer a data set with a z/OS variable-record format to ASCII, or when you transfer an ASCII file to a z/OS variable-record format.

USER='*username*'
where *username* is used to log in to the FTP server.

Restriction: The FTP access method does not support FTP proxy servers that require user ID authentication.

Interaction: If PROMPT is specified, but USER= is not, then the user is prompted for an ID.

Tip: You can specify a proxy server and credentials for an FTP server when using the FTP access method. The user ID and password that you need to log in to the FTP server is sent via the proxy server by using the **user="*userid@ftpservername*" pass="*password*" host="*proxy.server.xxx.com*"** syntax. Both anonymous and user ID validation are supported.

Featured in:

TERMSTR='*eol-char*'
where *eol-char* is the line delimiter to use when RECFM=V. There are three valid values:

CRLF	carriage return (CR) followed by line feed (LF).
LF	line feed only (the default).
NULL	NULL character (0x00).

Default: LF

Restriction: Use this option only when RECFM=V.

WAIT_MILLISECONDS=*milliseconds*
specifies the FTP response time in milliseconds.

Default: 1,000 milliseconds

Tip: If you receive a "connection closed; transfer aborted" or "network name is no longer available" message in the log, use the WAIT_MILLISECONDS option to increase the response time.

Comparisons

As with the FTP **get** and **put** commands, the FTP access method lets you download and upload files. However, this method directly reads files into your SAS session without first storing them on your system.

Examples

Example 1: Retrieving a Directory Listing This example retrieves a directory listing from a host named **mvshost1** for user **smythe**, and prompts **smythe** for a password:

```
filename dir ftp '' ls user='smythe'
         host='mvshost1.mvs.sas.com' prompt;

data _null_;
   infile dir;
   input;
   put _INFILE_;
run;
```

Note: The quotation marks are empty because no file is being transferred. Because quotation marks are required by the syntax, however, you must include them. △

Example 2: Reading a File from a Remote Host This example reads a file called **sales** in the directory **/u/kudzu/mydata** from the remote UNIX host **hp720**:

```
filename myfile ftp 'sales' cd='/u/kudzu/mydata'
         user='guest' host='hp720.hp.sas.com'
         recfm=v prompt;

data mydata / view=mydata;   /* Create a view */
   infile myfile;
   input x $10. y 4.;
run;

proc print data=mydata;     /* Print the data */
run;
```

Example 3: Creating a File on a Remote Host This example creates a file called **test.dat** in a directory called **c:\remote** for the user **bbailey** on the host **winnt.pc**:

```
filename create ftp 'c:\remote\test.dat'
         host='winnt.pc'
         user='bbailey' prompt recfm=v;

data _null_;
   file create;
   do i=1 to 10;
      put i=;
   end;
run;
```

Example 4: Reading an S370V-Format File on z/OS This example reads an S370V-format file from a z/OS system. See RCMD= on page 1546 for more information about RCMD='**site rdw**'.

```
filename viewdata ftp 'sluggo.stat.data'
         user='sluggo' host='zoshost1'
         s370v prompt rcmd='site rdw';

data mydata / view=mydata;   /* Create a view */
   infile viewdata;
   input x $ebcdic8.;
run;

proc print data=mydata;     /* Print the data */
run;
```

Example 5: Anonymously Logging In to FTP This example shows how to log in to FTP anonymously, if the host accepts anonymous logins.

Note: Some anonymous FTP servers require a password. If required, your e-mail address is usually used. See PASS= on page 1546 under "FTP Options." △

```
filename anon ftp '' ls host='130.96.6.1'
         user='anonymous';

data _null_;
   infile anon;
   input;
```

```
   list;
run;
```

Note: The quotation marks following the argument FTP are empty. A filename is needed only when transferring a file, not when routing a command. The quotation marks, however, are required. △

Example 6: Using an Encoded Password This example shows you how to use an encoded password in the FILENAME statement.

In a separate SAS session, use the PWENCODE procedure to encode your password and make note of the output.

```
proc pwencode in= "MyPass1";
run;
```

The following output appears in the SAS log:

```
(sas001)TX1QYXNzMQ==
```

You can now use the entire encoded password string in your batch program.

```
filename myfile ftp 'sales' cd='/u/kudzu/mydata'
   user='tjbarry' host='hp720.hp.mycompany.com'
   pass="(sas001)TX1QYXMZ==";
```

Example 7: Importing a Transport Data Set

This example uses the CIMPORT procedure to import a transport data set from a host named **myshost1**for user **calvin**. The new data set will reside locally in the SASUSER library. Note that user and password can be SAS macro variables. If you specify a fully qualified data set name, then use double quotation marks and single quotation marks. Otherwise, the system will append the profile prefix to the name that you specify.

```
   %let user=calvin;
   %let pw=xxxxx;
   filename inp ftp "'calvin.mat1.cpo'" user="&user"
            pass="&pw" rcmd='binary'
            host='mvshost1';

  proc cimport library=sasuser infile=inp;
  run;
```

Example 8: Transporting a SAS Library This example uses the CPORT procedure to transport a SAS library to a host named **mvshost1** for user **calvin**. It will create a new sequential file on the host called **userid.mat64.cpo** with the recfm of **fb**, lrecl of 80, and blocksize of 8000.

```
filename inp ftp 'mat64.cpo' user='calvin'
         pass="xxxx" host='mvshost1'
         lrecl=80 recfm=f blocksize=8000
         rcmd='site blocksize=800 recfm=fb lrecl=80';

proc cport library=mylib file=inp;
run;
```

Example 9: Creating a Transport Library with Transport Engine This example creates a new SAS library on host **mvshost1**. The FILENAME statement assigns a fileref to the new data set. Note the use of the RCMD= option to specify important file attributes. The LIBNAME statement uses a libref that is the same as the fileref and assigns it to the XPORT engine. The PROC COPY step copies all data sets from the SAS library that are referenced by MYLIB to the XPORT engine. Output from the PROC CONTENTS step confirms that the copy was successful:

```
filename inp ftp  'mat65.cpo' user='calvin'
         pass="xxxx"  host='mvshost1'
         lrecl=80 recfm=f blocksize=8000
     rcmd='site blocksize=8000 recfm=fb lrecl=80';

libname mylib 'SAS-library';
libname inp xport;

proc copy in=mylib out=inp mt=data;
run;

proc contents data=inp._all_;
run;
```

Note: For more information about the XPORT engine, see "The Transport Engine" in *SAS Language Reference: Concepts* and "XPORT Engine Limitations" in *Moving and Accessing SAS Files*. △

Example 10: Reading and Writing from Directories This example reads the file **ftpmem1** from a directory on a UNIX host, and writes the file **ftpout1** to a different directory on another UNIX host.

```
filename indir ftp '/usr/proj2/dir1' DIR
                   host="host1.mycompany.com"
                   user="xxxx" prompt;

filename outdir ftp '/usr/proj2/dir2' DIR FILEEXT
                    host="host2.mycompany.com"
                    user="xxxx" prompt;

data _null_;
   infile indir(ftpmem1) truncover;
   input;
   file outdir(ftpout1);
   put _infile_;
run;
```

The file **ftpout1** is written to **/usr/proj2/dir2/ftpout1.DATA**. Note that a member type of DATA is appended to the **ftpout1** file because the FILEEXT option was specified in the output file's FILENAME statement. For more information, see FILEEXT on page 1544 .

Note: The DIR option is not needed for some ODS destinations. △

The following example writes an output file and transfers it to an ODS-specified destination. The DIR option is not needed.

```
filename output ftp "~user/ftpdir/" host="host.fyi.company.com" user="userid"
pass="userpass" recfm=s debug;
```

```
ods listing close;
ods html body='body.html' path=output;
proc print data=sashelp.class;run;
ods html close;
ods listing;
```

To export multiple graph files to a remote directory location, the DIR option must be specified in the FILENAME statement. Accordingly, when creating external graph files with the ODS HTML destination, two FILENAME statements are needed: one for the HTML files, and one for the graph files. The following example illustrates the need for two FILENAME statements.

```
filename output1 ftp "~user/dir" fileext host="host.unx.company.com"
   user="userid" pass="userpass" recfm=s debug;
filename output2 ftp "~user/dir" dir fileext host="host.unx.company.com"
   user="userid" pass="userpass" recfm=s debug;
ods listing close;
ods html body='body.html' path=output1 gpath=output2
   frame='frames.html' contents='contents.html';
proc gtestit;run;quit;
ods html close;
ods listing;
```

Example 11: Using a Proxy Server This example uses a proxy server with the FTP access method. The user ID and password are sent via the proxy server.

```
filename test ftp ' ' ls
   host='proxy.server.xxx.com'
   user='userid@ftpservername'
   pass='xxxxxx'
   cd='pubsdir/';

data _null_;
   infile test truncover;
   input a $256.;
   put a=;
run;
```

See Also

Statements:

"FILENAME Statement" on page 1520
"FILENAME Statement, CATALOG Access Method" on page 1526
"FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532
"FILENAME Statement, SOCKET Access Method" on page 1559
"FILENAME Statement, SFTP Access Method" on page 1554
"FILENAME Statement, URL Access Method" on page 1563
"LIBNAME Statement" on page 1656

FILENAME Statement, SFTP Access Method

Enables you to access remote files by using the SFTP protocol.

Valid: anywhere

Category: Data Access

Syntax

FILENAME *fileref* SFTP '*external-file*' <*sftp-options*>;

Arguments

fileref
is a valid fileref.

Tip: The association between a fileref and an external file lasts only for the duration of the SAS session or until you change it or discontinue it with another FILENAME statement. You can change the fileref for a file as often as you want.

SFTP
specifies the access method that enables you to use Secure File Transfer Protocol (SFTP) to read from or write to a file from any host computer that you can connect to on a network with an OpenSSH SSHD server running.

'*external-file*'
specifies the physical name of an external file that you want to read from or write to. The physical name is the name that is recognized by the operating environment.

Operating Environment Information: For details about specifying the physical names of external files, see the SAS documentation for your operating environment. △

Tip: If you are not transferring a file but performing a task such as retrieving a directory listing, then you do not need to specify an external filename. Instead, put empty quotation marks in the statement.

Tip: You can associate a fileref with a single file or with an aggregate file storage location.

sftp-options
specifies details that are specific to your operating environment such as file attributes and processing attributes.

Operating Environment Information: For more information about some of these SFTP options, see the SAS documentation for your operating environment. △

SFTP Options

BATCHFILE='*path*'
specifies the fully qualified pathname and the filename of the batch file that contains the SFTP commands. These commands are submitted when the SFTP access method is executed. After the batch file processing ends, the SFTP connection is closed.

Requirement: The path must be enclosed in quotation marks.

Tip: After the batch file processing ends, the SFTP connection is closed and the filename assignment is no longer available. If subsequent DATA step processing requires the FILENAME SFTP statement, then another FILENAME SFTP statement is required.

Featured in: Example 5 on page 1558

CD='*directory*'
issues a command that changes the working directory for the file transfer to the *directory* that you specify.

DEBUG
writes informational messages to the SAS log.

DIR
enables you to access directory files. Specify the directory name in the external-file argument. You must use valid directory syntax for the specified host.

Interaction: The CD and DIR options are mutually exclusive. If both are specified, SFTP ignores the CD option and SAS writes an informational note to the log.

Tip: If you want SFTP to create the directory, then use the NEW option in conjunction with the DIR option. The NEW option will be ignored if the directory exists.

Tip: If the NEW option is omitted and you specify an invalid directory, then a new directory will not be created and you will receive an error message.

HOST='*host*'
where *host* is the network name of the remote host with the OpenSSH SSHD server running.

You can specify either the name of the host (for example, **`server.pc.mydomain.com`**) or the IP address of the computer (for example, **`2001:db8::`**).

LRECL=*lrecl*
where *lrecl* is the logical record length of the data.

Default: 256

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

LS
issues the LS command to the SFTP server. LS returns the contents of the working directory as records with no file attributes.

Restriction: The LS option will not display files with leading periods, for example *.xAuthority*.

Interaction: The LS and LSA options are mutually exclusive. If you specify both options, the LSA option takes precedence.

Tip: To return a listing of a subset of files, use the LSFILE= option in addition to LS.

LSA
issues the LS command to the SFTP server. LSA returns all the contents of the working directory as records with no file attributes.

Interaction: The LS and LSA options are mutually exclusive. If you specify both options, the LSA option takes precedence.

Interaction: To display files without leading periods, for example *.xAuthority*, use the LS= option.

Tip: To return a listing of a subset of files, use the LSFILE= option in addition to LSA.

LSFILE='*character-string*'
in combination with the LS option, specifies a character string that enables you to request a listing of a subset of files from the working directory. Enclose the character string in quotation marks.

Restriction: LSFILE= can be used only if LS or LSA is specified.

Tip: You can specify a wildcard as part of '*character-string* '.

Example: This statement lists all of the files that start with **sales** and end with **sas**:

```
filename myfile sftp '' ls lsfile='sales*.sas'
         other-sftp-options;
```

MGET
transfers multiple files, similar to the SFTP command MGET.

Tip: The whole transfer is treated as one file. However, as the transfer of each new file is started, the EOV= variable is set to 1.

NEW
specifies that you want SFTP to create the directory when you use the DIR option.

Restriction: The NEW option is not available under z/OS.

Tip: The NEW option will be ignored if the directory exists.

OPTIONS=
specifies SFTP configuration options such as port numbers.

PATH
specifies the location of the SFTP executable if it is not installed in the PATH or $PATH search path.

Tip: It is recommended that the OpenSSH "SFTP" executable or PUTTY "PSFTP" executable be installed in a directory that is accessible via the PATH or $PATH search path.

RECFM=*recfm*
where *recfm* is one of two record formats:

F
is fixed-record format. Thus, all records are of size LRECL with no line delimiters.

V
is variable-record format (the default). In this format, records have varying lengths, and they are separated by newlines. Data is transferred in image (binary) mode.

Default: V

USER='*username*'
specifies the user name.

Requirement: The *username* is required by the PUTTY client on the Windows host.

Tip: The *username* is not typically required on LINUX or UNIX hosts when using public key authentication.

Tip: Public key authentication using an SSH agent is the recommended way to connect to a remote SSHD server.

WAIT_MILLISECONDS=*milliseconds*
specifies the SFTP response time in milliseconds.

Default: 1,500 milliseconds

Tip: If you receive a timeout message in the log, use the WAIT_MILLISECONDS option to increase the response time.

Details

The Basics

The Secure File Transfer Protocol (SFTP) provides a secure connection and file transfers between two hosts (client and server) over a network. Both commands and data are encrypted. The client machine initiates a connection with the remote host (OpenSSH SSHD server).

With the SFTP access method, you can read from or write to any host computer that you can connect to on a network with an OpenSSH SSHD server running. The client and server applications can reside on the same computer or on different computers that are connected by a network.

Specific implementation details are dependent on the OpenSSH SSHD server version and how that site is configured.

The SFTP access method relies on default send and reply messages to OpenSSH commands. Custom installs of OpenSSH that modify these messages will disable the SFTP access method.

You must have the applicable client software installed to use the SFTP access method. The SFTP access method supports only the following SSH clients.

- OpenSSH - UNIX
- PUTTY – Windows

Note: Password validation is not supported for the SFTP access method. △

Note: Public key authentication using an SSH agent is the recommended way to connect to a remote SSHD server. △

Note: If you have trouble running the SFTP access method try to manually validate SFTP client access to an OpenSSH SSHD server without involving the SAS system. Manually validating SFTP client access without involving the SAS system will ensure that your SSH/SSHD configuration and key authentication is setup correctly. △

SFTP Access Methods and SFTP Prompts

The SFTP access method supports only the following prompts. Changing the prompt will disable the SFTP access method.

- For OpenSSH:

  sftp>

  sftp >

- For PUTTY:

  psftp>

Comparisons

As with the SFTP **`get`** and **`put`** commands, the SFTP access method lets you download and upload files. However, this method directly reads files into your SAS session without first storing them on your system.

Examples

Example 1: Connecting to an SSHD Server at a Standard Port This example reads a file called **test.dat** using the SFTP access method after connecting to the SSHD server a standard port:

```
filename myfile sftp '/users/xxxx/test.dat' host="unixhost1";

data _null_;
   infile myfile truncover;
   input a $25.;
run;
```

Example 2: Connecting to an SSHD Server at a Nonstandard Port This example reads a file called **test.dat** using the SFTP access method after connecting to the SSHD server at port 4117:

```
filename myfile sftp '/users/xxxx/test.dat' host="unixhost1" options="-oPort=4117";

data _null_;
   infile myfile truncover;
   input a $25.;;
run;
```

Example 3: Connecting a Windows PUTTY Client to an SSHD Server This example writes a file called **test.dat** using the SFTP access method after connecting a Windows PUTTY client to the SSHD server with a userid of **userid**:

```
filename outfile sftp '/users/xxxx/test.dat' host="unixhost1" user="userid";

data _null_;
   file outfile;
   do i=1 to 10;
      put i=;
   end;
run;
```

Example 4: Reading Files from a Directory on the Remote Host This example reads the files **test.dat** and **test2.dat** from a directory on the remote host.

```
filename infile sftp '/users/xxxx/' host="unixhost1" dir;

data _null_;
   infile infile(test.dat) truncover;
   input a $25.;

   infile infile(test2.dat) truncover;
   input b $25.;
run;
```

Example 5: Using a Batch File In this example, when the INFILE statement is processed, the batch file associated with the FILENAME SFTP statement, **sftpcmds**, is executed.

```
filename process sftp ' ' host="unixhost1" user="userid"
   batchfile="c:/stfpdir/sftpcmds.bat";
```

```
data _null_;
   infile process;
run;
```

See Also

Statements:

"FILENAME Statement" on page 1520
"FILENAME Statement, CATALOG Access Method" on page 1526
"FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532
"FILENAME Statement, FTP Access Method" on page 1542
"FILENAME Statement, SOCKET Access Method" on page 1559
"FILENAME Statement, URL Access Method" on page 1563
"LIBNAME Statement" on page 1656

Barrett, Daniel J., Richard E. Silverman, and Robert G. Byrnes. 2005. *SSH, The Secure Shell: A Definitive Guide*. Sebastopol, CA: O'Reilly

FILENAME Statement, SOCKET Access Method

Enables you to read from or write to a TCP/IP socket.

Valid: anywhere

Category: Data Access

Syntax

❶**FILENAME** *fileref* SOCKET '*hostname*:*portno*'
<*tcpip-options*>;

❷**FILENAME** *fileref* SOCKET ':*portno*' SERVER
<*tcpip-options*>;

Arguments

fileref
is a valid fileref.

Tip: The association between a fileref and an external file lasts only for the duration of the SAS session or until you change it or discontinue it with another FILENAME statement. You can change the fileref for a file as often as you want.

SOCKET
specifies the access method that enables you to read from or write to a Transmission Control Protocol/Internet Protocol (TCP/IP) socket.

'*hostname*:*portno*'
is the name or IP address of the host and the TCP/IP port number to connect to.

Tip: Use this specification for client access to the socket.

'*:portno*'
is the port number to create for listening.

Tip: Use this specification for server mode.

Tip: If you specify `:0`, the system will choose a number.

SERVER
sets the TCP/IP socket to be a listening socket, thereby enabling the system to act as a server that is waiting for a connection.

Tip: The system accepts all connections serially; only one connection is active at any one time.

See Also: The RECONN= option description on page 1561 under *TCPIP-Options*.

TCPIP-Options

BLOCKSIZE=*blocksize*
where *blocksize* is the size of the socket data buffer in bytes.

Default: 8192

ENCODING=*encoding-value*
specifies the encoding to use when reading from or writing to the socket. The value for ENCODING= indicates that the socket has a different encoding from the current session encoding.

When you read data from a socket, SAS transcodes the data from the specified encoding to the session encoding. When you write data to a socket, SAS transcodes the data from the session encoding to the specified encoding.

For valid encoding values, see "Encoding Values for SAS Language Elements" in *SAS National Language Support (NLS): Reference Guide*.

LRECL=*lrecl*
where *lrecl* is the logical record length.

Default: 256

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

RECFM=*recfm*
where *recfm* is one of three record formats:

F
is fixed record format. Thus, all records are of size LRECL with no line delimiters. Data are transferred in image (binary) mode.

S
is stream record format.

Tip: Data are transferred in image (binary) mode.

Interactions: The amount of data that is read is controlled by the current LRECL value or the value of the NBYTE= variable in the INFILE statement. The NBYTE= option specifies a variable equal to the amount of data to be read. This amount must be less than or equal to LRECL.

See Also: The NBYTE= option on page 1598 in the INFILE statement.

V
is variable record format (the default).

Tip: In this format, records have varying lengths, and they are transferred in text (stream) mode.

Tip: Any record larger than LRECL is truncated.

Default: V

RECONN=*conn-limit*
where *conn-limit* is the maximum number of connections that the server will accept.

Explanation: Because only one connection can be active at a time, a connection must be disconnected before the server can accept another connection. When a new connection is accepted, the EOV= variable is set to 1. The server will continue to accept connections, one at a time, until *conn-limit* has been reached.

TERMSTR='*eol-char*'
where *eol-char* is the line delimiter to use when RECFM=V. There are three valid values:

CRLF	carriage return (CR) followed by line feed (LF).
LF	line feed only (the default).
NULL	NULL character (0x00).

Default: LF

Restriction: Use this option only when RECFM=V.

Details

The Basics A TCP/IP socket is a communication link between two applications. The *server* application creates the socket and waits for a connection. The *client* application connects to the socket. With the SOCKET access method, you can use SAS to communicate with another application over a socket in either client or server mode. The client and server applications can reside on the same computer or on different computers that are connected by a network.

As an example, you can develop an application using Microsoft Visual Basic that communicates with a SAS session that uses the TCP/IP sockets. Note that Visual Basic does not provide inherent TCP/IP support. You can obtain a custom control (VBX) from SAS Technical Support (free of charge) that allows a Visual Basic application to communicate through the sockets.

❶Using the SOCKET Access Method in Client Mode

In client mode, a local SAS application can use the SOCKET access method to communicate with a remote application that acts as a server (and waits for a connection). Before you can connect to a server, you must know:

- the network name or IP address of the host computer running the server.
- the port number that the remote application is listening to for new connections.

The remote application can be another SAS application, but it does not need to be. When the local SAS application connects to the remote application through the TCP/IP socket, the two applications can communicate by reading from and writing to the socket as if it were an external file. If at any time the remote side of the socket is disconnected, the local side will also automatically terminate.

❷Using the SOCKET Access Method in Server Mode When the local SAS application is in server mode, it remains in a wait state until a remote application connects to it. To use the SOCKET access method in server mode, you need to know only the port number that you want the server to listen to for a connection. Typically, servers use *well-known ports* to listen for connections. These port numbers are reserved by the system for specific server applications. For more information about how well-known ports are

defined on your system, refer to the documentation for your TCP/IP software or ask your system administrator.

If the server application does not use a well-known port, then the system assigns a port number when it establishes the socket from the local application. However, because any client application that waits to connect to the server must know the port number, you should try to use a well-known port.

While a local SAS server application is waiting for a connection, SAS is in a wait state. Each time a new connection is established, the EOV= variable in the DATA step is set to 1. Because the server accepts only one connection at a time, no new connections can be established until the current connection is closed. The connection closes automatically when the remote client application disconnects. The SOCKET access method continues to accept new connections until it reaches the limit set in the RECONN option.

Examples

Example 1: Communicating between Two SAS Applications Over a TCP/IP Socket This example shows how two SAS applications can talk over a TCP/IP socket. The local application is in server mode; the remote application is the client that connects to the server. This example assumes that the server host name is **hp720.unx.sas.com**, that the well-known port number is 5000, and that the server allows a maximum of three connections before closing the socket.

Here is the program for the server application:

```
filename local socket ':5000' server reconn=3;
   /*The server is using a reserved */
   /*port number of 5000.           */

data tcpip;
   infile local eov=v;
   input x $10;
   if v=1 then
      do;              /* new connection when v=1 */
         put 'new connection received';
      end;
   output;
run;
```

Here is the program for the remote client application:

```
filename remote socket 'hp720.unx.sas.com:5000';

data _null_;
   file remote;
   do i=1 to 10;
      put i;
   end;
run;
```

See Also

Statements:

"FILENAME Statement" on page 1520

FILENAME Statement, URL Access Method

Enables you to access remote files by using the URL access method.

Valid: anywhere
Category: Data Access

Syntax

FILENAME *fileref* URL '*external-file*'
<*url-options*>;

Arguments

fileref
is a valid fileref.

Tip: The association between a fileref and an external file lasts only for the duration of the SAS session or until you change it or discontinue it with another FILENAME statement. You can change the fileref for a file as often as you want.

URL
specifies the access method that enables you to read a file from any host computer that you can connect to on a network with a URL server running.

Alias: HTTP

'external-file'
specifies the name of the file that you want to read from on a URL server. The Secure Socket Layer (SSL) protocol, https, can also be used to access the files. The file must be specified in one of these formats:

http://*hostname/file*

https://*hostname/file*

http://*hostname:portno/file*

https://*hostname:portno/file*

Operating Environment Information: For details about specifying the physical names of external files, see the SAS documentation for your operating environment. △

url-options
can be any of the following:

AUTHDOMAIN="*auth-domain*"
specifies the name of an authentication domain metadata object in order to connect to the proxy or Web server. The authentication domain references

credentials (user ID and password) without your having to explicitly specify the credentials. The *auth-domain* name is case sensitive, and it must be enclosed in double quotation marks.

An administrator creates authentication domain definitions while creating a user definition with the User Manager in SAS Management Console. The authentication domain is associated with one or more login metadata objects that provide access to the proxy or Web server and is resolved by the BASE engine calling the SAS Metadata Server and returning the authentication credentials.

Requirement: The authentication domain and the associated login definition must be stored in a metadata repository, and the metadata server must be running in order to resolve the metadata object specification.

Interaction: If you specify AUTHDOMAIN=, you do not need to specify USER= and PASS=.

See also: For more information about creating and using authentication domains, see the discussion on credential management in the *SAS Intelligence Platform: Security Administration Guide*.

BLOCKSIZE=*blocksize*
where *blocksize* is the size of the URL data buffer in bytes.

Default: 8K

DEBUG
writes debugging information to the SAS log.

Tip: The result of the HELP command is returned as records.

HEADERS=*fileref*
specifies the fileref to which the header information is written when a file is opened by using the URL access method. The header information is the same information that is written to the SAS log.

Requirement: The fileref must be defined in a previous FILENAME statement.

Interaction: If you specify the HEADERS= option without specifying the DEBUG option, the DEBUG option is automatically turned on.

Interaction: By default, log information is overwritten. To append the log information, you must specify the MOD option in the FILENAME statement that creates the fileref.

LRECL=*lrecl*
where *lrecl* is the logical record length of the data.

Default: 256

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

PASS='*password*'
where *password* is the password to use with the user name that is specified in the USER option.

Tip: You can specify the PROMPT option instead of the PASS option, which tells the system to prompt you for the password.

Tip: To use an encoded password, use the PWENCODE procedure in order to disguise the text string, and then enter the encoded password for the PASS= option. For more information, see the PWENCODE Procedure in the *Base SAS Procedures Guide*.

PPASS='*password*'
where *password* is the password to use with the user name that is specified in the PUSER option. The PPASS option is used to access the proxy server.

Tip: You can specify the PROMPT option instead of the PPASS option, which tells the system to prompt you for the password.

Tip: To use an encoded password, use the PWENCODE procedure to disguise the text string, and then enter the encoded password for the PASS= option. For more information, see the PWENCODE procedure in the *Base SAS Procedures Guide*.

PROMPT
specifies to prompt for the user login password if necessary.

Tip: If you specify PROMPT, you do not need to specify PASS= or PPASS=.

PROXY=*url*
specifies the Uniform Resource Locator (URL) for the proxy server in one of these forms:

http://*hostname*/

http://*hostname:portno*/

PUSER='*username*'
where *username* is used to log on to the URL proxy server.

Tip: If you specify **`puser='*'`**, then the user is prompted for an ID.

Interaction: If you specify the PUSER option, the USER option goes to the Web server regardless of whether you specify a proxy server.

Interaction: If PROMPT is specified, but PUSER is not, the user is prompted for an ID as well as a password.

RECFM=*recfm*
where *recfm* is one of three record formats:

F — is fixed-record format. Thus, all records are of size LRECL with no line delimiters. Data is transferred in image (binary) mode.

S — is stream-record format. Data is transferred in image (binary) mode.

Alias: N

Tip: The amount of data that is read is controlled by the current LRECL value or the value of the NBYTE= variable in the INFILE statement. The NBYTE= option specifies a variable that is equal to the amount of data to be read. This amount must be less than or equal to LRECL.

See Also: The NBYTE= option on page 1598 in the INFILE statement.

V — is variable-record format (the default). In this format, records have varying lengths, and they are transferred in text (stream) mode.

Tip: Any record larger than LRECL is truncated.

Default: V

TERMSTR='*eol-char*'
where *eol-char* is the line delimiter to use when RECFM=V. There are four valid values:

CR — carriage return (CR).

CRLF — carriage return (CR) followed by line feed (LF).

LF — line feed only (the default).

NULL NULL character (0x00).

Default: LF

Restriction: Use this option only when RECFM=V.

USER='*username*'
where *username* is used to log on to the URL server.

Tip: If you specify **user='*'**, then the user is prompted for an ID.

Interaction: If you specify the USER option but do not specify the PUSER option, where the USER option goes depends on whether you specify a proxy server. If you do not specify a proxy server, USER goes to the Web server. If you do specify a proxy server, USER will go to the proxy server.

If you specify the PUSER option, the USER option goes to the Web server regardless of whether you specify a proxy server.

Interaction: If PROMPT is specified, but USER or PUSER is not, the user is prompted for an ID as well as a password.

Details

The Secure Sockets Layer (SSL) protocol is used when the URL begins with "https" instead of "http". The SSL protocol provides network security and privacy. Developed by Netscape Communications, SSL uses encryption algorithms that include RC2, RC4, DES, tripleDES, IDEA, and MD5. Not limited to providing only encryption services, SSL can also perform client and server authentication and use message authentication codes. SSL is supported by both Netscape Navigator and Internet Explorer. Many Web sites use the protocol to provide confidential user information such as credit card numbers. The SSL protocol is application independent, enabling protocols such as HTTP, FTP, and Telnet to be layered transparently above it. SSL is optimized for HTTP.

Operating Environment Information: Using the FILENAME statement requires information that is specific to your operating environment. The URL access method is fully documented here, but for more information about how to specify filenames, see the SAS documentation for your operating environment. △

Examples

Example 1: Accessing a File at a Web Site This example accesses document **test.dat**at site **www.a.com**:

```
filename foo url 'http://www.a.com/test.dat'
         proxy='http://www.gt.sas.com';
```

Example 2: Specifying a User ID and a Password This example accesses document **file1.html** at site **www.b.com** using the SSL protocol and requires a user ID and password:

```
filename foo url 'https://www.b.com/file1.html'
         user='jones' prompt;
```

Example 3: Reading the First 15 Records from a URL File This example reads the first 15 records from a URL file and writes them to the SAS log with a PUT statement:

```
filename foo url
    'http://support.sas.com/techsup/service_intro.html';
```

```
data _null_;
   infile foo length=len;
   input record $varying200. len;
   put record $varying200. len;
   if _n_=15 then stop;
run;
```

See Also

Statements:

"FILENAME Statement" on page 1520

"FILENAME Statement, CATALOG Access Method" on page 1526

"FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532

"FILENAME Statement, FTP Access Method" on page 1542

"FILENAME Statement, SOCKET Access Method" on page 1559

"FILENAME Statement, SFTP Access Method" on page 1554

"Secure Sockets Layer (SSL)r" in *Encryption in SAS*

FILENAME Statement, WebDAV Access Method

Enables you to access remote files by using the WebDAV protocol.

Valid: Anywhere

Category: Data Access

Restriction: Access to WebDAV servers is not supported on Open VMS.

Syntax

FILENAME *filref* WEBDAV '*external-file*' <*webdav-options*>;

Arguments

fileref
is a valid fileref.

Tip: The association between a fileref and an external file lasts only for the duration of the SAS session or until you change it or discontinue it with another FILENAME statement. You can change the fileref for a file as often as you want.

WEBDAV
specifies the access method that enables you to use WebDAV (Web Distributed Authoring and Versioning) to read from or write to a file from any host machine that you can connect to on a network with a WebDAV server running.

'*external-file*'
specifies the name of the file that you want to read from or write to a WebDAV server. The external file must be in one of these forms:

http://*hostname/path-to-the-file*

https://*hostname/path-to-the-file*

http://*hostname*:*port/path-to-the-file*

https://*hostname*:*port/path-to-the-file*

Requirement: When using the HTTPS communication protocol, you must use the SSL (Secure Sockets Layer) protocol that provides secure network communications. For more information, see *Encryption in SAS*.

Operating Environment Information: For details about specifying the physical names of external files, see the SAS documentation for your operating environment. △

WebDAV Options

webdav-options can be any of the following:

DEBUG
writes debugging information to the SAS log.

DIR
enables you to access directory files. Specify the directory name in the external-file argument. You must use valid directory syntax for the specified host.

Tip: See the FILEEXT option on page 1568 for information about specifying filename extensions.

ENCODING='*encoding-value*'
specifies the encoding to use when SAS is reading from or writing to an external file. The value for ENCODING= indicates that the external file has a different encoding from the current session encoding.

When you read data from an external file, SAS transcodes the data from the specified encoding to the session encoding. When you write data to an external file, SAS transcodes the data from the session encoding to the specified encoding.

Default: SAS assumes that an external file is in the same encoding as the session encoding.

See Also: "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*

FILEEXT
specifies that a file extension is automatically appended to the filename when you use the DIR option.

Interaction: The autocall macro facility always passes the extension .SAS to the file access method as the extension to use when opening files in the autocall library. The DATA step always passes the extension .DATA. If you define a fileref for an autocall macro library and the files in that library have a file extension of .SAS, use the FILEEXT option. If the files in that library do not have an extension, do not use the FILEEXT option. For example, if you define a fileref for an input file in the DATA step and the file **x** has an extension of .DATA, you would use the FILEEXT option to read the file X.DATA. If you use the INFILE or FILE statement, enclose the member name and extension in quotation marks to preserve case.

Tip: The FILEEXT option will be ignored if you specify a file extension in the FILE or INFILE statement.

See Also: LOWCASE_MEMNAME option on page 1569

LOCALCACHE="*directory name*"
specifies a directory where a temporary subdirectory is created to hold local copies of the server files. Each fileref has its own unique subdirectory. If a directory is not specified, then the subdirectories are created in the SAS Work directory. SAS deletes the temporary files when the SAS program completes.

Default: SAS Work directory

LOCKDURATION=*n*
specifies the number of minutes that the files that are written through the WebDAV fileref are locked. SAS unlocks the files when the SAS program successfully finishes executing. If the SAS program fails, then the locks expire after the time allotted.

Default: 30 minutes

LOWCASE_MEMNAME
enables autocall macro retrieval of lowercase directory or member names from WebDAV servers.

Restriction: SAS autocall macro retrieval always searches for uppercase directory member names. Mixed-case directory or member names are not supported.

See Also: FILEEXT option on page 1568

LRECL=*lrecl*
where *lrecl* is the logical record length of the data.

Default: 256

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

MOD
Places the file in update mode and appends updates to the bottom of the file.

PASS='*password*'
where *password* is the password to use with the user name that is specified in the USER option. The password is case sensitive and it must be enclosed in single or double quotation marks.

Alias: PASSWORD=, PW=, PWD=

Tip: To use an encoded password, use the PWENCODE procedure in order to disguise the text string, and then enter the encoded password for the PASS= option. For more information, see "The PWENCODE Procedure" in the *Base SAS Procedures Guide*.

PROXY=*url*
specifies the Uniform Resource Locator (URL) for the proxy server in one of these forms:

http://*hostname*/

http://*hostname:port*/

RECFM=*recfm*
where *recfm* is one of two record formats:

S — is stream-record format. Data is transferred in image (binary) mode.

Tip: The amount of data that is read is controlled by the current LRECL value or the value of the NBYTE= variable in the INFILE statement. The NBYTE= option specifies a variable that is equal to the amount of data to be read. This amount must be less than or equal to LRECL.

See Also: The NBYTE= option on page 1598 in the INFILE statement.

V — is variable-record format (the default). In this format, records have varying lengths, and they are transferred in text (stream) mode.

Tip: Any record larger than LRECL is truncated.

Default: V

USER='*username*'
where *username* is used to log on to the URL server. The user ID is case sensitive and it must be enclosed in single or double quotation marks.

Alias: UID=

Details

When you access a WebDAV server to update a file, the file is pulled from the WebDAV server to your local disk storage for processing. When this processing is complete, the file is pushed back to the WebDAV server for storage. The file is removed from the local disk storage when it is pushed back.

The Secure Sockets Layer (SSL) protocol is used when the URL begins with "https" instead of "http". The SSL protocol provides network security and privacy. Developed by Netscape Communications, SSL uses encryption algorithms that include RC2, RC4, DES, tripleDES, IDEA, and MD5. Not limited to providing only encryption services, SSL can also perform client and server authentication and use message authentication codes. SSL is supported by both Netscape Navigator and Internet Explorer. Many Web sites use the protocol to provide confidential user information such as credit card numbers. The SSL protocol is application independent, which enables protocols such as HTTP, FTP, and Telnet to be layered transparently above it. SSL is optimized for HTTP.

Note: WebDAV servers have defined levels of permissions at both the directory and file level. The WebDAV access method honors those permissions. For example, if a file is available as read-only, the user will not be able to modify it. △

Operating Environment Information: Using the FILENAME statement requires information that is specific to your operating environment. The WebDAV access method is fully documented here, but for more information about how to specify filenames, see the SAS documentation for your operating environment. △

Examples

Example 1: Accessing a File at a Web Site This example accesses the file **`rawFile.txt`** at site **`www.mycompany.com`**.

```
filename foo webdav 'https://www.mycompany.com/production/files/rawFile.txt'
   user='wong' pass='jd75ld';

data _null_;
infile foo;
input a $80.;
run;
```

Example 2: Using a Proxy Server This example accesses the file **acctgfile.dat** by using the proxy server **otherwebsvr:80**.

```
filename foo webdav 'https://webserver.com/webdav/acctgfile.dat'
   user='sanchez' pass='239sk349exz'
   proxy='http://otherwebsvr.com:80';

data _null_;
infile foo;
input a $80.;
run;
```

Example 3: Writing to a New Member of a Directory This example writes the file **SHOES** to the directory **TESTING**.

```
filename writeit webdav
   "https://webserver.com:8443/webdav/testing/"
    dir user="webuser" pass=XXXXXXXXX;

    data _null_;
       file writeit(shoes);
       set sashelp.shoes;
       put region $25. product $14.;
    run;
```

Example 4: Reading from a Member of a Directory This example reads the file **SHOES** from the directory **TESTING1**.

```
filename readit webdav
   "https://webserver.com:8443/webdav/testing1/"
   dir user="webuser" pass=XXXXXXXXX;

   data shoes;
      length region $25  product $14;
      infile readit(shoes);
      input region $25. product $14.;
   run;
```

Example 5: Using a WebDAV Location as an Autocall Macro Library By default, the autocall macro facility expects uppercase filenames. This example accesses the file **MYTEST** in the autocall macro library **WRITEIT**.

```
filename writeit webdav
   "https://webserver.com/webdav/macrolib"
      dir fileext user="webuser" pass=XXXXXXXXX;
options SASAUTOS=(writeit);

/* expects a file called MYTEST.SAS */
%MYTEST;
```

Example 6: Accessing a Lowercased Autocall Macro Member The following example accesses the file **testmem.sas** in the autocall macro library **LIST**. The LOWCASE_MEMNAME option is used to access the file, which is in lowercase.

```
filename list webdav "https://t1234.na.fyi.com:8443/accounting/"
   dir fileext user="xxxxx" pass="xxxxx" LOWCASE_MEMNAME;
options sasautos=(list);
%testmem;
```

Example 7: Using a %INCLUDE Statement and Macro Invocation to Access a Lowercased Autocall Macro Member The following example accesses the file **testmem.sas** in the autocall macro library **MYTEST**. Because the file is accessed by using the %INCLUDE statement, case sensitivity is preserved.

```
filename mytest webdav "https://t1234.na.fyi.com:8443/payroll/"
   dir user="xxxxxx" pass="xxxxx";
%include mytest(testmem.sas) /source2;
%testmem;
```

If the filename was in uppercase, the reference to the filename in the %INCLUDE statement and macro call needs to be uppercase.

```
%include mytest(TESTMEM.SAS) /source2;
%TESTMEM;
```

Example 8: Accessing a File with a Mixed-Case Name The following example accesses the file **fileNOext** from the **production** directory. Because the file is quoted in the INFILE statement, case sensitivity is preserved and the file extension is ignored.

```
filename test webdav "https://t1234.na.fyi.com:8443/production"
     dir user="xxxxxx" pass="xxxxx";
data _null_;
   infile test('fileNOext');
   input;
   list;
run;
```

Example 9: Using the FILEEXT Option to Automatically Attach a File Extension The following example accesses the file **testmem.sas** from the **sales** directory. The FILEEXT option automatically adds **.DATA** as the file extension. The member name that is read is **testmem.DATA**.

```
filename listing webdav "https://t1234.na.fyi.com:8443/sales"
     dir fileext user="xxxxxx" pass="xxxxx";
data _null_;
   infile listing(testmem);
   input;
   list;
run;
```

See Also

Statements:

"FILENAME Statement" on page 1520

"FILENAME Statement, CATALOG Access Method" on page 1526

"FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532

"FILENAME Statement, FTP Access Method" on page 1542
"FILENAME Statement, SOCKET Access Method" on page 1559
"FILENAME Statement, URL Access Method" on page 1563
"LIBNAME Statement for WebDAV Server Access" on page 1665

FOOTNOTE Statement

Writes up to 10 lines of text at the bottom of the procedure or DATA step output.

Valid: anywhere
Category: Output Control
Requirement: You must specify the FOOTNOTE option if you use a FILE statement.
See: FOOTNOTE Statement in the documentation for your operating environment.

Syntax

FOOTNOTE<*n* > <*ods-format-options*> <'*text*' | "*text*" >;

Without Arguments

Using FOOTNOTE without arguments cancels all existing footnotes.

Arguments

n
specifies the relative line to be occupied by the footnote.

Tip: For footnotes, lines are pushed up from the bottom. The FOOTNOTE statement with the highest number appears on the bottom line.

Range: *n* can range from 1 to 10.

Default: If you omit *n*, SAS assumes a value of 1.

ods-format-options
specifies formatting options for the ODS HTML, RTF, and PRINTER(PDF) destinations.

BOLD
specifies that the footnote text is bold font weight.

ODS Destinations: HTML, RTF, PRINTER

COLOR=*color*
specifies the footnote text color.

Alias: C

ODS Destinations: HTML, RTF, PRINTER

Featured in: Example 3 on page 1783

BCOLOR=*color*
specifies the background color of the footnote block.

ODS Destinations: HTML, RTF, PRINTER

FONT=*font-face*
specifies the font to use. If you supply multiple fonts, then the destination device uses the first one that is installed on your system.

Alias: F

ODS Destinations: HTML, RTF, PRINTER

HEIGHT=*size*
specifies the point size.

Alias: H

ODS Destinations: HTML, RTF, PRINTER

Featured in: Example 3 on page 1783

ITALIC
specifies that the footnote text is in italic style.

ODS Destinations: HTML, RTF, PRINTER

JUSTIFY= CENTER | LEFT | RIGHT
specifies justification.

CENTER
specifies center justification.

Alias: C

LEFT
specifies left justification.

Alias: L

RIGHT
specifies right justification.

Alias: R

Alias: J

ODS Destinations: HTML, RTF, PRINTER

Featured in: Example 3 on page 1783

LINK='*url*'
specifies a hyperlink.

Tip: The visual properties for LINK= always come from the current style.

ODS Destinations: HTML, RTF, PRINTER

UNDERLIN= 0 | 1 | 2 | 3
specifies whether the subsequent text is underlined. 0 indicates no underlining. 1, 2, and 3 indicates underlining.

Alias: U

Tip: ODS generates the same type of underline for values 1, 2, and 3. However, SAS/GRAPH uses values 1, 2, and 3 to generate increasingly thicker underlines.

ODS Destinations: HTML, RTF, PRINTER

Note: The defaults for how ODS renders the FOOTNOTE statement come from style elements that relate to system footnotes in the current style. The FOOTNOTE statement syntax with *ods-format-options* is a way to override the settings that are provided by the current style.

The current style varies according to the ODS destination. For more information about how to determine the current style, see "What Are Style Definitions, Style Elements, and Style Attributes?" and "Concepts: Style Definitions and the TEMPLATE Procedure" in the *SAS Output Delivery System: User's Guide*. △

Tip: You can specify these options by letter, word, or words by preceding each letter or word of the *text* by the option.

For example, this code will make the footnote "Red, White, and Blue" appear in different colors.

```
footnote color=red "Red," color=white "White, and" color=blue "Blue";
```

'text' | *"text"*

specifies the text of the footnote in single or double quotation marks

Tip: For compatibility with previous releases, SAS accepts some text without quotation marks. When you write new programs or update existing programs, *always* enclose text in quotation marks.

Tip: If you use an automatic macro variable in the title text, you must enclose the title text in double quotation marks. The SAS macro facility will resolve the macro variable only if the text is in double quotation marks.

Tip: If you use single quotation marks (') or double quotation marks ("") together (with no space in between them) as the string of text, SAS will output a single quotation mark (') or double quotation mark ("), respectively.

Details

A FOOTNOTE statement takes effect when the step or RUN group with which it is associated executes. After you specify a footnote for a line, SAS repeats the same footnote on all pages until you cancel or redefine the footnote for that line. When a FOOTNOTE statement is specified for a given line, it cancels the previous FOOTNOTE statement for that line and for all footnote lines with higher numbers.

Operating Environment Information: The maximum footnote length that is allowed depends on the operating environment and the value of the LINESIZE= system option. Refer to the SAS documentation for your operating environment for more information. △

Comparisons

You can also create footnotes with the FOOTNOTES window. For more information, refer to the online Help for the window.

You can modify footnotes with the Output Delivery System. See Example 3 on page 1783.

Examples

These examples of a FOOTNOTE statement result in the same footnote:

□ `footnote8 "Managers' Meeting";`

□ `footnote8 'Managers'' Meeting';`

These are examples of FOOTNOTE statements that use some of the formatting options for the ODS HTML, RTF, and PRINTER(PDF) destinations. For the complete example, see Example 3 on page 1783.

```
footnote j=left height=20pt
       color=red "Prepared "
       c='#FF9900' "on";

footnote2 j=center color=blue
          height=24pt "&SYSDATE9";
footnote3 link='http://support.sas.com' "SAS";
```

See Also

Statement:

"TITLE Statement" on page 1779

"The TEMPLATE Procedure" in the *SAS Output Delivery System: User's Guide*

FORMAT Statement

Associates formats with variables.

Valid: in a DATA step or PROC step

Category: Information

Type: Declarative

Syntax

FORMAT *variable-1* <. . . *variable-n*> <*format*> <DEFAULT=*default-format*>;

FORMAT *variable-1* <. . . *variable-n*> *format* <DEFAULT=*default-format*>;

FORMAT *variable-1* <. . . *variable-n*> *format variable-1* <. . . *variable-n*> *format*;

Arguments

variable

names one or more variables for SAS to associate with a format. You must specify at least one *variable*.

Tip: To disassociate a format from a variable, use the variable in a FORMAT statement without specifying a format in a DATA step or in PROC DATASETS. In a DATA step, place this FORMAT statement after the SET statement. See Example 3 on page 1579. You can also use PROC DATASETS.

format

specifies the format that is listed for writing the values of the variables.

Tip: Formats that are associated with variables by using a FORMAT statement behave like formats that are used with a colon modifier in a subsequent PUT statement. For details on using a colon modifier, see "PUT Statement, List" on page 1731.

See also: "Formats by Category" on page 99

DEFAULT=*default-format*

specifies a temporary default format for displaying the values of variables that are not listed in the FORMAT statement. These default formats apply only to the current DATA step; they are not permanently associated with variables in the output data set.

A DEFAULT= format specification applies to

- variables that are not named in a FORMAT or ATTRIB statement
- variables that are not permanently associated with a format within a SAS data set

- □ variables that are not written with the explicit use of a format.

Default: If you omit DEFAULT=, SAS uses BEST*w*. as the default numeric format and $*w*. as the default character format.

Restriction: Use this option only in a DATA step.

Tip: A DEFAULT= specification can occur anywhere in a FORMAT statement. It can specify either a numeric default, a character default, or both.

Featured in: Example 1 on page 1577

Details

The FORMAT statement can use standard SAS formats or user-written formats that have been previously defined in PROC FORMAT. A single FORMAT statement can associate the same format with several variables, or it can associate different formats with different variables. If a variable appears in multiple FORMAT statements, SAS uses the format that is assigned last.

You use a FORMAT statement in the DATA step to permanently associate a format with a variable. SAS changes the descriptor information of the SAS data set that contains the variable. You can use a FORMAT statement in some PROC steps, but the rules are different. For more information, see *Base SAS Procedures Guide*.

Comparisons

Both the ATTRIB and FORMAT statements can associate formats with variables, and both statements can change the format that is associated with a variable. You can use the FORMAT statement in PROC DATASETS to change or remove the format that is associated with a variable. You can also associate, change, or disassociate formats and variables in existing SAS data sets through the windowing environment.

Examples

Example 1: Assigning Formats and Defaults This example uses a FORMAT statement to assign formats and default formats for numeric and character variables. The default formats are not associated with variables in the data set but affect how the PUT statement writes the variables in the current DATA step.

```
data tstfmt;
   format W $char3.
          Y 10.3
          default=8.2 $char8.;
   W='Good morning.';
   X=12.1;
   Y=13.2;
   Z='Howdy-doody';
   put W/X/Y/Z;
run;

proc contents data=tstfmt;
run;

proc print data=tstfmt;
run;
```

The following output shows a partial listing from PROC CONTENTS, as well as the report that PROC PRINT generates.

Output 6.5 Partial Listing from PROC CONTENTS and the PROC PRINT Report

```
                        The SAS System                          3

                     CONTENTS PROCEDURE

      -----Alphabetic List of Variables and Attributes-----

         #    Variable    Type    Len     Pos    Format
         ----------------------------------------------
         1    W           Char      3      16    $CHAR3.
         3    X           Num       8       8
         2    Y           Num       8       0    10.3
         4    Z           Char     11      19
```

Output 6.6 PROC PRINT Report

```
                        The SAS System                          4

        OBS    W               Y       X          Z

          1    Goo         13.200    12.1    Howdy-doody
```

The default formats apply to variables X and Z while the assigned formats apply to the variables W and Y.

The PUT statement produces this result:

```
----+----1----+----2
Goo
12.10
13.200
Howdy-do
```

Example 2: Associating Multiple Variables with a Single Format This example uses the FORMAT statement to assign a single format to multiple variables.

```
data report;
   input Item $ 1--6 Material $ 8--14 Investment 16--22 Profit 24--31;
   format Item Material $upcase9. Investment  Profit  dollar15.2;
   datalines;
shirts cotton  2256354 83952175
ties   silk    498678  2349615
suits  silk    9482146 69839563
belts  leather 7693    14893
shoes  leather 7936712 22964
;
run;
options pageno=1 nodate ls=80 ps=64;

proc print data=report;
   title 'Profit Summary: Kellam Manufacturing Company';
run;
```

Output 6.7 Results from Associating Multiple Variables with a Single Format

Profit Summary: Kellam Manufacturing Company 1

Obs	Item	Material	Investment	Profit
1	SHIRTS	COTTON	$2,256,354.00	$83,952,175.00
2	TIES	SILK	$498,678.00	$2,349,615.00
3	SUITS	SILK	$9,482,146.00	$69,839,563.00
4	BELTS	LEATHER	$7,693.00	$14,893.00
5	SHOES	LEATHER	$7,936,712.00	$22,964.00

Example 3: Removing a Format This example disassociates an existing format from a variable in a SAS data set. The order of the FORMAT and the SET statements is important.

```
data rtest;
   set rtest;
   format x;
run;
```

See Also

Statement:

"ATTRIB Statement" on page 1448

"The DATASETS Procedure" in *Base SAS Procedures Guide*

GO TO Statement

Directs program execution immediately to the statement label that is specified and, if followed by a RETURN statement, returns execution to the beginning of the DATA step.

Valid: in a DATA step

Category: Control

Type: Executable

Alias: GOTO

Syntax

GO TO *label*;

Arguments

label

specifies a statement label that identifies the GO TO destination. The destination must be within the same DATA step. You must specify the *label* argument.

Comparisons

The GO TO statement and the LINK statement are similar. However, a GO TO statement is often used without a RETURN statement, whereas a LINK statement is usually used with an explicit RETURN statement. The action of a subsequent RETURN statement differs between the GO TO and LINK statements. A RETURN statement after a LINK statement returns execution to the statement that follows the LINK statement. A RETURN after a GO TO statement returns execution to the beginning of the DATA step (unless a LINK statement precedes the GO TO statement. In that case, execution continues with the first statement after the LINK statement).

GO TO statements can often be replaced by DO-END and IF-THEN/ELSE programming logic.

Examples

Use the GO TO statement as shown here.

- In this example, if the condition is true, the GO TO statement instructs SAS to jump to a label called ADD and to continue execution from there. If the condition is false, SAS executes the PUT statement and the statement that is associated with the GO TO label:

```
data info;
   input x;
   if 1<=x<=5 then go to add;
   put x=;
   add: sumx+x;
   datalines;
7
6
323
;
```

  Because every DATA step contains an implied RETURN at the end of the step, program execution returns to the top of the step after the sum statement is executed. Therefore, an explicit RETURN statement at the bottom of the DATA step is not necessary.

- If you do not want the sum statement to execute for observations that do not meet the condition, rewrite the code and include an explicit return statement.

```
data info;
   input x;
   if 1<=x<=5 then go to add;
   put x=;
   return;
      /* SUM statement not executed */
      /* if x<1 or x>5              */
   add: sumx+x;
   datalines;
7
6
323
;
```

See Also

Statements:

"DO Statement" on page 1491
"Labels, Statement" on page 1651
"LINK Statement" on page 1669
"RETURN Statement" on page 1752

IF Statement, Subsetting

Continues processing only those observations that meet the condition of the specified expression.

Valid: in a DATA step
Category: Action
Type: Executable

Syntax

IF *expression*;

Arguments

expression
is any SAS expression.

Details

The subsetting IF statement causes the DATA step to continue processing only those raw data records or those observations from a SAS data set that meet the condition of the expression that is specified in the IF statement. That is, if the expression is true for the observation or record (its value is neither 0 nor missing), SAS continues to execute statements in the DATA step and includes the current observation in the data set. The resulting SAS data set or data sets contain a subset of the original external file or SAS data set.

If the expression is false (its value is 0 or missing), no further statements are processed for that observation or record, the current observation is not written to the data set, and the remaining program statements in the DATA step are not executed. SAS immediately returns to the beginning of the DATA step because the subsetting IF statement does not require additional statements to stop processing observations.

Comparisons

□ The subsetting IF statement is equivalent to this IF-THEN statement:

```
if not (expression)
   then delete;
```

- When you create SAS data sets, use the subsetting IF statement when it is easier to specify a condition for including observations. When it is easier to specify a condition for excluding observations, use the DELETE statement.
- The subsetting IF and the WHERE statements are not equivalent. The two statements work differently and produce different output data sets in some cases. The most important differences are summarized as follows:
  - The subsetting IF statement selects observations that have been read into the program data vector. The WHERE statement selects observations before they are brought into the program data vector. The subsetting IF might be less efficient than the WHERE statement because it must read each observation from the input data set into the program data vector.
  - The subsetting IF statement and WHERE statement can produce different results in DATA steps that interleave, merge, or update SAS data sets.
  - When the subsetting IF statement is used with the MERGE statement, the SAS System selects observations after the current observations are combined. When the WHERE statement is used with the MERGE statement, the SAS System applies the selection criteria to each input data set before combining the current observations.
  - The subsetting IF statement can select observations from an existing SAS data set or from raw data that are read with the INPUT statement. The WHERE statement can select observations only from existing SAS data sets.
  - The subsetting IF statement is executable; the WHERE statement is not.

Examples

- This example results in a data set that contains only those observations with the value **F** for the variable SEX:

```
if sex='F';
```

- This example results in a data set that contains all observations for which the value of the variable AGE is not missing or 0:

```
if age;
```

See Also

Data Set Options:

"WHERE= Data Set Option" on page 67

Statements:

"DELETE Statement" on page 1486

"IF-THEN/ELSE Statement" on page 1582

"WHERE Statement" on page 1792

"WHERE-Expression Processing" in *SAS Language Reference: Concepts*

IF-THEN/ELSE Statement

Executes a SAS statement for observations that meet specific conditions.

Valid: in a DATA step

Category: Control

Type: Executable

Syntax

IF *expression* **THEN** *statement*;
<**ELSE** *statement*;>

Arguments

expression
is any SAS expression and is a required argument.

statement
can be any executable SAS statement or DO group.

Details

SAS evaluates the expression in an IF-THEN statement to produce a result that is either non-zero, zero, or missing. A non-zero and nonmissing result causes the expression to be true; a result of zero or missing causes the expression to be false.

If the conditions that are specified in the IF clause are met, the IF-THEN statement executes a SAS statement for observations that are read from a SAS data set, for records in an external file, or for computed values. An optional ELSE statement gives an alternative action if the THEN clause is not executed. The ELSE statement, if used, must immediately follow the IF-THEN statement.

Using IF-THEN statements *without* the ELSE statement causes SAS to evaluate all IF-THEN statements. Using IF-THEN statements *with* the ELSE statement causes SAS to execute IF-THEN statements until it encounters the first true statement. Subsequent IF-THEN statements are not evaluated.

Note: For greater efficiency, construct your IF-THEN/ELSE statement with conditions of decreasing probability. △

Comparisons

- Use a SELECT group rather than a series of IF-THEN statements when you have a long series of mutually exclusive conditions.
- Use subsetting IF statements, without a THEN clause, to continue processing only those observations or records that meet the condition that is specified in the IF clause.

Examples

These examples show different ways of specifying the IF-THEN/ELSE statement.

```
if x then delete;
```

```
if status='OK' and type=3 then count+1;
```

```
if age ne agecheck then delete;
```

```
□ if x=0 then
     if y ne 0 then put 'X ZERO, Y NONZERO';
     else put 'X ZERO, Y ZERO';
  else put 'X NONZERO';

□ if answer=9 then
     do;
        answer=.;
        put 'INVALID ANSWER FOR ' id=;
     end;
  else
     do;
        answer=answer10;
        valid+1;
     end;

□ data region;
     input city $ 1-30;
     if city='New York City'
        or city='Miami' then
        region='ATLANTIC COAST';
     else if city='San Francisco'
        or city='Los Angeles' then
           region='PACIFIC COAST';
     datalines;
  ...more data lines...
  ;
```

See Also

Statements:

"DO Statement" on page 1491
"IF Statement, Subsetting" on page 1581
"SELECT Statement" on page 1760

%INCLUDE Statement

Brings a SAS programming statement, data lines, or both, into a current SAS program.

Valid: anywhere

Category: Program Control

Alias: %INC

See: %INCLUDE Statement in the documentation for your operating environment.

Syntax

%INCLUDE *source(s)*
</<SOURCE2> <S2=*length*> <*operating-environment-options*>>;

Arguments

source(s)
describes the location of the information that you want to access with the %INCLUDE statement. There are three possible sources:

Source	Definition
file-specification	specifies an external file
internal-lines	specifies lines that are entered earlier in the same SAS job or session
keyboard-entry	specifies statements or data lines that you enter directly from the keyboard

file-specification
identifies an entire external file that you want to bring into your program.

Restriction: You cannot selectively include lines from an external file.

Tip: Including external sources is useful in all types of SAS processing: batch, windowing, interactive line, and noninteractive.

Operating Environment Information: The character length allowed for filenames is operating environment specific. For complete details on specifying the physical names of external files, see the SAS documentation for your operating environment. △

File-specification can have these forms:

'external-file'
specifies the physical name of an external file that is enclosed in quotation marks. The physical name is the name by which the operating environment recognizes the file.

fileref
specifies a fileref that has previously been associated with an external file.

Tip: You can use a FILENAME statement or function or an operating environment command to make the association.

fileref (filename-1 <, "filename–2.xxx", ... filename-n>)
specifies a fileref that has previously been associated with an aggregate storage location. Follow the fileref with one or more filenames that reside in that location. Enclose the filenames in one set of parentheses, and separate each filename with a comma, space.

This example instructs SAS to include the files “testcode1.sas”, “testcode2.sas” and “testcode3.txt.” These files are located in aggregate storage location “mylib.” You do not need to specify the file extension for testcode1 and testcode2 because they are the default .SAS extension. You must enclose testcode3.txt in quotation marks with the whole filename specified because it has an extension other than .SAS:

```
%include mylib(testcode1, testcode2,
               "testcode3.txt");
```

Note: A file that is located in an aggregate storage location and has a name that is not a valid SAS name must have its name enclosed in quotation marks. △

Tip: You can use a FILENAME statement or function or an operating environment command to make the association.

Operating Environment Information: Different operating environments call an aggregate grouping of files by different names, such as a directory, a MACLIB, a text library, or a partitioned data set. For information about accessing files from a storage location that contains several files, see the SAS documentation for your operating environment. △

Tip: You can verify the existence of *file-specification* by using the SYSERR macro variable if the ERRORCHECK option is set to STRICT.

internal-lines
includes lines that are entered earlier in the same SAS job or session.
To include internal lines, use any of the following:

n includes line *n*.

n-m or *n:m* includes lines *n* through *m*.

Tip: Including internal lines is most useful in interactive line mode processing.

Tip: Use a %LIST statement to determine the line numbers that you want to include.

Tip: Although you can use the %INCLUDE statement to access previously submitted lines when you run SAS in a windowing environment, it might be more practical to recall lines in the Program Editor with the RECALL command and then submit the lines with the SUBMIT command.

Note: The SPOOL system option controls internal access to previously submitted lines when you run SAS in interactive line mode, noninteractive mode, and batch mode. By default, the SPOOL system option is set to NOSPOOL. The SPOOL system option must be in effect in order to use %INCLUDE statements with internal line references. Use the OPTIONS procedure to determine the current setting of the SPOOL system option on your system. △

keyboard-entry
is a method for preparing a program so that you can interrupt the current program’s execution, enter statements or data lines from the keyboard, and then resume program processing.

Tip: Use this method when you run SAS in noninteractive or interactive line mode. SAS pauses during processing and prompts you to enter statements from the keyboard.

Tip: Use this argument to include source from the keyboard:

* prompts you to enter data from the keyboard. Place an asterisk (*) after the %INCLUDE statement in your code:

```
proc print;
   %include *;
run;
```

To resume processing the original source program, enter a %RUN statement from the keyboard.

Restriction: The asterisk (*) cannot be used to specify keyboard entry if you use the Enhanced Editor in the Microsoft Windows operating environment.

Tip: You can use a %INCLUDE * statement in a batch job by creating a file with the fileref SASTERM that contains the statements that you would otherwise enter from the keyboard. The %INCLUDE * statement causes SAS to read from the file that is referenced by SASTERM. Insert a %RUN statement into the file that is referenced by SASTERM where you want SAS to resume reading from the original source.

Note: The fileref SASTERM must have been previously associated with an external file in a FILENAME statement or function or an operating environment command. △

SOURCE2

causes the SAS log to show the source statements that are being included in your SAS program.

Tip: The SAS log also displays the fileref and the filename of the source and the level of nesting (1, 2, 3, and so on).

Tip: The SAS system option SOURCE2 produces the same results. When you specify SOURCE2 in a %INCLUDE statement, it overrides the setting of the SOURCE2 system option for the duration of the include operation.

S2=*length*

specifies the length of the record to be used for input. *Length* can have these values:

S sets S2 equal to the current setting of the S= SAS system option.

0 tells SAS to use the setting of the SEQ= system option to determine whether the line contains a sequence field. If the line does contain a sequence field, SAS determines line length by excluding the sequence field from the total length.

n specifies a number greater than zero that corresponds to the length of the line to be read, when the file contains fixed-length records. When the file contains variable-length records, *n* specifies the column in which to begin reading data.

Tip: Text input from the %INCLUDE statement can be either fixed or variable length.

- Fixed-length records are either unsequenced or sequenced at the end of each record. For fixed-length records, the value given in S2= is the ending column of the data.
- Variable-length records are either unsequenced or sequenced at the beginning of each record. For variable-length records, the value given in S2= is the starting column of the data.

Interaction: The S2= system option also specifies the length of secondary source statements that are accessed by the %INCLUDE statement, and it is effective for the duration of your SAS session. The S2= option in the %INCLUDE statement

affects only the current include operation. If you use the option in the %INCLUDE statement, it overrides the system option setting for the duration of the include operation.

See Also: For a detailed discussion of fixed- and variable-length input records, see "S= System Option" on page 1987 and "S2= System Option" on page 1990.

operating-environment-options

Operating Environment Information: Operating environments can support various options for the %INCLUDE statement. See the documentation for your operating environment for a list of these options and their functions. △

Details

What %INCLUDE Does When you execute a program that contains the %INCLUDE statement, SAS executes your code, including any statements or data lines that you bring into the program with %INCLUDE.

Operating Environment Information: Use of the %INCLUDE statement is dependent on your operating environment. See the documentation for your operating environment for more information about additional software features and methods of referring to and accessing your files. See your documentation before you run the examples for this statement. △

Three Sources of Data The %INCLUDE statement accesses SAS statements and data lines from three possible sources:

- external files
- lines entered earlier in the same job or session
- lines entered from the keyboard.

When Useful The %INCLUDE statement is most often used when running SAS in interactive line mode, noninteractive mode, or batch mode. Although you can use the %INCLUDE statement when you run SAS using windows, it might be more practical to use the INCLUDE and RECALL commands to access data lines and program statements, and submit these lines again.

Rules for Using %INCLUDE

- You can specify any number of sources in a %INCLUDE statement, and you can mix the types of included sources. Note, however, that although it is possible to include information from multiple sources in one %INCLUDE statement, it might be easier to understand a program that uses separately coded %INCLUDE statements for each source.
- The %INCLUDE statement must begin at a statement boundary. That is, it must be the first statement in a SAS job or must immediately follow a semicolon ending another statement. A %INCLUDE statement cannot immediately follow a DATALINES, DATALINES4, CARDS, or CARDS4 statement (or PARMCARDS or PARMCARDS4 statement in procedures that use those statements). However, you can include data lines with the %INCLUDE statement using one of these methods:
  - Make the DATALINES, DATALINES4, or CARDS, CARDS4 statement the first line in the file that contains the data.
  - Place the DATALINES, DATALINES4, or CARDS, CARDS4 statement in one file, and the data lines in another file. Use both sources in a single %INCLUDE statement.

  The %INCLUDE statement can be nested within a file that has been accessed with %INCLUDE. The maximum number of nested %INCLUDE statements that

you can use depends on system-specific limitations of your operating environment (such as available memory or the number of files you can have open concurrently).

- Because %INCLUDE is a global statement and global statements are not executable, the %INCLUDE statement cannot be used in conditional logic.
- The maximum line length is 32K bytes.

Comparisons

The %INCLUDE statement executes statements immediately. The INCLUDE command brings the included lines into the Program Editor window but does not execute them. You must issue the SUBMIT command to execute those lines.

Examples

Example 1: Including an External File

- This example stores a portion of a program in a file and includes it in a program to be written later. This program is stored in a file named MYFILE:

```
data monthly;
   input x y month $;
   datalines;
1 1 January
2 2 February
3 3 March
4 4 April
;
```

This program includes an external file named MYFILE and submits the DATA step that it contains before the PROC PRINT step executes:

```
%include 'MYFILE';

proc print;
run;
```

- To reference a file by using a fileref rather than the actual filename, you can use the FILENAME statement (or a command recognized by your operating environment) to assign a fileref:

```
filename in1 'MYFILE';
```

You can later access MYFILE with the fileref IN1:

```
%inc in1;
```

- If you want to use many files that are stored in a directory, PDS, or MACLIB (or whatever your operating environment calls an aggregate storage location), you can assign the fileref to the larger storage unit and then specify the filename. For example, this FILENAME statement assigns the fileref STORAGE to an aggregate storage location:

```
filename storage
   'aggregate-storage-location';
```

You can later include a file using this statement:

```
%inc storage(MYFILE);
```

- You can also access several files or members from this storage location by listing them in parentheses after the fileref in a single %INCLUDE statement. Separate filenames with a comma or a blank space. The following %INCLUDE statement demonstrates this method:

  ```
  %inc storage(file-1,file-2,file-3);
  ```

  When the file does not have the default .SAS extension, you can access it using quotation marks around the complete filename listed inside the parentheses.

- ```
  %inc storage("file-1.txt","file-2.dat",
                "file-3.cat");
  ```

Example 2: Including Previously Submitted Lines This %INCLUDE statement causes SAS to process lines 1, 5, 9 through 12, and 13 through 16 as if you had entered them again from your keyboard:

```
%include 1 5 9-12 13:16;
```

Example 3: Including Input from the Keyboard The method shown in this example is valid only when you run SAS in noninteractive mode or interactive line mode.

Restriction: The asterisk (*) cannot be used to specify keyboard entry if you use the Enhanced Editor in the Microsoft Windows operating environment.

This example uses %INCLUDE to add a customized TITLE statement when PROC PRINT executes:

```
data report;
   infile file-specification;
   input month $ salesamt $;
run;

proc print;
   %include *;
run;
```

When this DATA step executes, %INCLUDE with the asterisk causes SAS to issue a prompt for statements that are entered at the keyboard. You can enter statements such as

```
where month= 'January';

title 'Data for month of January';
```

After you enter statements, you can use %RUN to resume processing by typing

```
%run;
```

The %RUN statement signals to SAS to leave keyboard-entry mode and resume reading and executing the remaining SAS statements from the original program.

Example 4: Using %INCLUDE with Several Entries in a Single Catalog This example submits the source code from three entries in the catalog MYLIB.INCLUDE. When no entry type is specified, the default is CATAMS.

```
filename dir catalog 'mylib.include';
%include dir(mem1);
%include dir(mem2);
%include dir(mem3);
```
```

See Also

Statements:

"%LIST Statement" on page 1672
"%RUN Statement" on page 1754

INFILE Statement

Specifies an external file to read with an INPUT statement.

Valid: in a DATA Step
Category: File-handling
Type: Executable
See: INFILE Statement in the documentation for your operating environment.

Syntax

INFILE *file-specification* <*device-type*> <*options*> <*operating-environment-options*>;

INFILE *DBMS-specifications;*

Arguments

file-specification
identifies the source of the input data records, which is an external file or instream data. *File-specification* can have these forms:

'*external-file*'
specifies the physical name of an external file. The physical name is the name that the operating environment uses to access the file.

fileref
specifies the fileref of an external file.

Requirement: You must have previously associated the fileref with an external file in a FILENAME statement, FILENAME function, or an appropriate operating environment command.

See: "FILENAME Statement" on page 1520

fileref(file)
specifies a fileref of an aggregate storage location and the name of a file or member, enclosed in parentheses, that resides in that location.

Requirement: A file that is located in an aggregate storage location and has a name that is not a valid SAS name must have its name enclosed in quotation marks.

Requirement: You must have previously associated the fileref with an external file in a FILENAME statement, a FILENAME function, or an appropriate operating environment command.

See: "FILENAME Statement" on page 1520

Operating Environment Information: Different operating environments call an aggregate grouping of files by different names, such as a directory, a MACLIB, or a partitioned data set. For details about how to specify external files, see the SAS documentation for your operating environment. △

CARDS | CARDS4
for a definition, see DATALINES.

Alias: DATALINES | DATALINES4

DATALINES | DATALINES4
specifies that the input data immediately follows the DATALINES or DATALINES4 statement in the DATA step. Using DATALINES allows you to use the INFILE statement options to control how the INPUT statement reads instream data lines.

Alias: CARDS | CARDS4

Featured in: Example 1 on page 1605

Tip: You can verify the existence of *file-specification* by using the SYSERR macro variable if the ERRORCHECK option is set to STRICT.

device-type
specifies the type of device or the access method that is used if the fileref points to an input or output device or location that is not a physical file:

DISK
specifies that the device is a disk drive.

Tip: When you assign a fileref to a file on disk, you are not required to specify DISK.

DUMMY
specifies that the output to the file is discarded.

Tip: Specifying DUMMY can be useful for testing.

GTERM
indicates that the output device type is a graphics device that will receive graphics data.

PIPE
specifies an unnamed pipe.

Note: Some operating environments do not support pipes. △

PLOTTER
specifies an unbuffered graphics output device.

PRINTER
specifies a printer or printer spool file.

TAPE
specifies a tape drive.

TEMP
creates a temporary file that exists only as long as the filename is assigned. The temporary file can be accessed only through the logical name and is available only while the logical name exists.

Restriction: Do not specify a physical pathname. If you do, SAS returns an error.

Tip: Files that are manipulated by the TEMP device can have the same attributes and behave identically to DISK files.

TERMINAL
specifies the user's terminal.

UPRINTER
specifies a Universal Printing printer definition name.

Tip: If you do not specify the printer name in the FILENAME statement, the PRINTERPATH options control which Universal Printer is used and the destination of the output.

Alias: DEVICE=

Requirement: *device-type* must appear immediately after the physical path. DEVICE=*device-type* can appear anywhere in the statement.

Operating Environment Information: Additional specifications might be required when you specify some devices. See the SAS documentation for your operating environment before specifying a value other than DISK. Values in addition to the ones listed here might be available in some operating environments. △

Options

BLKSIZE=*block-size*
specifies the block size of the input file.

Default: Dependent on the operating environment

Operating Environment Information: For details, see the SAS documentation for your operating environment. △

COLUMN=*variable*
names a variable that SAS uses to assign the current column location of the input pointer. Like automatic variables, the COLUMN= variable is not written to the data set.

Alias: COL=

See Also: LINE= on page 1596

Featured in: Example 8 on page 1610

DELIMITER= *delimiter(s)*
specifies an alternate delimiter (other than a blank) to be used for LIST input, where *delimiter(s)* is

'*list-of-delimiting-characters*'
specifies one or more characters to read as delimiters.

Requirement: Enclose the list of characters in quotation marks.

Featured in: Example 1 on page 1605

character-variable
specifies a character variable whose value becomes the delimiter.

Alias: DLM=

Default: blank space

Tip: The delimiter is case sensitive.

See: "Reading Delimited Data" on page 1602

See Also: DLMSTR=, DSD (delimiter-sensitive data) on page 1594

Featured in: Example 1 on page 1605

DLMSTR= *delimiter*
specifies a character string as an alternate delimiter (other than a blank) to be used for LIST input, where *delimiter* is

'*delimiting-string*'
specifies a character string to read as a delimiter.

Requirement: Enclose the string in quotation marks.

Featured in: Example 1 on page 1605

character-variable
specifies a character variable whose value becomes the delimiter.

Default: blank space

Interaction: If you specify more than one DLMSTR= option in the INFILE statement, the DLMSTR= option that is specified last will be used. If you specify both the DELIMITER= and DLMSTR= options, the option that is specified last will be used.

Interaction: If you specify RECFM=N, make sure that the LRECL is large enough to hold the largest input item. Otherwise, it might be possible for the delimiter to be split across the record boundary.

Tip: The delimiter is case sensitive. To make the delimiter case insensitive, use the DLMSOPT='I' option.

See: "Reading Delimited Data" on page 1602

See Also: DELIMITER= on page 1593, DLMSOPT= on page 1594, and DSD on page 1594

Featured in: Example 1 on page 1605

DLMSOPT= '*option(s)*'
specifies parsing options for the DLMSTR= option where *option(s)* can be the following:

I
specifies that case-insensitive comparisons will be done.

T
specifies that trailing blanks of the string delimiter will be removed.

Tip: The *T* option is useful when you use a variable as the delimiter string.

Tip: You can specify either *I*, *T*, or both.

Requirement: The DLMSOPT= option has an effect only when used with the DLMSTR= option.

See Also: DLMSTR= on page 1593

Featured in: Example 1 on page 1605

DSD (delimiter-sensitive data)
specifies that when data values are enclosed in quotation marks, delimiters within the value are treated as character data. The DSD option changes how SAS treats delimiters when you use LIST input and sets the default delimiter to a comma. When you specify DSD, SAS treats two consecutive delimiters as a missing value and removes quotation marks from character values.

Interaction: Use the DELIMITER= or DLMSTR= option to change the delimiter.

Tip: Use the DSD option and LIST input to read a character value that contains a delimiter within a string that is enclosed in quotation marks. The INPUT statement treats the delimiter as a valid character and removes the quotation marks from the character string before the value is stored. Use the tilde (~) format modifier to retain the quotation marks.

See: "Reading Delimited Data" on page 1602

See Also: DELIMITER= on page 1593, DLMSTR= on page 1593

Featured in: Example 1 on page 1605 and Example 2 on page 1607

ENCODING= '*encoding-value*'
specifies the encoding to use when reading from the external file. The value for ENCODING= indicates that the external file has a different encoding from the current session encoding.

When you read data from an external file, SAS transcodes the data from the specified encoding to the session encoding.

For valid encoding values, see "Encoding Values in SAS Language Elements" in the *SAS National Language Support (NLS): Reference Guide*.

Default: SAS assumes that an external file is in the same encoding as the session encoding.

Featured in: Example 11 on page 1613

END=*variable*
specifies a variable that SAS sets to 1 when the current input data record is the last in the input file. Until SAS processes the last data record, the END= variable is set to 0. Like automatic variables, this variable is not written to the data set.

Restriction: You cannot use the END= option with

- the UNBUFFERED option
- the DATALINES or DATALINES4 statement
- an INPUT statement that reads multiple input data records.

Tip: Use the option EOF= on page 1595 when END= is invalid.

Featured in: Example 5 on page 1608

EOF=*label*
specifies a statement label that is the object of an implicit GO TO when the INFILE statement reaches end of file. When an INPUT statement attempts to read from a file that has no more records, SAS moves execution to the statement label indicated.

Interaction: Use EOF= instead of the END= option with

- the UNBUFFERED option
- the DATALINES or DATALINES4 statement
- an INPUT statement that reads multiple input data records.

Tip: The EOF= option is useful when you read from multiple input files sequentially.

See Also: END= on page 1595, EOV= on page 1595, and UNBUFFERED on page 1600

EOV=*variable*
specifies a variable that SAS sets to 1 when the first record in a file in a series of concatenated files is read. The variable is set only after SAS encounters the next file. Like automatic variables, the EOV= variable is not written to the data set.

Tip: Reset the EOV= variable back to 0 after SAS encounters each boundary.

See Also: END= on page 1595 and EOF= on page 1595

EXPANDTABS | NOEXPANDTABS
specifies whether to expand tab characters to the standard tab setting, which is set at 8-column intervals that start at column 9.

Default: NOEXPANDTABS

Tip: EXPANDTABS is useful when you read data that contains the tab character that is native to your operating environment.

FILENAME=*variable*
specifies a variable that SAS sets to the physical name of the currently opened input file. Like automatic variables, the FILENAME= variable is not written to the data set.

Tip: Use a LENGTH statement to make the variable length long enough to contain the value of the filename.

See Also: FILEVAR= on page 1596

Featured in: Example 5 on page 1608

FILEVAR=*variable*
specifies a variable whose change in value causes the INFILE statement to close the current input file and open a new one. When the next INPUT statement executes, it reads from the new file that the FILEVAR= variable specifies. Like automatic variables, this variable is not written to the data set.

Restriction: The FILEVAR= variable must contain a character string that is a physical filename.

Interaction: When you use the FILEVAR= option, the *file-specification* is just a placeholder, not an actual filename or a fileref that has been previously assigned to a file. SAS uses this placeholder for reporting processing information to the SAS log. It must conform to the same rules as a fileref.

Tip: Use FILEVAR= to dynamically change the currently opened input file to a new physical file.

See Also: "Updating External Files in Place" on page 1601

Featured in: Example 5 on page 1608

FIRSTOBS=*record-number*
specifies a record number that SAS uses to begin reading input data records in the input file.

Default: 1

Tip: Use FIRSTOBS= with OBS= to read a range of records from the middle of a file.

Example: This statement processes record 50 through record 100:

```
infile file-specification firstobs=50 obs=100;
```

FLOWOVER
causes an INPUT statement to continue to read the next input data record if it does not find values in the current input line for all the variables in the statement. FLOWOVER is the default behavior of the INPUT statement.

See: "Reading Past the End of a Line" on page 1604

See Also: MISSOVER on page 1597, STOPOVER on page 1599, and TRUNCOVER on page 1599

LENGTH=*variable*
specifies a variable that SAS sets to the length of the current input line. SAS does not assign the variable a value until an INPUT statement executes. Like automatic variables, the LENGTH= variable is not written to the data set.

Tip: This option in conjunction with the $VARYING informat on page 1297 is useful when the field width varies.

Featured in: Example 4 on page 1608 and Example 7 on page 1610

LINE=*variable*
specifies a variable that SAS sets to the line location of the input pointer in the input buffer. Like automatic variables, the LINE= variable is not written to the data set.

Range: 1 to the value of the N= option

Interaction: The value of the LINE= variable is the current relative line number within the group of lines that is specified by the N= option or by the #*n* line pointer control in the INPUT statement.

See Also: COLUMN= on page 1593 and N= on page 1597

Featured in: Example 8 on page 1610

LINESIZE=*line-size*
specifies the record length that is available to the INPUT statement.

Operating Environment Information: Values for *line-size* are dependent on the operating environment record size. For details, see the SAS documentation for your operating environment. △

Alias: LS=

Range: up to 32767

Interaction: If an INPUT statement attempts to read past the column that is specified by the LINESIZE= option, then the action that is taken depends on whether the FLOWOVER, MISSOVER, SCANOVER, STOPOVER, or TRUNCOVER option is in effect. FLOWOVER is the default.

Tip: Use LINESIZE= to limit the record length when you do not want to read the entire record.

Example: If your data lines contain a sequence number in columns 73 through 80, then use this INFILE statement to restrict the INPUT statement to the first 72 columns:

```
infile file-specification linesize=72;
```

LRECL=*logical-record-length*
specifies the logical record length.

Operating Environment Information: Values for *logical-record-length* are dependent on the operating environment. For details, see the SAS documentation for your operating environment. △

Default: Dependent on the file characteristics of your operating environment

Restriction: LRECL is not valid when you use the DATALINES file specification.

Interaction: Alternatively, you can specify a global logical record length by using the LRECL= system option"LRECL= System Option" on page 1941.

Tip: LRECL= specifies the physical line length of the file. LINESIZE= tells the INPUT statement how much of the line to read.

MISSOVER
prevents an INPUT statement from reading a new input data record if it does not find values in the current input line for all the variables in the statement. When an INPUT statement reaches the end of the current input data record, variables without any values assigned are set to missing.

Tip: Use MISSOVER if the last field or fields might be missing and you want SAS to assign missing values to the corresponding variable.

See: "Reading Past the End of a Line" on page 1604

See Also: FLOWOVER on page 1596, SCANOVER on page 1599, STOPOVER on page 1599, and TRUNCOVER on page 1599

Featured in: Example 2 on page 1607

N=*available-lines*
specifies the number of lines that are available to the input pointer at one time.

Default: The highest value following a # pointer control in any INPUT statement in the DATA step. If you omit a # pointer control, then the default value is 1.

Interaction: This option affects only the number of lines that the pointer can access at a time; it has no effect on the number of lines an INPUT statement reads.

Tip: When you use # pointer controls in an INPUT statement that are less than the value of N=, you might get unexpected results. To prevent unexpected results, include a # pointer control that equals the value of the N= option. Here is an example:

```
infile 'external file' n=5;
input #2 name : $25. #3 job : $25. #5;
```

The INPUT statement includes a **#5** pointer control, even though no data is read from that record.

Featured in: Example 8 on page 1610

NBYTE=*variable*
specifies the name of a variable that contains the number of bytes to read from a file when you are reading data in stream record format (RECFM=S in the FILENAME statement).

Default: The LRECL value of the file

Interaction: If the number of bytes to read is set to -1, then the FTP and SOCKET access methods return the number of bytes that are currently available in the input buffer.

See: The RECFM= option on page 1560 in the FILENAME statement, SOCKET access method, and the RECFM= option on page 1547 in the FILENAME statement, FTP access method

OBS=*record-number* | MAX

record-number	specifies the record number of the last record to read in an input file that is read sequentially.
MAX	specifies the maximum number of observations to process, which will be at least as large as the largest signed, 32–bit integer. The absolute maximum depends on your host operating environment.

Default: MAX

Tip: Use OBS= with FIRSTOBS= to read a range of records from the middle of a file.

Example: This statement processes only the first 100 records in the file:

```
infile file-specification obs=100;
```

PAD | NOPAD
controls whether SAS pads the records that are read from an external file with blanks to the length that is specified in the LRECL= option.

Default: NOPAD

See Also: LRECL= on page 1597

PRINT | NOPRINT
specifies whether the input file contains carriage-control characters.

Tip: To read a file in a DATA step without having to remove the carriage-control characters, specify PRINT. To read the carriage-control characters as data values, specify NOPRINT.

RECFM=*record-format*
specifies the record format of the input file.

Operating Environment Information: Values for *record-format* are dependent on the operating environment. For details, see the SAS documentation for your operating environment. △

SCANOVER
causes the INPUT statement to scan the input data records until the character string that is specified in the @'*character-string*' expression is found.

Interaction: The MISSOVER, TRUNCOVER, and STOPOVER options change how the INPUT statement behaves when it scans for the @'*character-string*' expression and reaches the end of the record. By default (FLOWOVER option), the INPUT statement scans the next record while these other options cause scanning to stop.

Tip: It is redundant to specify both SCANOVER and FLOWOVER.

See: "Reading Past the End of a Line" on page 1604

See Also: FLOWOVER on page 1596, MISSOVER on page 1597, STOPOVER on page 1599, and TRUNCOVER on page 1599

Featured in: Example 3 on page 1607

SHAREBUFFERS
specifies that the FILE statement and the INFILE statement share the same buffer.

CAUTION:
When using SHAREBUFFERS, RECFM=V, and _INFILE_, use caution if you read a record with one length and update the file with a record of a different length. The length of the record can change by modifying _INFILE_. One option to avoid this potential problem is to pad or truncate _INFILE_ so that the original record length is maintained. △

Alias: SHAREBUFS

Tip: Use SHAREBUFFERS with the INFILE, FILE, and PUT statements to update an external file in place. Updating an external file in place saves CPU time because the PUT statement output is written straight from the input buffer instead of the output buffer.

Tip: Use SHAREBUFFERS to update specific fields in an external file instead of an entire record.

Featured in: Example 6 on page 1609

START=*variable*
specifies a variable whose value SAS uses as the first column number of the record that the PUT _INFILE_ statement writes. Like automatic variables, the START variable is not written to the data set.

See Also: _INFILE_ option in the PUT statement

STOPOVER
causes the DATA step to stop processing if an INPUT statement reaches the end of the current record without finding values for all variables in the statement. When an input line does not contain the expected number of values, SAS sets _ERROR_ to 1, stops building the data set as if a STOP statement has executed, and prints the incomplete data line.

Tip: Use FLOWOVER to reset the default behavior.

See: "Reading Past the End of a Line" on page 1604

See Also: FLOWOVER on page 1596, MISSOVER on page 1597, SCANOVER on page 1599, and TRUNCOVER on page 1599

Featured in: Example 2 on page 1607

TRUNCOVER
overrides the default behavior of the INPUT statement when an input data record is shorter than the INPUT statement expects. By default, the INPUT statement

automatically reads the next input data record. TRUNCOVER enables you to read variable-length records when some records are shorter than the INPUT statement expects. Variables without any values assigned are set to missing.

Tip: Use TRUNCOVER to assign the contents of the input buffer to a variable when the field is shorter than expected.

See: "Reading Past the End of a Line" on page 1604

See Also: FLOWOVER on page 1596, MISSOVER on page 1597, SCANOVER on page 1599, and STOPOVER on page 1599

Featured in: Example 3 on page 1607

UNBUFFERED
tells SAS not to perform a buffered ("look ahead") read.

Alias: UNBUF

Interaction: When you use UNBUFFERED, SAS never sets the END= variable to 1.

Tip: When you read instream data with a DATALINES statement, UNBUFFERED is in effect.

INFILE=*variable*
specifies a character variable that references the contents of the current input buffer for this INFILE statement. You can use the variable in the same way as any other variable, even as the target of an assignment. The variable is automatically retained and initialized to blanks. Like automatic variables, the _INFILE_= variable is not written to the data set.

Restriction: *variable* cannot be a previously defined variable. Ensure that the _INFILE_= specification is the first occurrence of this variable in the DATA step. Do not set or change the length of _INFILE_= variable with the LENGTH or ATTRIB statements. However, you can attach a format to this variable with the ATTRIB or FORMAT statement.

Interaction: The maximum length of this character variable is the logical record length for the specified INFILE statement. However, SAS does not open the file to know the LRECL= until before the execution phase. Therefore, the designated size for this variable during the compilation phase is 32,767.

Tip: Modification of this variable directly modifies the INFILE statement's current input buffer. Any PUT _INFILE_ (when this INFILE is current) that follows the buffer modification reflects the modified buffer contents. The _INFILE_= variable accesses only the current input buffer of the specified INFILE statement even if you use the N= option to specify multiple buffers.

Tip: To access the contents of the input buffer in another statement without using the _INFILE_= option, use the automatic variable _INFILE_.

Tip: The _INFILE_ variable does not have a fixed width. When you assign a value to the _INFILE_ variable, the length of the variable changes to the length of the value that is assigned.

Main Discussion: "Accessing the Contents of the Input Buffer" on page 1601

Featured in: Example 9 on page 1610 and Example 10 on page 1612

Operating Environment Options

Operating Environment Information: For descriptions of operating environment-specific options in the INFILE statement, see the SAS documentation for your operating environment. △

DBMS Specifications

DBMS-Specifications
enable you to read records from some DBMS files. You must license SAS/ACCESS software to be able to read from DBMS files. See the SAS/ACCESS documentation for the DBMS that you use.

Details

Operating Environment Information: The INFILE statement contains operating environment-specific material. See the SAS documentation for your operating environment before using this statement. △

How to Use the INFILE Statement Because the INFILE statement identifies the file to read, it must execute before the INPUT statement that reads the input data records. You can use the INFILE statement in conditional processing, such as an IF-THEN statement, because it is executable. The INFILE statement enables you to control the source of the input data records.

Usually, you use an INFILE statement to read data from an external file. When data is read from the job stream, you must use a DATALINES statement. However, to take advantage of certain data-reading options that are available only in the INFILE statement, you can use an INFILE statement with the file-specification DATALINES and a DATALINES statement in the same DATA step. See "Reading Long Instream Data Records" on page 1603 for more information.

When you use more than one INFILE statement for the same file specification and you use options in each INFILE statement, the effect is additive. To avoid confusion, use all the options in the first INFILE statement for a given external file.

Reading Multiple Input Files You can read from multiple input files in a single iteration of the DATA step in one of two ways:

- to keep multiple files open and change which file is read, use multiple INFILE statements.
- to dynamically change the current input file within a single DATA step, use the FILEVAR= option in an INFILE statement. The FILEVAR= option enables you to read from one file, close it, and then open another. See Example 5 on page 1608.

Updating External Files in Place You can use the INFILE statement in combination with the FILE statement to update records in an external file. Follow these steps:

1. Specify the INFILE statement before the FILE statement.
2. Specify the same fileref or physical filename in each statement.
3. Use options that are common to both the INFILE and FILE statements in the INFILE statement instead of the FILE statement. (Any such options that are used in the FILE statement are ignored.)

See Example 6 on page 1609.

To update individual fields within a record instead of the entire record, see the term SHAREBUFFERS under "Arguments" on page 1591.

Accessing the Contents of the Input Buffer In addition to the _INFILE_= variable, you can use the automatic _INFILE_ variable to reference the contents of the current input buffer for the most recent execution of the INFILE statement. This character variable is automatically retained and initialized to blanks. Like other automatic variables, _INFILE_ is not written to the data set.

When you specify the _INFILE_= option in an INFILE statement, then this variable is also indirectly referenced by the automatic _INFILE_ variable. If the automatic

INFILE variable is present and you omit _INFILE_= in a particular INFILE statement, then SAS creates an internal _INFILE_= variable for that INFILE statement. Otherwise, SAS does not create the _INFILE_= variable for a particular FILE.

During execution and at the point of reference, the maximum length of this character variable is the maximum length of the current _INFILE_= variable. However, because _INFILE_ merely references other variables whose lengths are not known until before the execution phase, the designated length is 32,767 during the compilation phase. For example, if you assign _INFILE_ to a new variable whose length is undefined, then the default length of the new variable is 32,767. You cannot use the LENGTH statement and the ATTRIB statement to set or override the length of _INFILE_. You can use the FORMAT statement and the ATTRIB statement to assign a format to _INFILE_.

Like other SAS variables, you can update the _INFILE_ variable in an assignment statement. You can also use a format with _INFILE_ in a PUT statement. For example, the following PUT statement writes the contents of the input buffer by using a hexadecimal format.

```
put _infile_ $hex100.;
```

Any modification of the _INFILE_ directly modifies the current input buffer for the current INFILE statement. The execution of any PUT _INFILE_ statement that follows this buffer modification will reflect the contents of the modified buffer.

INFILE only accesses the contents of the current input buffer for an INFILE statement, even when you use the N= option to specify multiple buffers. You can access all the N= buffers, but you must use an INPUT statement with the # line pointer control to make the desired buffer the current input buffer.

Reading Delimited Data By default, the delimiter that is used to read input data records with list input is a blank space. The delimiter-sensitive data (DSD) option, the DELIMITER= option, the DLMSTR= option, and the DLMSOPT= option affect how list input handles delimiters. The DELIMITER= or DLMSTR= option specifies that the INPUT statement use a character other than a blank as a delimiter for data values that are read with list input. When the DSD option is in effect, the INPUT statement uses a comma as the default delimiter.

To read a value as missing between two consecutive delimiters, use the DSD option. By default, the INPUT statement treats consecutive delimiters as a unit. When you use DSD, the INPUT statement treats consecutive delimiters separately. Therefore, a value that is missing between consecutive delimiters is read as a missing value. To change the delimiter from a comma to another value, use the DELIMITER= or DLMSTR= option.

For example, this DATA step program uses list input to read data that is separated with commas. The second data line contains a missing value. Because SAS allows consecutive delimiters with list input, the INPUT statement cannot detect the missing value.

```
data scores;
   infile datalines delimiter=',';
   input test1 test2 test3;
   datalines;
91,87,95
97,,92
,1,1
;
```

With the FLOWOVER option in effect, the data set SCORES contains two, not three, observations. The second observation is built incorrectly:

OBS	TEST1	TEST2	TEST3
1	91	87	95
2	97	92	1

To correct the problem, use the DSD option in the INFILE statement.

```
infile datalines dsd;
```

Now the INPUT statement detects the two consecutive delimiters and therefore assigns a missing value to variable TEST2 in the second observation.

OBS	TEST1	TEST2	TEST3
1	91	87	95
2	97	.	92
3	.	1	1

The DSD option also enables list input to read a character value that contains a delimiter within a quoted string. For example, if data is separated with commas, DSD enables you to place the character string in quotation marks and read a comma as a valid character. SAS does not store the quotation marks as part of the character value. To retain the quotation marks as part of the value, use the tilde (~) format modifier in an INPUT statement. See Example 1 on page 1605.

Note: Any time a text file originates from anywhere other than the local encoding environment, it might be necessary to specify the ENCODING= option on either EBCDIC or ASCII environments.

For example, when you read an EBCDIC text file on an ASCII platform, it is recommended that you specify the ENCODING= option in the INFILE statement. However, if you use the DSD and DLM options in the INFILE statement, the ENCODING= option is a requirement because these options require certain characters in the session encoding (such as quotation marks, commas, and blanks).

The use of encoding-specific informats should be reserved for use with true binary files. That is, files that contain both character and noncharacter fields. △

Reading Long Instream Data Records You can use the INFILE statement with the DATALINES file specification to process instream data. An INPUT statement reads the data records that follow the DATALINES statement. If you use the CARDIMAGE system option, or if this option is the default for your system, then SAS processes the data lines exactly like 80-byte punched card images that are padded with blanks. The default FLOWOVER option in the INFILE statement causes the INPUT statement to read the next record if it does not find values in the current record for all of the variables in the statement. To ensure that your data is processed correctly, use an external file for input when record lengths are greater than 80 bytes.

Note: The NOCARDIMAGE system option (see “CARDIMAGE System Option” on page 1859) specifies that data lines not be treated as if they were 80-byte card images.

The end of a data line is always treated as the end of the last token, except for strings that are enclosed in quotation marks. △

Reading Past the End of a Line By default, if the INPUT statement tries to read past the end of the current input data record, then it moves the input pointer to column 1 of the next record to read the remaining values. This default behavior is specified by the FLOWOVER option. A message is written to the SAS log:

```
NOTE: SAS went to a new line when INPUT
statement reached past the end of a line.
```

Several options are available to change the INPUT statement behavior when an end of line is reached. The STOPOVER option treats this condition as an error and stops building the data set. The MISSOVERand TRUNCOVER options do not allow the input pointer to go to the next record when the current INPUT statement is not satisfied. The SCANOVER option, used with @'character-string' scans the input record until it finds the specified *character-string*. The FLOWOVER option restores the default behavior.

The TRUNCOVER and MISSOVER options are similar. The MISSOVER option causes the INPUT statement to set a value to missing if the statement is unable to read an entire field because the value is shorter than the field length that is specified in the INPUT statement. The TRUNCOVER option writes whatever characters are read to the appropriate variable.

For example, an external file with variable-length records contains these records:

```
----+----1----+----2
1
22
333
4444
55555
```

The following DATA step reads this data to create a SAS data set. Only one of the input records is as long as the informatted length of the variable TESTNUM.

```
data numbers;
   infile 'external-file';
   input testnum 5.;
run;
```

This DATA step creates the three observations from the five input records because by default the FLOWOVER option is used to read the input records.

If you use the MISSOVER option in the INFILE statement, then the DATA step creates five observations. All the values that were read from records that were too short are set to missing. Use the TRUNCOVER option in the INFILE statement if you prefer to see what values were present in records that were too short to satisfy the current INPUT statement.

```
infile 'external-file' truncover;
```

The DATA step now reads the same input records and creates five observations. See Table 6.5 on page 1604 to compare the SAS data sets.

Table 6.5 The Value of TESTNUM Using Different INFILE Statement Options

OBS	FLOWOVER	MISSOVER	TRUNCOVER
1	22	.	1
2	4444	.	22

OBS	FLOWOVER	MISSOVER	TRUNCOVER
3	55555	.	333
4		.	4444
5		55555	55555

Comparisons

- The INFILE statement specifies the *input file* for any INPUT statements in the DATA step. The FILE statement specifies the *output file* for any PUT statements in the DATA step.
- An INFILE statement usually identifies data from an external file. A DATALINES statement indicates that data follows in the job stream. You can use the INFILE statement with the file specification DATALINES to take advantage of certain data-reading options that affect how the INPUT statement reads instream data.

Examples

Example 1: Changing How Delimiters Are Treated By default, the INPUT statement uses a blank as the delimiter. This DATA step uses a comma as the delimiter:

```
data num;
   infile datalines dsd;
   input x y z;
   datalines;
,2,3
4,5,6
7,8,9
;
```

The argument DATALINES in the INFILE statement allows you to use an INFILE statement option to read instream data lines. The DSD option sets the comma as the default delimiter. Because a comma precedes the first value in the first data line, a missing value is assigned to variable X in the first observation, and the value **2** is assigned to variable Y.

If the data uses multiple delimiters or a single delimiter other than a comma, then simply specify the delimiter values with the DELIMITER= option. In this example, the characters **a** and **b** function as delimiters:

```
 data nums;
    infile datalines dsd delimiter='ab';
    input X Y Z;
    datalines;
1aa2ab3
4b5bab6
7a8b9
;
```

The output that PROC PRINT generates shows the resulting NUM data set. Values are missing for variables in the first and second observations because DSD causes list input to detect two consecutive delimiters. If you omit DSD, the characters a, b, aa, ab, ba, or bb function as the delimiter and no variables are assigned missing values.

Output 6.8 The NUM Data Set

```
                          The SAS System                          1

                        OBS    X    Y    Z

                         1     1    .    2
                         2     4    5    .
                         3     7    8    9
```

If you want to use a string as the delimiter, specify the delimiter values with the DLMSTR= option. In this example, the string **PRD** is used as the delimiter. Note that the string contains uppercase characters. By using the DLMSOPT= option, **PRD**, **Prd**, **PRd**, **PrD**, **pRd**, **pRD**, **prD**, and **prd** are all valid delimiters.

```
 data test;
    infile datalines dsd dlmstr='PRD' dlmsopt='i';
    input X Y Z;
    datalines;
1PRD2PRd3
4PrD5Prd6
7pRd8pRD9
;
```

The output from PROC PRINT shows all the observations in the TEST data set.

Output 6.9 The TEST Data Set

```
                          The SAS System                          1

                        Obs    X    Y    Z

                         1     1    2    3
                         2     4    5    6
                         3     7    8    9
```

This DATA step uses modified list input and the DSD option to read data that is separated by commas and that might contain commas as part of a character value:

```
data scores;
   infile datalines dsd;
   input Name : $9. Score
         Team  : $25. Div $;
   datalines;
Joseph,76,"Red Racers, Washington",AAA
Mitchel,82,"Blue Bunnies, Richmond",AAA
Sue Ellen,74,"Green Gazelles, Atlanta",AA
;
```

The output that PROC PRINT generates shows the resulting SCORES data set. The delimiter (comma) is stored as part of the value of TEAM while the quotation marks are not.

Output 6.10 The SCORES Data Set

```
                          The SAS System                          1

 OBS    NAME          SCORE             TEAM               DIV

  1     Joseph          76     Red Racers, Washington     AAA
  2     Mitchel         82     Blue Bunnies, Richmond     AAA
  3     Sue Ellen       74     Green Gazelles, Atlanta    AA
```

Example 2: Handling Missing Values and Short Records with List Input This example demonstrates how to prevent missing values from causing problems when you read the data with list input. Some data lines in this example contain fewer than five temperature values. Use the MISSOVER option so that these values are set to missing.

```
data weather;
   infile datalines missover;
   input temp1-temp5;
   datalines;
97.9 98.1 98.3
98.6 99.2 99.1 98.5 97.5
96.2 97.3 98.3 97.6 96.5
;
```

SAS reads the three values on the first data line as the values of TEMP1, TEMP2, and TEMP3. The MISSOVER option causes SAS to set the values of TEMP4 and TEMP5 to missing for the first observation because no values for those variables are in the current input data record.

When you omit the MISSOVER option or use FLOWOVER, SAS moves the input pointer to line 2 and reads values for TEMP4 and TEMP5. The next time the DATA step executes, SAS reads a new line which, in this case, is line 3. This message appears in the SAS log:

```
NOTE: SAS went to a new line when INPUT statement
      reached past the end of a line.
```

You can also use the STOPOVER option in the INFILE statement. Using the STOPOVER option causes the DATA step to halt execution when an INPUT statement does not find enough values in a record of raw data:

```
 infile datalines stopover;
```

Because SAS does not find a TEMP4 value in the first data record, it sets _ERROR_ to 1, stops building the data set, and prints data line 1.

Example 3: Scanning Variable-Length Records for a Specific Character String This example shows how to use TRUNCOVER in combination with SCANOVER to pull phone numbers from a phone book. The phone number is always preceded by the word "phone:". Because the phone numbers include international numbers, the maximum length is 32 characters.

```
filename phonebk host-specific-path;
data _null_;
  file phonebk;
  input line $80.;
  put line;
  datalines;
    Jenny's Phone Book
```

```
    Jim Johanson phone: 619-555-9340
       Jim wants a scarf for the holidays.
    Jane Jovalley phone: (213) 555-4820
       Jane started growing cabbage in her garden.
       Her dog's name is Juniper.
    J.R. Hauptman phone: (49)12 34-56 78-90
       J.R. is my brother.
   ;
run;
```

Use @'phone:' to scan the lines of the file for a phone number and position the file pointer where the phone number begins. Use TRUNCOVER in combination with SCANOVER to skip the lines that do not contain 'phone:' and write only the phone numbers to the log.

```
data _null_;
   infile phonebk truncover scanover;
   input @'phone:' phone $32.;
   put phone=;
run;
```

The program writes the following lines to the SAS log:

```
phone=619-555-9340
phone=(213) 555-4820
phone=(49)12 34-56 78-90
```

Example 4: Reading Files That Contain Variable-Length Records

This example shows how to use LENGTH=, in combination with the $VARYING. informat, to read a file that contains variable-length records:

```
data a;
   infile file-specification length=linelen lrecl=510 pad;
   input firstvar 1-10 @;  /* assign LINELEN   */
   varlen=linelen-10;      /* Calculate VARLEN */
   input @11 secondvar $varying500. varlen;
run;
```

The following occurs in this DATA step:

- The INFILE statement creates the variable LINELEN but does not assign it a value.
- When the first INPUT statement executes, SAS determines the line length of the record and assigns that value to the variable LINELEN. The single trailing @ holds the record in the input buffer for the next INPUT statement.
- The assignment statement uses the two known lengths (the length of FIRSTVAR and the length of the entire record) to determine the length of VARLEN.
- The second INPUT statement uses the VARLEN value with the informat $VARYING500. to read the variable SECONDVAR.

See "$VARYING*w*. Informat" on page 1296 for more information.

Example 5: Reading from Multiple Input Files The following DATA step reads from two input files during each iteration of the DATA step. As SAS switches from one file to the next, each file remains open. The input pointer remains in place to begin reading from that location the next time an INPUT statement reads from that file.

```
data qtrtot(drop=jansale febsale marsale
                 aprsale maysale junsale);
```

```
     /* identify location of 1st file */
   infile file-specification-1;
     /* read values from 1st file     */
   input name $ jansale febsale marsale;
   qtr1tot=sum(jansale,febsale,marsale);

     /* identify location of 2nd file */
   infile file-specification-2;
     /* read values from 2nd file     */
   input @7 aprsale maysale junsale;
   qtr2tot=sum(aprsale,maysale,junsale);
run;
```

The DATA step terminates when SAS reaches an end of file on the shortest input file.

This DATA step uses FILEVAR= to read from a different file during each iteration of the DATA step:

```
data allsales;
   length fileloc myinfile $ 300;
   input fileloc $ ; /* read instream data       */

  /* The INFILE statement closes the current file
     and opens a new one if FILELOC changes value
     when INFILE executes                        */
   infile file-specification filevar=fileloc
          filename=myinfile end=done;

  /* DONE set to 1 when last input record read  */
   do while(not done);
  /* Read all input records from the currently  */
  /* opened input file, write to ALLSALES       */
     input name $ jansale febsale marsale;
     output;
   end;
   put 'Finished reading ' myinfile=;
   datalines;
external-file-1
external-file-2
external-file-3
;
```

The FILENAME= option assigns the name of the current input file to the variable MYINFILE. The LENGTH statement ensures that the FILENAME= variable and FILEVAR= variable have a length that is long enough to contain the value of the filename. The PUT statement prints the physical name of the currently open input file to the SAS log.

Example 6: Updating an External File This example shows how to use the INFILE statement with the SHAREBUFFERS option and the INPUT, FILE, and PUT statements to update an external file in place:

```
data _null_;
     /* The INFILE and FILE statements     */
     /* must specify the same file.       */
   infile file-specification-1 sharebuffers;
   file file-specification-1;
```

```
   input state $ 1-2 phone $ 5-16;
     /* Replace area code for NC exchanges */
   if state= 'NC' and substr(phone,5,3)='333' then
     phone='910-'||substr(phone,5,8);
   put phone 5-16;
run;
```

Example 7: Truncating Copied Records The LENGTH= option is useful when you copy the input file to another file with the PUT _INFILE_ statement. Use LENGTH= to truncate the copied records. For example, these statements truncate the last 20 columns from each input data record before the input data record is written to the output file:

```
data _null_;
   infile  file-specification-1 length=a;
   input;
   a=a-20;
   file file-specification-2;
   put _infile_;
run;
```

The START= option is also useful when you want to truncate what the PUT _INFILE_ statement copies. For example, if you do not want to copy the first 10 columns of each record, these statements copy from column 11 to the end of each record in the input buffer:

```
data _null_;
   infile file-specification start=s;
   input;
   s=11;
   file file-specification-2;
   put _infile_;
run;
```

Example 8: Listing the Pointer Location This DATA step assigns the value of the current pointer location in the input buffer to the variables LINEPT and COLUMNPT:

```
data _null_;
   infile datalines n=2 line=Linept col=Columnpt;
   input name $ 1-15 #2 @3 id;
   put linept= columnpt=;
   datalines;
J. Brooks
  40974
T. R. Ansen
  4032
;
```

These statements produce the following line for each execution of the DATA step because the input pointer is on the second line in the input buffer when the PUT statement executes:

```
Linept=2 Columnpt=9
Linept=2 Columnpt=8
```

Example 9: Working with Data in the Input Buffer The _INFILE_ variable always contains the most recent record that is read from an INPUT statement. This example illustrates the use of the _INFILE_ variable to

read an entire record that you want to parse without using the INPUT statement.

read an entire record that you want to write to the SAS log.

modify the contents of the input record before parsing the line with an INPUT statement.

The example file contains phone bill information. The numeric data, minutes, and charge are enclosed in angle brackets (< >).

```
filename phonbill host-specific-filename;
data _null_;
   file phonbill;
   input line $80.;
   put line;
   datalines;
   City Number Minutes Charge
   Jackson 415-555-2384 <25> <2.45>
   Jefferson 813-555-2356 <15> <1.62>
   Joliet 913-555-3223 <65> <10.32>
   ;
run;
```

The following code reads each record and parses the record to extract the minute and charge values.

```
data _null_;
   infile phonbill firstobs=2;
   input;
   city = scan(_infile_, 1, ' ');
   char_min = scan(_infile_, 3, ' ');
   char_min = substr(char_min, 2, length(char_min)-2);
   minutes = input(char_min, BEST12.);
   put city= minutes=;
run;
```

The program writes the following lines to the SAS log:

```
city=Jackson minutes=25
city=Jefferson minutes=15
city=Joliet minutes=65
```

The INPUT statement in the following code reads a record from the file. The automatic _INFILE_ variable is used in the PUT statement to write the record to the log.

```
data _null_;
   infile phonbill;
   input;
   put _infile_;
run;
```

The program writes the following lines to the SAS log:

```
City Number Minutes Charge
Jackson 415-555-2384 <25> <2.45>
Jefferson 813-555-2356 <15> <1.62>
Joliet 913-555-3223 <65> <10.32>
```

In the following code, the first INPUT statement reads and holds the record in the input buffer. The _INFILE_= option removes the angle brackets (< >) from the numeric data. The second INPUT statement parses the value in the buffer.

```
data _null_;
   length city number $16. minutes charge 8;
   infile phonbill firstobs=2;
   input @;
   _infile_ = compress(_infile_, '<>');
   input city number minutes charge;
   put city= number= minutes= charge=;
run;
```

The program writes the following lines to the SAS log:

```
city=Jackson number=415-555-2384 minutes=25 charge=2.45
city=Jefferson number=813-555-2356 minutes=15 charge=1.62
city=Joliet number=913-555-3223 minutes=65 charge=10.32
```

Example 10: Accessing the Input Buffers of Multiple Files This example uses both the _INFILE_ automatic variable and the _INFILE_= option to read multiple files and access the input buffers for each of them. The following code creates four files: three data files and one file that contains the names of all the data files. The second DATA step reads the filenames file, opens each data file, and writes the contents to the log. Because the PUT statement needs _INFILE_ for the filenames file and the data file, one of the _INFILE_ variables is referenced with *fname*.

```
data _null_;
   do i = 1 to 3;
      fname= 'external-data-file' || put(i,1.) || '.dat';
      file datfiles filevar=fname;
      do j = 1 to 5;
         put i j;
      end;

      file 'external-filenames-file';
      put fname;
   end;
run;

data _null_;
   infile 'external-filenames-file' _infile_=fname;
   input;

   infile datfiles filevar=fname end=eof;
   do while(^eof);
      input;
      put fname _infile_;
   end;
run;
```

The program writes the following lines to the SAS log:

```
NOTE: The infile 'external-filenames-file' is:
      File Name=external-filenames-file,
      RECFM=V, LRECL=256

NOTE: The infile DATFILES is:
      File Name=external-data-file1.dat,
      RECFM=V, LRECL=256
```

```
external-data-file1.dat 1 1
external-data-file1.dat 1 2
external-data-file1.dat 1 3
external-data-file1.dat 1 4
external-data-file1.dat 1 5

NOTE: The infile DATFILES is
      File Name=external-data-file2.dat,
      RECFM=V, LRECL=256

external-data-file2.dat 2 1
external-data-file2.dat 2 2
external-data-file2.dat 2 3
external-data-file2.dat 2 4
external-data-file2.dat 2 5

NOTE: The infile DATFILES is
      File Name=external-data-file3.dat,
      RECFM=V, LRECL=256

external-data-file3.dat 3 1
external-data-file3.dat 3 2
external-data-file3.dat 3 3
external-data-file3.dat 3 4
external-data-file3.dat 3 5
```

Example 11: Specifying an Encoding When Reading an External File This example creates a SAS data set from an external file. The external file's encoding is in UTF-8, and the current SAS session encoding is Wlatin1. By default, SAS assumes that the external file is in the same encoding as the session encoding, which causes the character data to be written to the new SAS data set incorrectly.

To tell SAS what encoding to use when reading the external file, specify the ENCODING= option. When you tell SAS that the external file is in UTF-8, SAS then transcodes the external file from UTF-8 to the current session encoding when writing to the new SAS data set. Therefore, the data is written to the new data set correctly in Wlatin1.

```
libname myfiles 'SAS-library';

filename extfile 'external-file';

data myfiles.unicode;
   infile extfile encoding="utf-8";
   input Make $ Model $ Year;
run;
```

See Also

Statements:

"FILENAME Statement" on page 1520

"INPUT Statement" on page 1617

"PUT Statement" on page 1708

INFORMAT Statement

Associates informats with variables.

Valid: in a DATA step or PROC step
Category: Information
Type: Declarative

Syntax

INFORMAT *variable-1* <*…variable-n*> <*informat*>;

INFORMAT <*variable-1*> <*… variable-n*> <DEFAULT=*default-informat*>;

INFORMAT *variable-1* <*…variable-n*> *informat* <DEFAULT=*default-informat*>;

Arguments

variable
specifies one or more variables to associate with an informat. You must specify at least one *variable* when specifying an *informat* or when including no other arguments. Specifying a variable is optional when using a DEFAULT= informat specification.

Tip: To disassociate an informat from a variable, use the variable's name in an INFORMAT statement without specifying an informat. Place the INFORMAT statement after the SET statement. See Example 3 on page 1617.

informat
specifies the informat for reading the values of the variables that are listed in the INFORMAT statement.

Tip: If an informat is associated with a variable by using the INFORMAT statement, and that same informat is not associated with that same variable in the INPUT statement, then that informat will behave like informats that you specify with a colon (:) modifier in an INPUT statement. SAS reads the variables by using list input with an informat. For example, you can use the : modifier with an informat to read character values that are longer than eight bytes, or numeric values that contain nonstandard values. For details, see "INPUT Statement, List" on page 1639.

See Also: "Informats by Category" on page 1273

Featured in: Example 2 on page 1616

DEFAULT= *default-informat*
specifies a temporary default informat for reading values of the variables that are listed in the INFORMAT statement. If no *variable* is specified, then the DEFAULT= informat specification applies a temporary default informat for reading values of all the variables of that type included in the DATA step. Numeric informats are applied to numeric variables, and character informats are applied to character variables. These default informats apply only to the current DATA step.

A DEFAULT= informat specification applies to

- variables that are not named in an INFORMAT or ATTRIB statement
- variables that are not permanently associated with an informat within a SAS data set

- variables that are not read with an explicit informat in the current DATA step.

Default: If you omit DEFAULT=, SAS uses *w.d* as the default numeric informat and $*w*. as the default character informat.

Restriction: Use this argument only in a DATA step.

Tip: A DEFAULT= specification can occur anywhere in an INFORMAT statement. It can specify either a numeric default, a character default, or both.

Featured in: Example 1 on page 1616

Details

The Basics An INFORMAT statement in a DATA step permanently associates an informat with a variable. You can specify standard SAS informats or user-written informats, previously defined in PROC FORMAT. A single INFORMAT statement can associate the same informat with several variables, or it can associate different informats with different variables. If a variable appears in multiple INFORMAT statements, SAS uses the informat that is assigned last.

CAUTION:
Because an INFORMAT statement defines the length of previously undefined character variables, you can truncate the values of character variables in a DATA step if an INFORMAT statement precedes a SET statement. △

How SAS Treats Variables when You Assign Informats with the INFORMAT Statement Informats that are associated with variables by using the INFORMAT statement behave like informats that are used with modified list input. SAS reads the variables by using the scanning feature of list input, but applies the informat. In modified list input, SAS

- does not use the value of *w* in an informat to specify column positions or input field widths in an external file
- uses the value of *w* in an informat to specify the length of previously undefined character variables
- ignores the value of *w* in numeric informats
- uses the value of *d* in an informat in the same way it usually does for numeric informats
- treats blanks that are embedded as input data as delimiters unless you change their status with a DLM= or DLMSTR= option specification in an INFILE statement.

If you have coded the INPUT statement to use another style of input, such as formatted input or column input, that style of input is not used when you use the INFORMAT statement.

Comparisons

- Both the ATTRIB and INFORMAT statements can associate informats with variables, and both statements can change the informat that is associated with a variable. You can also use the INFORMAT statement in PROC DATASETS to change or remove the informat that is associated with a variable. The SAS windowing environment allows you to associate, change, or disassociate informats and variables in existing SAS data sets.
- SAS changes the descriptor information of the SAS data set that contains the variable. You can use an INFORMAT statement in some PROC steps, but the

rules are different. See "The FORMAT Procedure" in *Base SAS Procedures Guide* for more information.

Examples

Example 1: Specifying Default Informats This example uses an INFORMAT statement to associate a default numeric informat:

```
data tstinfmt;
   informat default=3.1;
   input x;
   put x;
   datalines;
111
222
333
;
```

The PUT statement produces these results:

```
11.1
22.2
33.3
```

Example 2: Specifying Numeric and Character Informats This example associates a character informat and a numeric informat with SAS variables. Although the character variables do not fully occupy 15 column positions, the INPUT statement reads the data records correctly by using modified list input:

```
data name;
   informat FirstName LastName $15. n1 6.2 n2 7.3;
   input firstname lastname n1 n2;
   datalines;
Alexander Robinson 35 11
;

proc contents data=name;
run;

proc print data=name;
run;
```

The following output shows a partial listing from PROC CONTENTS, as well as the report PROC PRINT generates.

Output 6.11 Associating Numeric and Character Informats with SAS Variables

```
                        The SAS System                          3

                     CONTENTS PROCEDURE

      -----Alphabetic List of Variables and Attributes-----

        #    Variable     Type    Len    Pos    Informat
        ----------------------------------------------
        1    FirstName    Char     15     16    $15.
        2    LastName     Char     15     31    $15.
        3    n1           Num       8      0    6.2
        4    n2           Num       8      8    7.3
```

```
                        The SAS System                        4

        OBS    FirstName    LastName     n1        n2

         1     Alexander    Robinson    0.35    0.011
```

Example 3: Removing an Informat This example disassociates an existing informat. The order of the INFORMAT and SET statements is important.

```
data rtest;
   set rtest;
   informat x;
run;
```

See Also

Statements:

"ATTRIB Statement" on page 1448

"INPUT Statement" on page 1617

"INPUT Statement, List" on page 1639

INPUT Statement

Describes the arrangement of values in the input data record and assigns input values to the corresponding SAS variables.

Valid: in a DATA step

Category: File-handling

Type: Executable

Syntax

INPUT <*specification(s)*><@|@@>;

Without Arguments

The INPUT statement with no arguments is called a *null INPUT statement*. The null INPUT statement

- brings an input data record into the input buffer without creating any SAS variables
- releases an input data record that is held by a trailing @ or a double trailing @.

For an example, see Example 2 on page 1628.

Arguments

specification(s)
can include

variable
names a variable that is assigned input values.

(*variable-list*)
specifies a list of variables that are assigned input values.

Requirement: The (*variable-list*) is followed by an (*informat-list*).

See Also: "How to Group Variables and Informats" on page 1637

$
specifies to store the variable value as a character value rather than as a numeric value.

Tip: If the variable is previously defined as character, $ is not required.

Featured in: Example 1 on page 1628

pointer-control
moves the input pointer to a specified line or column in the input buffer.

See: "Column Pointer Controls" on page 1619 and "Line Pointer Controls" on page 1621

column-specifications
specifies the columns of the input record that contain the value to read.

Tip: Informats are ignored. Only standard character and numeric data can be read correctly with this method.

See: "Column Input" on page 1622

Featured in: Example 1 on page 1628

format-modifier
allows modified list input or controls the amount of information that is reported in the SAS log when an error in an input value occurs.

Tip: Use modified list input to read data that cannot be read with simple list input.

See: "When to Use List Input" on page 1640

See: "Format Modifiers for Error Reporting" on page 1621

Featured in: Example 6 on page 1631

informat.
specifies an informat to use to read the variable value.

Tip: You can use modified list input to read data with informats. Modified list input is useful when the data require informats but cannot be read with formatted input because the values are not aligned in columns.

See: "Formatted Input" on page 1623 and "List Input" on page 1622

Featured in: Example 2 on page 1638

(*informat-list*)
specifies a list of informats to use to read the values for the preceding list of variables.

Restriction: The (*informat-list*) must follow the (*variable-list*).

See: "How to Group Variables and Informats" on page 1637

@
holds an input record for the execution of the next INPUT statement within the same iteration of the DATA step. This line-hold specifier is called *trailing* @.

Restriction: The trailing @ must be the last item in the INPUT statement.

Tip: The trailing @ prevents the next INPUT statement from automatically releasing the current input record and reading the next record into the input buffer. It is useful when you need to read from a record multiple times.

See Also: "Using Line-Hold Specifiers" on page 1624

Featured in: Example 3 on page 1629

@@
holds the input record for the execution of the next INPUT statement across iterations of the DATA step. This line-hold specifier is called *double trailing* @.

Restriction: The double trailing @ must be the last item in the INPUT statement.

Tip: The double trailing @ is useful when each input line contains values for several observations, or when a record needs to be reread on the next iteration of the DATA step.

See Also: "Using Line-Hold Specifiers" on page 1624

Featured in: Example 4 on page 1630

Column Pointer Controls

@*n*
moves the pointer to column *n*.

Range: a positive integer

Tip: If *n* is not an integer, SAS truncates the decimal value and uses only the integer value. If *n* is zero or negative, the pointer moves to column 1.

Example: @15 moves the pointer to column 15:

```
input @15 name $10.;
```

Featured in: Example 7 on page 1631

@*numeric-variable*
moves the pointer to the column given by the value of *numeric-variable*.

Range: a positive integer

Tip: If *numeric-variable* is not an integer, SAS truncates the decimal value and only uses the integer value. If *numeric-variable* is zero or negative, the pointer moves to column 1.

Example: The value of the variable A moves the pointer to column 15:

```
a=15;
input @a name $10.;
```

Featured in: Example 5 on page 1630

@(*expression*)
moves the pointer to the column that is given by the value of *expression*.

Restriction: *Expression* must result in a positive integer.

Tip: If the value of *expression* is not an integer, SAS truncates the decimal value and only uses the integer value. If it is zero or negative, the pointer moves to column 1.

Example: The result of the expression moves the pointer to column 15:

```
b=5;
input @(b*3) name $10.;
```

@'*character-string*'
locates the specified series of characters in the input record and moves the pointer to the first column after *character-string*.

@*character-variable*
locates the series of characters in the input record that is given by the value of *character-variable* and moves the pointer to the first column after that series of characters.

Example: The following statement reads in the WEEKDAY character variable. The second @1 moves the pointer to the beginning of the input line. The value for SALES is read from the next non-blank column after the value of WEEKDAY:

```
input @1 day 1. @5 weekday $10.
      @1 @weekday sales 8.2;
```

Featured in: Example 6 on page 1631

@(*character-expression*)
locates the series of characters in the input record that is given by the value of *character-expression* and moves the pointer to the first column after the series.

Featured in: Example 6 on page 1631

+*n*
moves the pointer *n* columns.

Range: a positive integer or zero

Tip: If *n* is not an integer, SAS truncates the decimal value and uses only the integer value. If the value is greater than the length of the input buffer, the pointer moves to column 1 of the next record.

Example: This statement moves the pointer to column 23, reads a value for LENGTH from columns 23 through 26, advances the pointer five columns, and reads a value for WIDTH from columns 32 through 35:

```
input @23 length 4. +5 width 4.;
```

Featured in: Example 7 on page 1631

+*numeric-variable*
moves the pointer the number of columns that is given by the value of *numeric-variable*.

Range: a positive or negative integer or zero

Tip: If *numeric-variable* is not an integer, SAS truncates the decimal value and uses only the integer value. If *numeric-variable* is negative, the pointer moves backward. If the current column position becomes less than 1, the pointer moves to column 1. If the value is zero, the pointer does not move. If the value is greater than the length of the input buffer, the pointer moves to column 1 of the next record.

Featured in: Example 7 on page 1631

+(*expression*)
moves the pointer the number of columns given by *expression*.

Range: *expression* must result in a positive or negative integer or zero.

Tip: If *expression* is not an integer, SAS truncates the decimal value and uses only the integer value. If *expression* is negative, the pointer moves backward. If

the current column position becomes less than 1, the pointer moves to column 1. If the value is zero, the pointer does not move. If the value is greater than the length of the input buffer, the pointer moves to column 1 of the next record.

Line Pointer Controls

#*n*
moves the pointer to record *n*.

Range: a positive integer

Interaction: The N= option in the INFILE statement can affect the number of records the INPUT statement reads and the placement of the input pointer after each iteration of the DATA step. See the option N= on page 1597.

Example: The #2 moves the pointer to the second record to read the value for ID from columns 3 and 4:

```
input name $10. #2 id 3-4;
```

#*numeric-variable*
moves the pointer to the record that is given by the value of *numeric-variable*.

Range: a positive integer

Tip: If the value of *numeric-variable* is not an integer, SAS truncates the decimal value and uses only the integer value.

#(*expression*)
moves the pointer to the record that is given by the value of *expression*.

Range: *expression* must result in a positive integer.

Tip: If the value of *expression* is not an integer, SAS truncates the decimal value and uses only the integer value.

/
advances the pointer to column 1 of the next input record.

Example: The values for NAME and AGE are read from the first input record before the pointer moves to the second record to read the value of ID from columns 3 and 4:

```
input name age / id 3-4;
```

Format Modifiers for Error Reporting

?
suppresses printing the invalid data note when SAS encounters invalid data values.

See Also: "How Invalid Data is Handled" on page 1627

??
suppresses printing the messages and the input lines when SAS encounters invalid data values. The automatic variable _ERROR_ is not set to 1 for the invalid observation.

See Also: "How Invalid Data is Handled" on page 1627

Details

When to Use INPUT Use the INPUT statement to read raw data from an external file or in-stream data. If your data are stored in an external file, you can specify the file in an INFILE statement. The INFILE statement must execute before the INPUT

statement that reads the data records. If your data are in-stream, a DATALINES statement must precede the data lines in the job stream. If your data contain semicolons, use a DATALINES4 statement before the data lines. A DATA step that reads raw data can include multiple INPUT statements.

You can also use the INFILE statement to read in-stream data by specifying a filename of DATALINES in the INFILE statement before the INPUT statement. Using DATALINES in the INFILE statement allows you to use most of the options available in the INFILE statement with in-stream data.

To read data that are already stored in a SAS data set, use a SET statement. To read database or PC file-format data that are created by other software, use the SET statement after you access the data with the LIBNAME statement. See the SAS/ACCESS documentation for more information.

Operating Environment Information: LOG files that are generated under z/OS and captured with PROC PRINTTO contain an ASA control character in column 1. If you are using the INPUT statement to read a LOG file that was generated under z/OS, you must account for this character if you use column input or column pointer controls. △

Input Styles

There are four ways to describe a record's values in the INPUT statement:

- column
- list (simple and modified)
- formatted
- named.

Each variable value is read by using one of these input styles. An INPUT statement can contain any or all of the available input styles, depending on the arrangement of data values in the input records. However, once named input is used in an INPUT statement, you cannot use another input style.

Column Input With *column input*, the column numbers follow the variable name in the INPUT statement. These numbers indicate where the variable values are found in the input data records:

```
input name $ 1-8 age 11-12;
```

This INPUT statement can read the following data records:

```
----+----1----+----2----+
Peterson  21
Morgan    17
```

Because NAME is a character variable, a $ appears between the variable name and column numbers. For more information, see "INPUT Statement, Column" on page 1632.

List Input With *list input*, the variable names are simply listed in the INPUT statement. A $ follows the name of each character variable:

```
input name $ age;
```

This INPUT statement can read data values that are separated by blanks or aligned in columns (with at least one blank between):

```
----+----1----+----2----+
Peterson  21
Morgan  17
```

For more information, see "INPUT Statement, List" on page 1639.

Formatted Input With *formatted input*, an informat follows the variable name in the INPUT statement. The informat gives the data type and the field width of an input value. Informats also allow you to read data that are stored in nonstandard form, such as packed decimal, or numbers that contain special characters such as commas.

```
input name $char8. +2 income comma6.;
```

This INPUT statement reads these data records correctly:

```
----+----1----+----2----+
Peterson  21,000
Morgan    17,132
```

The pointer control of +2 moves the input pointer to the field that contains the value for the variable INCOME. For more information, see "INPUT Statement, Formatted" on page 1635.

Named Input With *named input*, you specify the name of the variable followed by an equal sign. SAS looks for a variable name and an equal sign in the input record:

```
input name= $ age=;
```

This INPUT statement reads the following data records correctly:

```
----+----1----+----2----+
name=Peterson age=21
name=Morgan age=17
```

For more information, see "INPUT Statement, Named" on page 1645.

Multiple Styles in a Single INPUT Statement

An INPUT statement can contain any or all of the different input styles:

```
input idno name $18. team $ 25-30 startwght endwght;
```

This INPUT statement reads the following data records correctly:

```
----+----1----+----2----+----3----+----
023 David Shaw         red    189 165
049 Amelia Serrano     yellow 189 165
```

The value of IDNO, STARTWGHT, and ENDWGHT are read with list input, the value of NAME with formatted input, and the value of TEAM with column input.

Note: Once named input is used in an INPUT statement, you cannot change input styles. △

Pointer Controls

As SAS reads values from the input data records into the input buffer, it keeps track of its position with a pointer. The INPUT statement provides three ways to control the movement of the pointer:

column pointer controls
: reset the pointer's column position when the data values in the data records are read.

line pointer controls
: reset the pointer's line position when the data values in the data records are read.

line-hold specifiers
hold an input record in the input buffer so that another INPUT statement can process it. By default, the INPUT statement releases the previous record and reads another record.

With column and line pointer controls, you can specify an absolute line number or column number to move the pointer or you can specify a column or line location relative to the current pointer position. Table 6.6 on page 1624 lists the pointer controls that are available with the INPUT statement.

Table 6.6 Pointer Controls Available in the INPUT Statement

Pointer Controls	Relative	Absolute
column pointer controls	+*n*	@*n*
	+*numeric-variable*	@*numeric-variable*
	+(*expression*)	@*(expression)*
		@'*character-string*'
		@*character-variable*
		@(*character-expression*)
line pointer controls	/	#*n*
		#*numeric-variable*
		#(*expression*)
line-hold specifiers	@	(not applicable)
	@@	(not applicable)

Note: Always specify pointer controls before the variable to which they apply. △

You can use the COLUMN= and LINE= options in the INFILE statement to determine the pointer's current column and line location.

Using Column and Line Pointer Controls Column pointer controls indicate the column in which an input value starts.

Use line pointer controls within the INPUT statement to move to the next input record or to define the number of input records per observation. Line pointer controls specify which input record to read. To read multiple data records into the input buffer, use the N= option in the INFILE statement to specify the number of records. If you omit N=, you need to take special precautions. For more information, see "Reading More Than One Record per Observation" on page 1626.

Using Line-Hold Specifiers Line-hold specifiers keep the pointer on the current input record when

- a data record is read by more than one INPUT statement (trailing @)
- one input line has values for more than one observation (double trailing @)
- a record needs to be reread on the next iteration of the DATA step (double trailing @).

Use a single trailing @ to allow the next INPUT statement to read from the same record. Use a double trailing @ to hold a record for the next INPUT statement across iterations of the DATA step.

Normally, each INPUT statement in a DATA step reads a new data record into the input buffer. When you use a trailing @, the following occurs:

- The pointer position does not change.
- No new record is read into the input buffer.
- The next INPUT statement for the same iteration of the DATA step continues to read the same record rather than a new one.

SAS releases a record held by a trailing @ when

- a null INPUT statement executes:

```
input;
```

- an INPUT statement without a trailing @ executes
- the next iteration of the DATA step begins.

Normally, when you use a double trailing @ (@@), the INPUT statement for the next iteration of the DATA step continues to read the same record. SAS releases the record that is held by a double trailing @

- immediately if the pointer moves past the end of the input record
- immediately if a null INPUT statement executes:

```
input;
```

- when the next iteration of the DATA step begins if an INPUT statement with a single trailing @ executes later in the DATA step:

```
input @;
```

Pointer Location After Reading Understanding the location of the input pointer after a value is read is important, especially if you combine input styles in a single INPUT statement. With column and formatted input, the pointer reads the columns that are indicated in the INPUT statement and stops in the next column. With list input, however, the pointer scans data records to locate data values and reads a blank to indicate that a value has ended. After reading a value with list input, the pointer stops in the second column after the value.

For example, you can read these data records with list, column, and formatted input:

```
----+----1----+----2----+----3
REGION1    49670
REGION2    97540
REGION3    86342
```

This INPUT statement uses list input to read the data records:

```
input region $ jansales;
```

After reading a value for REGION, the pointer stops in column 9.

```
----+----1----+----2----+----3
REGION1    49670
        ↑
```

These INPUT statements use column and formatted input to read the data records:

- column input

```
input region $ 1-7 jansales 12-16;
```

- formatted input

```
input region $7. +4 jansales 5.;
input region $7. @12 jansales 5.;
```

To read a value for the variable REGION, the INPUT statements instruct the pointer to read seven columns and stop in column 8.

```
----+----1----+----2----+----3
REGION1    49670
       ↑
```

Reading More Than One Record per Observation

The highest number that follows the # pointer control in the INPUT statement determines how many input data records are read into the input buffer. Use the N= option in the INFILE statement to change the number of records. For example, in this statement, the highest value after the # is 3:

```
input @31 age 3. #3 id 3-4 #2 @6 name $20.;
```

Unless you use N= in the associated INFILE statement, the INPUT statement reads three input records each time the DATA step executes.

When each observation has multiple input records but values from the last record are not read, you must use a # pointer control in the INPUT statement or N= in the INFILE statement to specify the last input record. For example, if there are four records per observation, but only values from the first two input records are read, use this INPUT statement:

```
input name $ 1-10 #2 age 13-14 #4;
```

When you have advanced to the next record with the / pointer control, use the # pointer control in the INPUT statement or the N= option in the INFILE statement to set the number of records that are read into the input buffer. To move the pointer back to an earlier record, use a # pointer control. For example, this statement requires the #2 pointer control, unless the INFILE statement uses the N= option, to read two records:

```
input a / b #1 @52 c #2;
```

The INPUT statement assigns A a value from the first record. The pointer advances to the next input record to assign B a value. Then the pointer returns from the second record to column 1 of the first record and moves to column 52 to assign C a value. The #2 pointer control identifies two input records for each observation so that the pointer can return to the first record for the value of C.

If the number of input records per observation varies, use the N= option in the INFILE statement to give the maximum number of records per observation. For more information, see the N= option on page 1597.

Reading Past the End of a Line When you use @ or + pointer controls with a value that moves the pointer to or past the end of the current record and the next value is to be read from the current column, SAS goes to column 1 of the next record to read it. It also writes this message to the SAS log:

```
NOTE: SAS went to a new line when INPUT statement
      reached past the end of a line.
```

You can alter the default behavior (the FLOWOVER option) in the INFILE statement.

Use the STOPOVER option in the INFILE statement to treat this condition as an error and to stop building the data set.

Use the MISSOVER option in the INFILE statement to set the remaining INPUT statement variables to missing values if the pointer reaches the end of a record.

Use the TRUNCOVER option in the INFILE statement to read column input or formatted input when the last variable that is read by the INPUT statement contains varying-length data.

Positioning the Pointer Before the Record When a column pointer control tries to move the pointer to a position before the beginning of the record, the pointer is positioned in column 1. For example, this INPUT statement specifies that the pointer is located in column –2 after the first value is read:

```
data test;
   input a @(a-3) b;
   datalines;
 2
;
```

Therefore, SAS moves the pointer to column 1 after the value of A is read. Both variables A and B contain the same value.

How Invalid Data is Handled

When SAS encounters an invalid character in an input value for the variable indicated, it

- sets the value of the variable that is being read to missing or the value that is specified with the INVALIDDATA= system option. For more information see "INVALIDDATA= System Option" on page 1931.
- prints an invalid data note in the SAS log.
- prints the input line and column number that contains the invalid value in the SAS log. Unprintable characters appear in hexadecimal. To help determine column numbers, SAS prints a rule line above the input line.
- sets the automatic variable _ERROR_ to 1 for the current observation.

The format modifiers for error reporting control the amount of information that is printed in the SAS log. Both the ? and ?? modifier suppress the invalid data message. However, the ?? modifier also resets the automatic variable _ERROR_ to 0. For example, these two sets of statements are equivalent:

- `input x ?? 10-12;`
- ```
  input x ? 10-12;
  _error_=0;
  ```

In either case, SAS sets invalid values of X to missing values. For information about the causes of invalid data, see *SAS Language Reference: Concepts*.

End-of-File

End-of-file occurs when an INPUT statement reaches the end of the data. If a DATA step tries to read another record after it reaches an end-of-file then execution stops. If you want the DATA step to continue to execute, use the END= or EOF= option in the INFILE statement. Then you can write SAS program statements to detect the end-of-file, and to stop the execution of the INPUT statement but continue with the DATA step. For more information, see "INFILE Statement" on page 1591.

Arrays

The INPUT statement can use array references to read input data values. You can use an array reference in a pointer control if it is enclosed in parentheses. See Example 6 on page 1631.
  ```

Use the array subscript asterisk (*) to input all elements of a previously defined explicit array. SAS allows single or multidimensional arrays. Enclose the subscript in braces, brackets, or parentheses. The form of this statement is

```
INPUT array-name{*};
```

You can use arrays with list, column, or formatted input. However, you cannot input values to an array that is defined with _TEMPORARY_ and that uses the asterisk subscript. For example, these statements create variables X1 through X100 and assign data values to the variables using the 2. informat:

```
array x{100};
input x{*} 2.;
```

Comparisons

- The INPUT statement reads raw data in external files or data lines that are entered in-stream (following the DATALINES statement) that need to be described to SAS. The SET statement reads a SAS data set, which already contains descriptive information about the data values.
- The INPUT statement reads data while the PUT statement writes data values, text strings, or both to the SAS log or to an external file.
- The INPUT statement can read data from external files; the INFILE statement points to that file and has options that control how that file is read.

Examples

Example 1: Using Multiple Styles of Input in One INPUT Statement This example uses several input styles in a single INPUT statement:

```
data club1;
   input Idno Name $18.
         Team $ 25-30 Startwght Endwght;
   datalines;
023 David Shaw         red    189 165
049 Amelia Serrano     yellow 189 165
... more data lines ...
;
```

Variable	Type of Input
Idno, Startwght, Endwght	list input
Name	formatted input
Team	column input

Example 2: Using a Null INPUT Statement This example uses an INPUT statement with no arguments. The DATA step copies records from the input file to the output file without creating any SAS variables:

```
data _null_;
   infile file-specification-1;
   file file-specification-2;
   input;
```

```
   put _infile_;
run;
```

Example 3: Holding a Record in the Input Buffer This example reads a file that contains two types of input data records and creates a SAS data set from these records. One type of data record contains information about a particular college course. The second type of record contains information about the students enrolled in the course. You need two INPUT statements to read the two records and to assign the values to different variables that use different formats. Records that contain class information have a C in column 1; records that contain student information have an S in column 1, as shown here:

```
----+----1----+----2----+
C HIST101 Watson
S Williams 0459
S Flores   5423
C MATH202 Sen
S Lee      7085
```

To know which INPUT statement to use, check each record as it is read. Use an INPUT statement that reads only the variable that tells whether the record contains class or student.

```
data schedule(drop=type);
   infile file-specification;
   retain Course Professor;
   input type $1. @;
   if type='C' then
      input course $ professor $;
   else if type='S' then
      do;
        input Name $10. Id;
        output schedule;
      end;
run;

proc print;
run;
```

The first INPUT statement reads the TYPE value from column 1 of every line. Because this INPUT statement ends with a trailing @, the next INPUT statement in the DATA step reads the same line. The IF-THEN statements that follow check whether the record is a class or student line before another INPUT statement reads the rest of the line. The INPUT statements without a trailing @ release the held line. The RETAIN statement saves the values about the particular college course. The DATA step writes an observation to the SCHEDULE data set after a student record is read.

The following output that PROC PRINT generates shows the resulting data set SCHEDULE.

Output 6.12 Data Set Schedule

```
                        The SAS System                          1

       OBS    Course     Professor    Name          Id

         1    HIST101     Watson      Williams      459
         2    HIST101     Watson      Flores       5423
         3    MATH202     Sen         Lee          7085
```

Example 4: Holding a Record Across Iterations of the DATA Step This example shows how to create multiple observations for each input data record. Each record contains several NAME and AGE values. The DATA step reads a NAME value and an AGE value, outputs an observation, and then reads another set of NAME and AGE values to output, and so on, until all the input values in the record are processed.

```
data test;
   input name $ age @@;
   datalines;
John 13 Monica 12 Sue 15 Stephen 10
Marc 22 Lily 17
;
```

The INPUT statement uses the double trailing @ to control the input pointer across iterations of the DATA step. The SAS data set contains six observations.

Example 5: Positioning the Pointer with a Numeric Variable This example uses a numeric variable to position the pointer. A raw data file contains records with the employment figures for several offices of a multinational company. The input data records are

```
----+----1----+----2----+----3----+
8      New York    1 USA 14
5   Cary           1 USA 2274
3 Chicago          1 USA 37
22 Tokyo           5 ASIA 80
5   Vancouver      2 CANADA 6
9       Milano     4 EUROPE 123
```

The first column has the column position for the office location. The next numeric column is the region category. The geographic region occurs before the number of employees in that office.

You determine the office location by combining the @*numeric-variable* pointer control with a trailing @. To read the records, use two INPUT statements. The first INPUT statement obtains the value for the @ *numeric-variable* pointer control. The second INPUT statement uses this value to determine the column that the pointer moves to.

```
data office (drop=x);
   infile file-specification;
   input x @;
   if 1<=x<=10 then
      input @x City $9.;
   else do;
      put 'Invalid input at line ' _n_;
      delete;
   end;
run;
```

The DATA step writes only five observations to the OFFICE data set. The fourth input data record is invalid because the value of X is greater than 10. Therefore, the second INPUT statement does not execute. Instead, the PUT statement writes a message to the SAS log and the DELETE statement stops processing the observation.

Example 6: Positioning the Pointer with a Character Variable This example uses character variables to position the pointer. The OFFICE data set, created in Example 5 on page 1630, contains a character variable CITY whose values are the office locations. Suppose you discover that you need to read additional values from the raw data file. By using another DATA step, you can combine the @*character-variable* pointer control with a trailing @ and the @*character-expression* pointer control to locate the values.

If the observations in OFFICE are still in the order of the original input data records, you can use this DATA step:

```
data office2;
   set office;
   infile file-specification;
   array region {5} $ _temporary_
         ('USA' 'CANADA' 'SA' 'EUROPE' 'ASIA');
   input @city Location : 2. @;
   input @(trim(region{location})) Population : 4.;
run;
```

The ARRAY statement assigns initial values to the temporary array elements. These elements correspond to the geographic regions of the office locations. The first INPUT statement uses an @*character-variable* pointer control. Each record is scanned for the series of characters in the value of CITY for that observation. Then the value of LOCATION is read from the next non-blank column. LOCATION is a numeric category for the geographic region of an office. The second INPUT statement uses an array reference in the @*character-expression* pointer control to determine the location POPULATION in the input records. The expression also uses the TRIM function to trim trailing blanks from the character value. This way an exact match is found between the character string in the input data and the value of the array element.

The following output that PROC PRINT generates shows the resulting data set OFFICE2.

Output 6.13 Data Set Office2

```
                    The SAS System                          1

          OBS    City          Location    Population

           1     New York           1            14
           2     Cary               1          2274
           3     Chicago            1            37
           4     Vancouver          2             6
           5     Milano             4           123
```

Example 7: Moving the Pointer Backward This example shows several ways to move the pointer backward.

- This INPUT statement uses the @ pointer control to read a value for BOOK starting at column 26. Then the pointer moves back to column 1 on the same line to read a value for COMPANY:

```
input @26 book $ @1 company;
```

- These INPUT statements use +*numeric-variable* or +(*expression*) to move the pointer backward one column. These two sets of statements are equivalent.
  - ```
    m=-1;
    input x 1-10 +m y 2.;
    ```
  - ```
    input x 1-10 +(-1) y 2.;
    ```

See Also

Statements:

"ARRAY Statement" on page 1440
"INPUT Statement, Column" on page 1632
"INPUT Statement, Formatted" on page 1635
"INPUT Statement, List" on page 1639
"INPUT Statement, Named" on page 1645

INPUT Statement, Column

Reads input values from specified columns and assigns them to the corresponding SAS variables.

Valid: in a DATA step
Category: File-handling
Type: Executable

Syntax

INPUT *variable* <$> *start-column* <– *end-column*>
<.*decimals*> <@ | @@>;

Arguments

variable
specifies a variable that is assigned input values.

$
indicates that the variable has character values rather than numeric values.

Tip: If the variable is previously defined as character, $ is not required.

start-column
specifies the first column of the input record that contains the value to read.

–end-column
specifies the last column of the input record that contains the value to read.

Tip: If the variable value occupies only one column, omit *end-column*.

Example: Because *end-column* is omitted, the values for the character variable GENDER occupy only column 16:

```
input name $ 1-10 pulse 11-13 waist 14-15
      gender $ 16;
```

.decimals
specifies the number of digits to the right of the decimal if the input value does not contain an explicit decimal point.

Tip: An explicit decimal point in the input value overrides a decimal specification in the INPUT statement.

Example: This INPUT statement reads the input data for a numeric variable using two decimal places:

Input Data	Statement	Results
`----+---1`		
`2314`	`input number 1-5 .2;`	`23.14`
`2`		`.02`
`400`		`4.00`
`-140`		`-1.40`
`12.234`		`12.234` *
`12.2`		`12.2` *

* The decimal specification in the INPUT statement is overridden by the input data value.

@
holds the input record for the execution of the next INPUT statement within the same iteration of the DATA step. This line-hold specifier is called *trailing* @.

Restriction: The trailing @ must be the last item in the INPUT statement.

Tip: The trailing @ prevents the next INPUT statement from automatically releasing the current input record and reading the next record into the input buffer. It is useful when you need to read from a record multiple times.

See: "Pointer Controls" on page 1623.

@@
holds the input record for the execution of the next INPUT statement across iterations of the DATA step. This line-hold specifier is called *double trailing* @.

Restriction: The double trailing @ must be the last item in the INPUT statement.

Tip: The double trailing @ is useful when each input line contains values for several observations.

See: "Using Line-Hold Specifiers" on page 1624.

Details

When to Use Column Input With column input, the column numbers that contain the value follow a variable name in the INPUT statement. To read with column input, data values must be in

- the same columns in all the input data records
- standard numeric form or character form.*

Useful features of column input are that

- Character values can contain embedded blanks.
- Character values can be from 1 to 32,767 characters long.
- Input values can be read in any order, regardless of their position in the record.
- Values or parts of values can be read multiple times. For example, this INPUT statement reads an ID value in columns 10 through 15 and then reads a GROUP value from column 13:

```
input id 10-15 group 13;
```

- Both leading and trailing blanks within the field are ignored. Therefore, if numeric values contain blanks that represent zeros or if you want to retain leading and trailing blanks in character values, read the value with an informat. See "INPUT Statement, Formatted" on page 1635.

Missing Values Missing data do not require a place-holder. The INPUT statement interprets a blank field as missing and reads other values correctly. If a numeric or character field contains a single period, the variable value is set to missing.

Reading Data Lines SAS always pads the data records that follow the DATALINES statement (in-stream data) to a fixed length in multiples of 80. The CARDIMAGE system option determines whether to read or to truncate data past the 80th column.

Reading Variable-Length Records By default, SAS uses the FLOWOVER option to read varying-length data records. If the record contains fewer values than expected, the INPUT statement reads the values from the next data record. To read varying-length data, you might need to use the TRUNCOVER option in the INFILE statement. The TRUNCOVER option is more efficient than the PAD option, which pads the records to a fixed length. For more information, see "Reading Past the End of a Line" on page 1604.

Examples

This DATA step demonstrates how to read input data records with column input:

```
data scores;
   input name $ 1-18 score1 25-27 score2 30-32
         score3 35-37;
   datalines;
Joseph                  11   32   76
Mitchel                 13   29   82
Sue Ellen               14   27   74
;
```

See Also

Statement:

"INPUT Statement" on page 1617

* See *SAS Language Reference: Concepts* for the definition of standard and nonstandard data values.

INPUT Statement, Formatted

Reads input values with specified informats and assigns them to the corresponding SAS variables.

Valid in a DATA step

Category: File-handling

Type: Executable

Syntax

INPUT <*pointer-control*> *variable informat.* <@ | @@>;

INPUT<*pointer-control*> (*variable-list*) (*informat-list*)
<@ | @@>;

INPUT <*pointer-control*> (*variable-list*) (<*n**> *informat.*)
<@ | @@>;

Arguments

pointer-control
moves the input pointer to a specified line or column in the input buffer.

See: "Column Pointer Controls" on page 1619 and "Line Pointer Controls" on page 1621

variable
specifies a variable that is assigned input values.

Requirement: The (*variable-list*) is followed by an (*informat-list*).

Featured in: Example 1 on page 1637

(*variable-list*)
specifies a list of variables that are assigned input values.

See: "How to Group Variables and Informats" on page 1637

Featured in: Example 2 on page 1638

informat.
specifies a SAS informat to use to read the variable values.

Tip: Decimal points in the actual input values override decimal specifications in a numeric informat.

See Also: Chapter 5, "Informats," on page 1257

Featured in: Example 1 on page 1637

(*informat-list*)
specifies a list of informats to use to read the values for the preceding list of variables In the INPUT statement, (*informat-list*) can include

informat.
specifies an informat to use to read the variable values.

pointer-control
specifies one of these pointer controls to use to position a value: @, #, /, or +.

*n**
specifies to repeat *n* times the next informat in an informat list.

Example: This statement uses the 7.2 informat to read GRADES1, GRADES2, and GRADES3 and the 5.2 informat to read GRADES4 and GRADES5:

```
input (grades1-grades5)(3*7.2, 2*5.2);
```

Restriction: The (*informat-list*) must follow the (*variable-list*).

See: "How to Group Variables and Informats" on page 1637

Featured in: Example 2 on page 1638

@
holds an input record for the execution of the next INPUT statement within the same iteration of the DATA step. This line-hold specifier is called *trailing* @.

Restriction: The trailing @ must be the last item in the INPUT statement.

Tip: The trailing @ prevents the next INPUT statement from automatically releasing the current input record and reading the next record into the input buffer. It is useful when you need to read from a record multiple times.

See: "Using Line-Hold Specifiers" on page 1624

@@
holds an input record for the execution of the next INPUT statement across iterations of the DATA step. This line-hold specifier is called *double trailing* @.

Restriction: The double trailing @ must be the last item in the INPUT statement.

Tip: The double trailing @ is useful when each input line contains values for several observations.

See: "Using Line-Hold Specifiers" on page 1624

Details

When to Use Formatted Input With formatted input, an informat follows a variable name and defines how SAS reads the values of this variable. An informat gives the data type and the field width of an input value. Informats also read data that are stored in nonstandard form, such as packed decimal, or numbers that contain special characters such as commas.* See "Definition of Informats" on page 1259 for descriptions of SAS informats.

Simple formatted input requires that the variables be in the same order as their corresponding values in the input data. You can use pointer controls to read variables in any order. For more information, see "INPUT Statement" on page 1617.

Missing Values Generally, SAS represents missing values in formatted input with a single period for a numeric value and with blanks for a character value. The informat that you use with formatted input determines how SAS interprets a blank. For example, $CHAR.*w* reads the blanks as part of the value, whereas BZ.*w* converts a blank to zero.

Reading Variable-Length Records By default, SAS uses the FLOWOVER option to read varying-length data records. If the record contains fewer values than expected, the INPUT statement reads the values from the next data record. To read varying-length data. you might need to use the TRUNCOVER option in the INFILE statement. For more information, see "Reading Past the End of a Line" on page 1604.

* See *SAS Language Reference: Concepts* for information about standard and nonstandard data values.

How to Group Variables and Informats When the input values are arranged in a pattern, you can group the informat list. A grouped informat list consists of two lists:

- the names of the variables to read enclosed in parentheses
- the corresponding informats separated by either blanks or commas and enclosed in parentheses.

Informat lists can make an INPUT statement shorter because the informat list is recycled until all variables are read and the numbered variable names can be used in abbreviated form. Using informat lists avoids listing the individual variables.

For example, if the values for the five variables SCORE1 through SCORE5 are stored as four columns per value without intervening blanks, this INPUT statement reads the values:

```
input (score1-score5) (4. 4. 4. 4. 4.);
```

However, if you specify more variables than informats, the INPUT statement reuses the informat list to read the remaining variables. A shorter version of the previous statement is

```
input (score1-score5) (4.);
```

You can use as many informat lists as necessary in an INPUT statement, but do not nest the informat lists. After all the values in the variable list are read, the INPUT statement ignores any directions that remain in the informat list. For an example, see Example 3 on page 1638.

The *n** modifier in an informat list specifies to repeat the next informat *n* times. For example,

```
input (name score1-score5) ($10. 5*4.);
```

How to Store Informats The informats that you specify in the INPUT statement are not stored with the SAS data set. Informats that you specify with the INFORMAT or ATTRIB statement are permanently stored. Therefore, you can read a data value with a permanently stored informat in a later DATA step without having to specify the informat or use PROC FSEDIT to enter data in the correct format.

Comparisons

When a variable is read with formatted input, the pointer movement is similar to the pointer movement of column input. The pointer moves the length that the informat specifies and stops at the next column. To read data with informats that are not aligned in columns, use *modified list input*. Using modified list input allows you to take advantage of the scanning feature in list input. See "When to Use List Input" on page 1640.

Examples

Example 1: Formatted Input with Pointer Controls This INPUT statement uses informats and pointer controls:

```
data sales;
   infile file-specification;
   input item $10. +5 jan comma5. +5 feb comma5.
         +5 mar comma5.;
run;
```

It can read these input data records:

```
----+----1----+----2----+----3----+----4
trucks          1,382     2,789     3,556
vans            1,265     2,543     3,987
sedans          2,391     3,011     3,658
```

The value for ITEM is read from the first 10 columns in a record. The pointer stops in column 11. The trailing blanks are discarded and the value of ITEM is written to the program data vector. Next, the pointer moves five columns to the right before the INPUT statement uses the COMMA5. informat to read the value of JAN. This informat uses five as the field width to read numeric values that contain a comma. Once again, the pointer moves five columns to the right before the INPUT statement uses the COMMA5. informat to read the values of FEB and MAR.

Example 2: Using Informat Lists This INPUT statement uses the character informat $10. to read the values of the variable NAME and uses the numeric informat 4. to read the values of the five variables SCORE1 through SCORE5:

```
data scores;
   input (name score1-score5) ($10. 5*4.);
   datalines;
Whittaker 121 114 137 156 142
Smythe    111 97  122 143 127
;
```

Example 3: Including More Informat Specifications Than Necessary This informat list includes more specifications than are necessary when the INPUT statement executes:

```
data test;
   input (x y z) (2.,+1);
   datalines;
2 24 36
0 20 30
   ;
```

The INPUT statement reads the value of X with the 2. informat. Then, the +1 column pointer control moves the pointer forward one column. Next, the value of Y is read with the 2. informat. Again, the +1 column pointer moves the pointer forward one column. Then, the value of Z is read with the 2. informat. For the third iteration, the INPUT statement ignores the +1 pointer control.

See Also

Statements:

"INPUT Statement" on page 1617

"INPUT Statement, List" on page 1639

INPUT Statement, List

Scans the input data record for input values and assigns them to the corresponding SAS variables.

Valid: in a DATA step

Category: File-handling

Type: Executable

Syntax

INPUT <*pointer-control*> *variable* <$> <&> <@ | @@>;

INPUT <*pointer-control*> *variable* <:|&|~>
<*informat.*> <@ | @@>;

Arguments

pointer-control
moves the input pointer to a specified line or column in the input buffer.

See: "Column Pointer Controls" on page 1619 and "Line Pointer Controls" on page 1621

Featured in: Example 2 on page 1643

variable
specifies a variable that is assigned input values.

$
indicates to store a variable value as a character value rather than as a numeric value.

Tip: If the variable is previously defined as character, $ is not required.

Featured in: Example 1 on page 1642

&
indicates that a character value can have one or more single embedded blanks. This format modifier reads the value from the next non-blank column until the pointer reaches two consecutive blanks, the defined length of the variable, or the end of the input line, whichever comes first.

Restriction: The & modifier must follow the variable name and $ sign that it affects.

Tip: If you specify an informat after the & modifier, the terminating condition for the format modifier remains two blanks.

See: "Modified List Input" on page 1641

Featured in: Example 2 on page 1643

:
enables you to specify an informat that the INPUT statement uses to read the variable value. For a character variable, this format modifier reads the value from the next non-blank column until the pointer reaches the next blank column, the defined length of the variable, or the end of the data line, whichever comes first. For a numeric variable, this format modifier reads the value from the next non-blank column until the pointer reaches the next blank column or the end of the data line, whichever comes first.

Tip: If the length of the variable has not been previously defined, then its value is read and stored with the informat length.

Tip: The pointer continues to read until the next blank column is reached. However, if the field is longer than the formatted length, then the value is truncated to the length of variable.

See: "Modified List Input" on page 1641

Featured in: Example 3 on page 1643 and Example 5 on page 1644

~

indicates to treat single quotation marks, double quotation marks, and delimiters in character values in a special way. This format modifier reads delimiters within quoted character values as characters instead of as delimiters and retains the quotation marks when the value is written to a variable.

Restriction: You must use the DSD option in an INFILE statement. Otherwise, the INPUT statement ignores this option.

See: "Modified List Input" on page 1641

Featured in: Example 5 on page 1644

informat.

specifies an informat to use to read the variable values.

Tip: Decimal points in the actual input values always override decimal specifications in a numeric informat.

See Also: "Definition of Informats" on page 1259

Featured in: Example 3 on page 1643 and Example 5 on page 1644

@

holds an input record for the execution of the next INPUT statement within the same iteration of the DATA step. This line-hold specifier is called *trailing* @.

Restriction: The trailing @ must be the last item in the INPUT statement.

Tip: The trailing @ prevents the next INPUT statement from automatically releasing the current input record and reading the next record into the input buffer. It is useful when you need to read from a record multiple times.

See: "Using Line-Hold Specifiers" on page 1624

@@

holds an input record for the execution of the next INPUT statement across iterations of the DATA step. This line-hold specifier is called *double trailing* @.

Restriction: The double trailing @ must be the last item in the INPUT statement.

Tip: The double trailing @ is useful when each input line contains values for several observations.

See: "Using Line-Hold Specifiers" on page 1624

Details

When to Use List Input List input requires that you specify the variable names in the INPUT statement in the same order that the fields appear in the input data records. SAS scans the data line to locate the next value but ignores additional intervening blanks. List input does not require that the data are located in specific columns. However, you must separate each value from the next by at least one blank unless the delimiter between values is changed. By default, the delimiter for data values is one blank space or the end of the input record. List input will not skip over any data values to read subsequent values, but it can ignore all values after a given point in the data

record. However, pointer controls enable you to change the order that the data values are read.

There are two types of list input:

- simple list input
- modified list input.

Modified list input makes the INPUT statement more versatile because you can use a format modifier to overcome several of the restrictions of simple list input. See "Modified List Input" on page 1641.

Simple List Input Simple list input places several restrictions on the type of data that the INPUT statement can read:

- By default, at least one blank must separate the input values. Use the DLM= or DLMSTR= option or the DSD option in the INFILE statement to specify a delimiter other than a blank.
- Represent each missing value with a period, not a blank, or two adjacent delimiters.
- Character input values cannot be longer than 8 bytes unless the variable is given a longer length in an earlier LENGTH, ATTRIB, or INFORMAT statement.
- Character values cannot contain embedded blanks unless you change the delimiter.
- Data must be in standard numeric or character format.*

Modified List Input List input is more versatile when you use format modifiers. The format modifiers are as follows:

Format Modifier	Purpose
&	reads character values that contain embedded blanks.
:	reads data values that need the additional instructions that informats can provide but that are not aligned in columns. **
~	reads delimiters within quoted character values as characters and retains the quotation marks.

** Use formatted input and pointer controls to quickly read data values that are aligned in columns.

For example, use the : modifier with an informat to read character values that are longer than 8 bytes or numeric values that contain nonstandard values.

Because list input interprets a blank as a delimiter, use modified list input to read values that contain blanks. The & modifier reads character values that contain single embedded blanks. However, the data values must be separated by two or more blanks. To read values that contain leading, trailing, or embedded blanks with list input, use the DLM= or DLMSTR= option in the INFILE statement to specify another character as the delimiter. See Example 5 on page 1644. If your input data use blanks as delimiters and they contain leading, trailing, or embedded blanks, you might need to use either column input or formatted input. If quotation marks surround the delimited values, you can use list input with the DSD option in the INFILE statement.

Comparisons

How Modified List Input and Formatted Input Differ *Modified list input* has a scanning feature that can use informats to read data which are not aligned in columns. *Formatted*

* See *SAS Language Reference: Concepts* for the information about standard and nonstandard data values.

input causes the pointer to move like that of column input to read a variable value. The pointer moves the length that is specified in the informat and stops at the next column.

This DATA step uses modified list input to read the first data value and formatted input to read the second:

```
data jansales;
   input item : $10. amount comma5.;
datalines;
trucks 1,382
vans 1,235
sedans 2,391
;
```

The value of ITEM is read with modified list input. The INPUT statement stops reading when the pointer finds a blank space. The pointer then moves to the second column after the end of the field, which is the correct position to read the AMOUNT value with formatted input.

Formatted input, on the other hand, continues to read the entire width of the field. This INPUT statement uses formatted input to read both data values:

```
input item $10. +1 amount comma5.;
```

To read this data correctly with formatted input, the second data value must occur after the 10th column of the first value, as shown here:

```
----+----1----+----2
trucks    1,382
vans      1,235
sedans    2,391
```

Also, after the value of ITEM is read with formatted input, you must use the pointer control +1 to move the pointer to the column where the value AMOUNT begins.

When Data Contains Quotation Marks When you use the DSD option in an INFILE statement, which sets the delimiter to a comma, the INPUT statement removes quotation marks before a value is written to a variable. When you also use the tilde (~) modifier in an INPUT statement, the INPUT statement maintains quotation marks as part of the value.

Examples

Example 1: Reading Unaligned Data with Simple List Input The INPUT statement in this DATA step uses simple list input to read the input data records:

```
data scores;
   input name $ score1 score2 score3 team $;
   datalines;
Joe 11 32 76 red
Mitchel 13 29 82 blue
Susan 14 27 74 green
;
```

The next INPUT statement reads only the first four fields in the previous data lines, which demonstrates that you are not required to read all the fields in the record:

```
input name $ score1 score2 score3;
```

Example 2: Reading Character Data That Contains Embedded Blanks The INPUT statement in this DATA step uses the & format modifier with list input to read character values that contain embedded blanks.

```
data list;
   infile file-specification;
   input name $ & score;
run;
```

It can read these input data records:

```
----+----1----+----2----+----3----+
Joseph   11 Joergensen  red
Mitchel  13 Mc Allister  blue
Su Ellen  14 Fischer-Simon  green
```

The & modifier follows the variable it affects in the INPUT statement. Because this format modifier follows NAME, at least two blanks must separate the NAME field from the SCORE field in the input data records.

You can also specify an informat with a format modifier, as shown here:

```
    input name $ & +3 lastname & $15. team $;
```

In addition, this INPUT statement reads the same data to demonstrate that you are not required to read all the values in an input record. The +3 column pointer control moves the pointer past the score value in order to read the value for LASTNAME and TEAM.

Example 3: Reading Unaligned Data with Informats This DATA step uses modified list input to read data values with an informat:

```
data jansales;
   input item : $10. amount;
   datalines;
trucks 1382
vans 1235
sedans 2391
;
```

The $10. informat allows a character variable of up to ten characters to be read.

Example 4: Reading Comma-Delimited Data with List Input and an Informat This DATA step uses the DELIMITER= option in the INFILE statement to read list input values that are separated by commas instead of blanks. The example uses an informat to read the date, and a format to write the date.

```
options pageno=1 nodate ls=80 ps=64;
data scores2;
   length Team $ 14;
   infile datalines delimiter=',';
   input Name $ Score1-Score3 Team $ Final_Date:MMDDYY10.;
   format final_date weekdate17.;
   datalines;
Joe,11,32,76,Red Racers,2/3/2007
Mitchell,13,29,82,Blue Bunnies,4/5/2007
Susan,14,27,74,Green Gazelles,11/13/2007
;
```

```
proc print data=scores2;
   var Name Team Score1-Score3 Final_Date;
   title 'Soccer Player Scores';
run;
```

Output 6.14 Output from Comma-Delimited Data

```
                              Soccer Player Scores                              1

Obs    Name            Team          Score1    Score2    Score3          Final_Date

 1     Joe         Red Racers           11        32        76       Mon, Feb 3, 2007
 2     Mitchell    Blue Bunnies         13        29        82       Sat, Apr 5, 2007
 3     Susan       Green Gazelles       14        27        74      Thu, Nov 13, 2007
```

Example 5: Reading Delimited Data with Modified List Input This DATA step uses the DSD option in an INFILE statement and the tilde (~) format modifier in an INPUT statement to retain the quotation marks in character data and to read a character in a string that is enclosed in quotation marks as a character instead of as a delimiter.

```
data scores;
   infile datalines dsd;
   input Name : $9. Score1-Score3
         Team ~ $25. Div $;
   datalines;
Joseph,11,32,76,"Red Racers, Washington",AAA
Mitchel,13,29,82,"Blue Bunnies, Richmond",AAA
Sue Ellen,14,27,74,"Green Gazelles, Atlanta",AA
;
```

The output that PROC PRINT generates shows the resulting SCORES data set. The values for TEAM contain the quotation marks.

Output 6.15 SCORES Data Set

```
                          The SAS System                          1

OBS Name      Score1 Score2 Score3           Team             Div

 1  Joseph       11     32     76   "Red Racers, Washington"  AAA
 2  Mitchel      13     29     82   "Blue Bunnies, Richmond"  AAA
 3  Sue Ellen    14     27     74   "Green Gazelles, Atlanta" AA
```

See Also

Statements:

"INFILE Statement" on page 1591

"INPUT Statement" on page 1617

"INPUT Statement, Formatted" on page 1635

INPUT Statement, Named

Reads data values that appear after a variable name that is followed by an equal sign and assigns them to corresponding SAS variables.

Valid: in a DATA step

Category: File-handling

Type: Executable

Syntax

INPUT <*pointer-control*> *variable*= <$> <@ | @@>;

INPUT <*pointer-control*> *variable*= *informat.* <@ | @@>;

INPUT *variable*= <$> *start-column* <*-end-column*>
<*.decimals*> <@ | @@>;

Arguments

pointer-control
moves the input pointer to a specified line or column in the input buffer.

See: "Column Pointer Controls" on page 1619 and "Line Pointer Controls" on page 1621

variable=
specifies a variable whose value is read by the INPUT statement. In the input data record, the field has the form

```
variable=value
```

Featured in: Example 3 on page 1647

$
indicates to store a variable value as a character value rather than as a numeric value.

Tip: If the variable is previously defined as character, $ is not required.

Featured in: Example 3 on page 1647

informat.
specifies an informat that indicates the data type of the input values, but not how the values are read.

Tip: Use the INFORMAT statement to associate an informat with a variable.

See: Chapter 5, "Informats," on page 1257

Featured in: Example 3 on page 1647

start-column
specifies the column that the INPUT statement uses to begin scanning in the input data records for the variable. The variable name does not have to begin here.

-end-column
determines the default length of the variable.

@
holds an input record for the execution of the next INPUT statement within the same iteration of the DATA step. This line-hold specifier is called *trailing* @.

Restriction: The trailing @ must be the last item in the INPUT statement.

Tip: The trailing @ prevents the next INPUT statement from automatically releasing the current input record and reading the next record into the input buffer. It is useful when you need to read from a record multiple times.

See: "Using Line-Hold Specifiers" on page 1624

@@
holds an input record for the execution of the next INPUT statement across iterations of the DATA step. This line-hold specifier is called *double trailing* @.

Restriction: The double trailing @ must be the last item in the INPUT statement.

Tip: The double trailing @ is useful when each input line contains values for several observations.

See: "Using Line-Hold Specifiers" on page 1624

Details

When to Use Named Input Named input reads the input data records that contain a variable name followed by an equal sign and a value for the variable. The INPUT statement reads the input data record at the current location of the input pointer. If the input data records contain data values at the start of the record that the INPUT statement cannot read with named input, use another input style to read them. However, once the INPUT statement starts to read named input, SAS expects that all the remaining values are in this form. See Example 3 on page 1647.

You do not have to specify the variables in the INPUT statement in the same order that they occur in the data records. Also, you do not have to specify a variable for each field in the record. However, if you do not specify a variable in the INPUT statement that another statement uses (for example, ATTRIB, FORMAT, INFORMAT, LENGTH statement) and it occurs in the input data record, the INPUT statement automatically reads the value. SAS writes a note to the log that the variable is uninitialized.

When you do not specify a variable for all the named input data values, SAS sets _ERROR_ to 1 and writes a note to the log. For example,

```
data list;
   input name=$ age=;
   datalines;
name=John age=34  gender=M
;
```

The note that SAS writes to the log states that GENDER is not defined and _ERROR_ is set to 1.

Restrictions

- After you start to read with named input, you cannot switch to another input style or use pointer controls. All the remaining values in the input data record must be in the form *variable=value*. SAS treats the values that are not in named input form as invalid data.

- If named input values continue after the end of the current input line, use a slash (/) at the end of the input line. The slash tells SAS to move the pointer to the next line and to continue to read with named input. For example,

  ```
  input name=$ age=;
  ```

  can read this input data record:

  ```
  name=John /
     age=34
  ```

- If you use named input to read character values that contain embedded blanks, put two blanks before and after the data value, as you would with list input. See Example 4 on page 1648.
- You cannot reference an array with an asterisk or an expression subscript.

Examples

Example 1: Using List and Named Input This DATA step uses list input with named input to read input data records.

```
data list;
   length name $ 20 gender $ 1;
   informat dob ddmmyy8.;
   input id name= gender= age= dob=;
   datalines;
4798 name=COLIN gender=m age=23 dob=16/02/75
2653 name=MICHELE gender=f age=46 dob=17/02/73
;
proc print data=list; run;
```

The INPUT statement uses list input to read the ID variable. The remaining variables NAME, GENDER, AGE, and DOB are read with named input. The LENGTH statement prevents the INPUT statement from truncating the character values for the variable name to a length of eight.

Example 2: Using Named Input with Variables in Random Order Using the same data as in the previous example, this DATA step also uses list input and named input to read input data records. However, in this example, the order of the values in the data is different for the two rows, except for the ID value, which must come first.

```
data list;
   length name $ 20 gender $ 1;
   informat dob ddmmyy8.;
   input id dob= name= age= gender=;
   datalines;
4798 gender=m name=COLIN age=23 dob=16/02/75
2653 name=MICHELE dob=17/02/73 age=46 gender=f
;
proc print data=list; run;
```

Example 3: Using Named Input with Another Input Style This DATA step uses list input and named input to read input data records:

```
data list;
   input id name=$20. gender=$;
   informat dob ddmmyy8.;
   datalines;
4798  gender=m name=COLIN age=23 dob=16/02/75
```

```
2653 name=MICHELE age=46 gender=f
;
proc print data=list; run;
```

The INPUT statement uses list input to read the first variable, ID. The remaining variables NAME, GENDER, and DOB are read with named input. These variables are not read in order. The $20. informat with NAME= prevents the INPUT statement from truncating the character value to a length of eight. The INPUT statement reads the DOB= field because the INFORMAT statement refers to this variable. It skips the AGE= field altogether. SAS writes notes to the log that DOB is uninitialized, AGE is not defined, and _ERROR_ is set to 1.

Example 4: Reading Character Variables with Embedded Blanks This DATA step reads character variables that contain embedded blanks with named input:

```
data list2;
   informat header $30. name $15.;
   input header= name=;
   datalines;
header=  age=60 AND UP  name=PHILIP
;
```

Two spaces precede and follow the value of the variable HEADER, which is **AGE=60 AND UP**. The field also contains an equal sign.

See Also

Statement:

"INPUT Statement" on page 1617

KEEP Statement

Specifies the variables to include in output SAS data sets.

Valid: in a DATA step

Category: Information

Type: Declarative

Syntax

KEEP *variable-list*;

Arguments

variable-list
specifies the names of the variables to write to the output data set.

Tip: List the variables in any form that SAS allows.

Details

The KEEP statement causes a DATA step to write only the variables that you specify to one or more SAS data sets. The KEEP statement applies to all SAS data sets that are created within the same DATA step and can appear anywhere in the step. If no KEEP or DROP statement appears, all data sets that are created in the DATA step contain all variables.

Note: Do not use both the KEEP and DROP statements within the same DATA step. △

Comparisons

- The KEEP *statement* cannot be used in SAS PROC steps. The KEEP= *data set option* can.
- The KEEP *statement* applies to all output data sets that are named in the DATA statement. To write different variables to different data sets, you must use the KEEP= *data set option*.
- The DROP statement is a parallel statement that specifies variables to omit from the output data set.
- The KEEP and DROP statements select variables to include in or exclude from output data sets. The subsetting IF statement selects observations.
- Do not confuse the KEEP statement with the RETAIN statement. The RETAIN statement causes SAS to hold the value of a variable from one iteration of the DATA step to the next iteration. The KEEP statement does not affect the value of variables but only specifies which variables to include in any output data sets.

Examples

- These examples show the correct syntax for listing variables in the KEEP statement:
  - `keep name address city state zip phone;`
  - `keep rep1-rep5;`
- This example uses the KEEP statement to include only the variables NAME and AVG in the output data set. The variables SCORE1 through SCORE20, from which AVG is calculated, are not written to the data set AVERAGE.

```
data average;
   keep name avg;
   infile file-specification;
   input name $ score1-score20;
   avg=mean(of score1-score20);
run;
```

See Also

Data Set Option:

"KEEP= Data Set Option" on page 36

Statements:

"DROP Statement" on page 1499

"IF Statement, Subsetting" on page 1581

"RETAIN Statement" on page 1747

LABEL Statement

Assigns descriptive labels to variables.

Valid: in a DATA step

Category: Information

Type: Declarative

Syntax

LABEL *variable-1=label-1 . . . <variable-n=label-n>*;

LABEL *variable-1=' ' … <variable-n=' '>*;

Arguments

variable
specifies the variable that you want to label.

Tip: You can specify additional pairs of labels and variables.

label
specifies a label of up to 256 characters, including blanks.

Tip: You can specify additional pairs of labels and variables.

Tip: For more information about including quotation marks as part of the label, see "Character Constants" in *SAS Language Reference: Concepts*.

Restriction: If the label includes a semicolon (;) or an equal sign (=), you must enclose the label in either single or double quotation marks.

Restriction: If the label includes single quotation marks ('), then you must enclose the label in double quotation marks.

' '
removes a label from a variable. Enclose a single blank space in quotation marks to remove an existing label.

Details

Using a LABEL statement in a DATA step permanently associates labels with variables by affecting the descriptor information of the SAS data set that contains the variables. You can associate any number of variables with labels in a single LABEL statement.

You can use a LABEL statement in a PROC step, but the rules are different. See the *Base SAS Procedures Guide* for more information.

Comparisons

Both the ATTRIB and LABEL statements can associate labels with variables and change a label that is associated with a variable.

Examples

Example 1: Specifying Labels Here are several LABEL statements:

```
label compound=Type of Drug;
label date="Today's Date";
label n='Mark''s Experiment Number';
label score1="Grade on April 1 Test"
      score2="Grade on May 1 Test";
```

Example 2: Removing a Label This example removes an existing label:

```
data rtest;
   set rtest;
   label x=' ';
run;
```

See Also

Statement:

"ATTRIB Statement" on page 1448

Labels, Statement

Identifies a statement that is referred to by another statement.

Valid: in a DATA step

Category: Control

Type: Declarative

Syntax

label: *statement*;

Arguments

label

specifies any SAS name, which is followed by a colon (:). You must specify the *label* argument.

statement

specifies any executable statement, including a null statement (;). You must specify the *statement* argument.

Restriction: No two statements in a DATA step can have the same label.

Restriction: If a statement in a DATA step is labeled, it should be referenced by a statement or option in the same step.

Tip: A null statement can have a label:

```
ABC:;
```

Details

The statement label identifies the destination of either a GO TO statement, a LINK statement, the HEADER= option in a FILE statement, or the EOF= option in an INFILE statement.

Comparisons

The LABEL statement assigns a descriptive label to a variable. A statement label identifies a statement or group of statements that are referred to in the same DATA step by another statement, such as a GO TO statement.

Examples

In this example, if Stock=0, the GO TO statement causes SAS to jump to the statement that is labeled reorder. When Stock is not 0, execution continues to the RETURN statement and then returns to the beginning of the DATA step for the next observation.

```
data Inventory Order;
   input Item $ Stock @;
      /* go to label reorder: */
   if Stock=0 then go to reorder;
   output Inventory;
   return;
      /* destination of GO TO statement */
   reorder: input Supplier $;
   put 'ORDER ITEM ' Item 'FROM ' Supplier;
   output Order;
   datalines;
milk  0 A
bread 3 B
;
```

See Also

Statements:

"GO TO Statement" on page 1579

"LINK Statement" on page 1669

Statement Options:

HEADER= option in the FILE statement on page 1509

EOF= option in the INFILE statement on page 1595

LEAVE Statement

Stops processing the current loop and resumes with the next statement in the sequence.

Valid: in a DATA step

Category: Control

Type: Executable

Syntax

LEAVE;

Without Arguments

The LEAVE statement stops the processing of the current DO loop or SELECT group and continues DATA step processing with the next statement following the DO loop or SELECT group.

Details

You can use the LEAVE statement to exit a DO loop or SELECT group prematurely based on a condition.

Comparisons

- The LEAVE statement causes processing of the current loop to end. The CONTINUE statement stops the processing of the current iteration of a loop and resumes with the next iteration.
- You can use the LEAVE statement in a DO loop or in a SELECT group. You can use the CONTINUE statement only in a DO loop.

Examples

This DATA step demonstrates using the LEAVE statement to stop the processing of a DO loop under a given condition. In this example, the IF/THEN statement checks the value of BONUS. When the value of BONUS reaches 500, the maximum amount allowed, the LEAVE statement stops the processing of the DO loop.

```
data week;
   input name $ idno start_yr status $ dept $;
   bonus=0;
   do year= start_yr to 1991;
      if bonus ge 500 then leave;
      bonus+50;
   end;
   datalines;
Jones 9011 1990 PT PUB
Thomas 876 1976 PT HR
Barnes 7899 1991 FT TECH
Harrell 1250 1975 FT HR
Richards 1002 1990 FT DEV
```

```
Kelly 85 1981 PT PUB
Stone 091 1990 PT MAIT
;
```

LENGTH Statement

Specifies the number of bytes for storing variables.

Valid: in a DATA step

Category: Information

Type: Declarative

See: LENGTH Statement in the documentation for your operating environment.

Syntax

LENGTH *variable-specification(s)*<DEFAULT=*n*>;

Arguments

variable-specification
is a required argument and has the form
variable(s)<$>*length*
where

variable
specifies one or more variables that are to be assigned a length. This includes any variables in the DATA step, including those dropped from the output data set.

Restriction: Array references are not allowed.

Tip: If the variable is character, the length applies to the program data vector and the output data set. If the variable is numeric, the length applies only to the output data set.

$
specifies that the preceding variables are character variables.

Default: SAS assumes that the variables are numeric.

length
specifies a numeric constant that is the number of bytes used for storing variable values.

Range: For numeric variables, 2 to 8 or 3 to 8, depending on your operating environment. For character variables, 1 to 32767 under all operating environments.

DEFAULT=*n*
changes the default number of bytes that SAS uses to store the values of any newly created numeric variables.

Default: 8

Range: 2 to 8 or 3 to 8, depending on your operating environment.

CAUTION:
Avoid shortening numeric variables that contain fractions. The precision of a numeric variable is closely tied to its length, especially when the variable contains fractional values. You can safely shorten variables that contain integers according to the rules that are given in the SAS documentation for your operating environment, but shortening variables that contain fractions might eliminate important precision. △

Details

In general, the length of a variable depends on

- whether the variable is numeric or character
- how the variable was created
- whether a LENGTH or ATTRIB statement is present.

Subject to the rules for assigning lengths, lengths that are assigned with the LENGTH statement can be changed in the ATTRIB statement and vice versa. See "SAS Variables" in *SAS Language Reference: Concepts* for information about assigning lengths to variables.

Operating Environment Information: Valid variable lengths depend on your operating environment. For details, see the SAS documentation for your operating environment. △

Comparisons

The ATTRIB statement can assign the length as well as other attributes of variables.

Examples

This example uses a LENGTH statement to set the length of the character variable NAME to 25. It also changes the default number of bytes that SAS uses to store the values of newly created numeric variables from 8 to 4. The TRIM function removes trailing blanks from LASTNAME before it is concatenated with a comma (,) , a blank space, and the value of FIRSTNAME. If you omit the LENGTH statement, SAS sets the length of NAME to 32.

```
data testlength;
   informat FirstName LastName $15. n1 6.2;
   input firstname lastname n1 n2;
   length name $25 default=4;
   name=trim(lastname)||', '||firstname;
   datalines;
Alexander Robinson 35 11
;

proc contents data=testlength;
run;

proc print data=testlength;
run;
```

The following output shows a partial listing from PROC CONTENTS, as well as the report that PROC PRINT generates.

Output 6.16 Setting the Length of a Variable

```
                          The SAS System                          3

                       CONTENTS PROCEDURE

      -----Alphabetic List of Variables and Attributes-----

        #    Variable     Type    Len     Pos    Informat
        ------------------------------------------------
        1    FirstName    Char     15       8    $15.
        2    LastName     Char     15      23    $15.
        3    n1           Num       4       0    6.2
        4    n2           Num       4       4
        5    name         Char     25      38
```

```
                          The SAS System                          4

OBS   FirstName   LastName      n1      n2          name

 1    Alexander   Robinson   0.35000   11   Robinson, Alexander
```

See Also

Statement:

"ATTRIB Statement" on page 1448

For information about the use of the LENGTH statement in PROC steps, see *Base SAS Procedures Guide*

LIBNAME Statement

Associates or disassociates a SAS library with a libref (a shortcut name), clears one or all librefs, lists the characteristics of a SAS library, concatenates SAS libraries, or concatenates SAS catalogs.

Valid: Anywhere

Category: Data Access

See: LIBNAME Statement in the documentation for your operating environment

Syntax

❶**LIBNAME** *libref* <*engine*> '*SAS-library*'
< *options* > <*engine*/*host-options*>;

❷**LIBNAME** *libref* CLEAR | _ALL_ CLEAR;

❸**LIBNAME** *libref* LIST | _ALL_ LIST;

❹❺**LIBNAME** *libref* <*engine*> (*library-specification-1* <. . . *library-specification-n*>)
< *options* >;

Arguments

libref
is a shortcut name or a "nickname" for the aggregate storage location where your SAS files are stored. It is any SAS name when you are assigning a new libref. When you are disassociating a libref from a SAS library or when you are listing attributes, specify a libref that was previously assigned.

Range: 1 to 8 characters

Tip: The association between a libref and a SAS library lasts only for the duration of the SAS session or until you change it or discontinue it with another LIBNAME statement.

'SAS-library'
must be the physical name for the SAS library. The physical name is the name that is recognized by the operating environment. Enclose the physical name in single or double quotation marks.

Operating Environment Information: For details about specifying the physical names of files, see the SAS documentation for your operating environment. △

library-specification
is two or more SAS libraries that are specified by physical names, previously assigned librefs, or a combination of the two. Separate each specification with either a blank or a comma and enclose the entire list in parentheses.

'SAS-library'
is the physical name of a SAS library, enclosed in quotation marks.

libref
is the name of a previously assigned libref.

Restriction: When concatenating libraries, you cannot specify options that are specific to an engine or an operating environment.

Featured in: Example 2 on page 1663

See Also: "Rules for Library Concatenation" on page 1662

engine
is an engine name.

Tip: Usually, SAS automatically determines the appropriate engine to use for accessing the files in the library. If you want to create a new library with an engine other than the default engine, then you can override the automatic selection.

See: For a list of valid engines, see the SAS documentation for your operating environment. For background information about engines, see *SAS Language Reference: Concepts*.

CLEAR
disassociates one or more currently assigned librefs.

Tip: Specify *libref* to disassociate a single libref. Specify _ALL_ to disassociate all currently assigned librefs.

ALL
specifies that the CLEAR or LIST argument applies to all currently assigned librefs.

LIST
writes the attributes of one or more SAS libraries to the SAS log.

Tip: Specify *libref* to list the attributes of a single SAS library. Specify _ALL_ to list the attributes of all SAS libraries that have librefs in your current session.

Options

ACCESS=READONLY | TEMP

READONLY
: assigns a read-only attribute to an entire SAS library. SAS will not allow you to open a data set in the library in order to update information or write new information.

TEMP
: specifies that the SAS library be treated as a scratch library. That is, the system will not consume CPU cycles to ensure that the files in a TEMP library do not become corrupted.

  Tip: Use ACCESS=TEMP to save resources only when the data is recoverable.

Operating Environment Information: Some operating environments support LIBNAME statement options that have similar functions to the ACCESS= option. See the SAS documentation for your operating environment. △

COMPRESS=NO | YES | CHAR | BINARY
: controls the compression of observations in output SAS data sets for a SAS library.

  NO
  : specifies that the observations in a newly created SAS data set be uncompressed (fixed-length records).

  YES | CHAR
  : specifies that the observations in a newly created SAS data set be compressed (variable-length records) by SAS using RLE (Run Length Encoding). RLE compresses observations by reducing repeated consecutive characters (including blanks) to two-byte or three-byte representations.

    Tip: Use this compression algorithm for character data.

  BINARY
  : specifies that the observations in a newly created SAS data set be compressed (variable-length records) by SAS using RDC (Ross Data Compression). RDC combines run-length encoding and sliding-window compression to compress the file.

    Tip: This method is highly effective for compressing medium to large (several hundred bytes or larger) blocks of binary data (numeric variables). Because the compression function operates on a single record at a time, the record length needs to be several hundred bytes or larger for effective compression.

  Interaction: For the COPY procedure, the default value CLONE uses the compression attribute from the input data set for the output data set instead of the value specified in the COMPRESS= option. For more information about CLONE and NOCLONE, see the COPY statement in the DATASETS procedure in the *Base SAS Procedures Guide*. This interaction does not apply when using SAS/SHARE or SAS/CONNECT.

CVPBYTES=*bytes*
: specifies the number of bytes to expand character variable lengths when processing a SAS data file that requires transcoding.

  See: "CVPBYTES=, CVPENGINE=, and CVPMULTIPLIER= Options" in the *SAS National Language Support (NLS): Reference Guide*

CVPENGINE | CVPENG=*engine*
: specifies the engine to use in order to process a SAS data file that requires transcoding.

See: "CVPBYTES=, CVPENGINE=, and CVPMULTIPLIER= Options" in the *SAS National Language Support (NLS): Reference Guide*

CVPMULTIPLIER|CVPMULT=*multiplier*
specifies a multiplier value in order to expand character variable lengths when processing a SAS data file that requires transcoding.

See: "CVPBYTES=, CVPENGINE=, and CVPMULTIPLIER= Options" in the *SAS National Language Support (NLS): Reference Guide*

INENCODING=ANY | ASCIIANY | EBCDICANY | *encoding-value*
overrides the encoding when you are reading (input processing) SAS data sets in the SAS library.

See: "INENCODING= and OUTENCODING= Options" in the *SAS National Language Support (NLS): Reference Guide*

OUTENCODING=
OUTENCODING=ANY | ASCIIANY | EBCDICANY | *encoding-value*
overrides the encoding when you are creating (output processing) SAS data sets in the SAS library.

See: The "INENCODING= and OUTENCODING= Options" in the *SAS National Language Support (NLS): Reference Guide*

OUTREP=*format*
specifies the data representation for the SAS library, which is the form in which data is stored in a particular operating environment. Different operating environments use different standards or conventions for storing floating-point numbers (for example, IEEE or IBM Mainframe); for character encoding (ASCII or EBCDIC); for the ordering of bytes in memory (big Endian or little Endian); for word alignment (4-byte boundaries or 8-byte boundaries); for integer data-type length (16-bit, 32-bit, or 64-bit); and for doubles (byte-swapped or not).

Native data representation refers to an environment in which the data representation is comparable to the CPU that is accessing the file. For example, a file that is in Windows data representation is native to the Windows operating environment.

By default, SAS creates a new SAS data set by using the native data representation of the CPU that is running SAS. Specifying the OUTREP= option enables you to create files within the native environment that use a foreign data representation. For example, in a UNIX environment, you can create a SAS data set that uses a Windows data representation. Existing data sets that are written to the library are given the new data representation.

Interaction: For the COPY procedure, the default value CLONE uses the data representation from the input data set instead of the value specified in the OUTREP= option. For more information about CLONE and NOCLONE, see the COPY statement in the DATASETS procedure in the *Base SAS Procedures Guide*. This interaction does not apply when using SAS/SHARE or SAS/CONNECT.

Interaction: The COPY procedure (with NOCLONE) and the MIGRATE procedure can use the LIBNAME option OUTREP= for DATA, VIEW, ACCESS, MDDB and DMDB member types. Otherwise, only DATA member types are affected by the OUTREP= LIBNAME option.

Interaction: Transcoding could result in character data loss when encodings are incompatible. For information about encoding and transcoding, see *SAS National Language Support (NLS): Reference Guide*.

Values for OUTREP= are listed in the following table:

Table 6.7 Data Representation Values for OUTREP= Option

OUTREP= Value	Alias*	Environment
ALPHA_TRU64	ALPHA_OSF	Compaq Tru64 UNIX
ALPHA_VMS_32	ALPHA_VMS	OpenVMS on Alpha
ALPHA_VMS_64		OpenVMS on Alpha
HP_IA64	HP_ITANIUM	HP UX on Itanium 64-bit platform
HP_UX_32	HP_UX	HP UX on 32-bit platform
HP_UX_64		HP UX on 64-bit platform
INTEL_ABI		ABI UNIX on Intel 32-bit platform
LINUX_32	LINUX	Linux for Intel Architecture on 32-bit platform
LINUX_IA64		Linux for Itanium-based system on 64-bit platform
LINUX_X86_64		LINUX on x64 64-bit platform
MIPS_ABI		ABI UNIX on 32-bit platform
MVS_32	MVS	z/OS on 32-bit platform
OS2		OS/2 on Intel 32-bit platform
RS_6000_AIX_32	RS_6000_AIX	AIX UNIX on 32-bit RS/6000
RS_6000_AIX_64		AIX UNIX on 64-bit RS/6000
SOLARIS_32	SOLARIS	Solaris on SPARC 32-bit platform
SOLARIS_64		Solaris on SPARC 64-bit platform
SOLARIS_X86_64		Solaris on x64 64-bit platform
VAX_VMS		OpenVMS VAX
VMS_IA64		OpenVMS for HP Integrity servers 64-bit platform
WINDOWS_32	WINDOWS	Microsoft Windows on 32-bit platform
WINDOWS_64		Microsoft Windows 64-bit Edition (for both Itanium-based systems and x64)

* It is recommended that you use the current values. The aliases are available for compatibility only.

REPEMPTY=YES | NO
controls replacement of like-named temporary or permanent SAS data sets when the new one is empty.

YES
specifies that a new empty data set with a given name replace an existing data set with the same name. This is the default.

Interaction: When REPEMPTY=YES and REPLACE=NO, then the data set is not replaced.

NO
specifies that a new empty data set with a given name not replace an existing data set with the same name.

Tip: Use REPEMPTY=NO to prevent the following syntax error from replacing the existing data set MYLIB.B with the new empty data set MYLIB.B that is created by mistake:

```
libname libref SAS-library REPEMPTY=NO;
data mylib.a set mylib.b;
```

Tip: For both the convenience of replacing existing data sets with new ones that contain data and the protection of not overwriting existing data sets with new empty ones that are created by mistake, set REPLACE=YES and REPEMPTY=NO.

Comparison: For an individual data set, the REPEMPTY= data set option overrides the setting of the REPEMPTY= option in the LIBNAME statement.

See Also: "REPEMPTY= Data Set Option" on page 54

Engine Host Options

engine-host-options
are one or more options that are listed in the general form *keyword=value*.

Operating Environment Information: For a list of valid specifications, see the SAS documentation for your operating environment. △

Restriction: When concatenating libraries, you cannot specify options that are specific to an engine or an operating environment.

Details

❶Associating a Libref with a SAS Library The association between a libref and a SAS library lasts only for the duration of the SAS session or until you change the libref or discontinue it with another LIBNAME statement. The simplest form of the LIBNAME statement specifies only a libref and the physical name of a SAS library:

LIBNAME *libref 'SAS-library'*;

See Example 1 on page 1663.

An engine specification is usually not necessary. If the situation is ambiguous, SAS uses the setting of the ENGINE= system option to determine the default engine. If all data sets in the library are associated with a single engine, then SAS uses that engine as the default. In either situation, you can override the default by specifying another engine with the ENGINE= system option:

LIBNAME *libref engine 'SAS-library'*
<*options* > <*engine*/*host-options*>;

Operating Environment Information: Using the LIBNAME statement requires host-specific information. See the SAS documentation for your operating environment before using this statement. △

❷Disassociating a Libref from a SAS Library To disassociate a libref from a SAS library, use a LIBNAME statement by specifying the libref and the CLEAR option. You can clear a single, specified libref or all current librefs.

LIBNAME *libref* CLEAR | _ALL_ CLEAR;

❸Writing SAS Library Attributes to the SAS Log Use a LIBNAME statement to write the attributes of one or more SAS libraries to the SAS log. Specify *libref* to list the attributes of one SAS library; use _ALL_ to list the attributes of all SAS libraries that have been assigned librefs in your current SAS session.

LIBNAME *libref* LIST | _ALL_ LIST;

❹Concatenating SAS Libraries When you logically concatenate two or more SAS libraries, you can reference them all with one libref. You can specify a library with its physical filename or its previously assigned libref.

LIBNAME *libref* <*engine*> (*library–specification–1* <. . . *library-specification-n*>) < *options* >;

In the same LIBNAME statement you can use any combination of specifications: librefs, physical filenames, or a combination of librefs and physical filenames. See Example 2 on page 1663.

❺Concatenating SAS Catalogs When you logically concatenate two or more SAS libraries, you also concatenate the SAS catalogs that have the same name. For example, if three SAS libraries each contain a catalog named CATALOG1, then when you concatenate them, you create a catalog concatenation for the catalogs that have the same name. See Example 3 on page 1664.

LIBNAME *libref* <*engine*> (*library–specification–1* <. . . *library-specification-n*>) < *options* >;

Rules for Library Concatenation After you create a library concatenation, you can specify the libref in any context that accepts a simple (non-concatenated) libref. These rules determine how SAS files (that is, members of SAS libraries) are located among the concatenated libraries:

1. When a SAS file is opened for input or update, the concatenated libraries are searched and the first occurrence of the specified file is used.
2. When a SAS file is opened for output, it is created in the first library that is listed in the concatenation.

   Note: A new SAS file is created in the first library even if there is a file with the same name in another part of the concatenation. △
3. When you delete or rename a SAS file, only the first occurrence of the file is affected.
4. Anytime a list of SAS files is displayed, only one occurrence of a filename is shown.

   Note: Even if the name occurs multiple times in the concatenation, only the first occurrence is shown. △
5. A SAS file that is logically connected to another file (such as an index to a data set) is listed only if the parent file resides in that same library. For example, if library ONE contains A.DATA, and library TWO contains A.DATA and A.INDEX, only A.DATA from library ONE is listed. (See rule 4.)
6. If any library in the concatenation is sequential, then all of the libraries are treated as sequential.
7. The attributes of the first library that is specified determine the attributes of the concatenation. For example, if the first SAS library that is listed is “read only,” then the entire concatenated library is “read only.”
8. If you specify any options or engines, they apply only to the libraries that you specified with the complete physical name, not to any library that you specified with a libref.

9 If you alter a libref after it has been assigned in a concatenation, it will not affect the concatenation.

Comparisons

- Use the LIBNAME statement to reference a SAS library. Use the FILENAME statement to reference an external file. Use the LIBNAME, SAS/ACCESS statement to access DBMS tables.
- Use the CATNAME statement to concatenate SAS catalogs. Use the LIBNAME statement to concatenate SAS catalogs. The CATNAME statement enables you to specify the names of the catalogs that you want to concatenate. The LIBNAME statement concatenates all like-named catalogs in the specified SAS libraries.

Examples

Example 1: Assigning and Using a Libref This example assigns the libref SALES to an aggregate storage location that is specified in quotation marks as a physical filename. The DATA step creates SALES.QUARTER1 and stores it in that location. The PROC PRINT step references it by its two-level name, SALES.QUARTER1.

```
libname sales 'SAS-library';

data sales.quarter1;
   infile 'your-input-file';
   input salesrep $20. +6 jansales febsales
         marsales;
run;

proc print data=sales.quarter1;
run;
```

Example 2: Logically Concatenating SAS Libraries

- This example concatenates three SAS libraries by specifying the physical filename of each:

```
libname allmine ('file-1' 'file-2'
                 'file-3');
```

- This example assigns librefs to two SAS libraries, one that contains SAS 6 files and one that contains SAS 9 files. This technique is useful for updating your files and applications from SAS 6 to SAS 9, while allowing you to have convenient access to both sets of files:

```
libname v6 'v6--SAS-library';
libname v9 'v9--SAS-library';
libname allmine (v9 v6);
```

- This example shows that you can specify both librefs and physical filenames in the same concatenation specification:

```
libname allmine (v9 v6 'some-filename');
```

Example 3: Concatenating SAS Catalogs This example concatenates three SAS libraries by specifying the physical filename of each and assigns the libref ALLMINE to the concatenated libraries:

```
libname allmine ('file-1' 'file-2'
                 'file-3');
```

If each library contains a SAS catalog named MYCAT, then using ALLMINE.MYCAT as a libref.catref provides access to the catalog entries that are stored in all three catalogs named MYCAT. To logically concatenate SAS catalogs with different names, see "CATNAME Statement" on page 1459.

Example 4: Permanently Storing Data Sets with One-Level Names If you want the convenience of specifying only a one-level name for permanent, not temporary, SAS files, then use the USER= system option. This example stores the data set QUARTER1 permanently without using a LIBNAME statement first to assign a libref to a storage location:

```
options user='SAS-library';

data quarter1;
   infile 'your-input-file';
   input salesrep $20. +6 jansales febsales
         marsales;
run;

proc print data=quarter1;
run;
```

See Also

Data Set Options:

ENCODING in the *SAS National Language Support (NLS): Reference Guide*

Statements:

- "CATNAME Statement" on page 1459 for a discussion of concatenating SAS catalogs
- "FILENAME Statement" on page 1520
- "LIBNAME Statement" for character variable processing in order to transcode a SAS file in *SAS National Language Support (NLS): Reference Guide*
- "LIBNAME Statement" for the Output Delivery System (ODS) in *SAS Output Delivery System: User's Guide*
- "LIBNAME Statement" for SAS metadata in *SAS Language Interfaces to Metadata*
- "LIBNAME Statement" for Scalable Performance Data (SPD) in *SAS Scalable Performance Data Engine: Reference*
- "LIBNAME statement" for XML documents in *SAS XML LIBNAME Engine: User's Guide*
- "LIBNAME Statement" for SAS/ACCESS in *SAS/ACCESS for Relational Databases: Reference*
- "LIBNAME Statement" for SAS/CONNECT in *SAS/CONNECT User's Guide*
- "LIBNAME Statement" for SAS/CONNECT, TCP/IP pipes in *SAS/CONNECT User's Guide*

"LIBNAME Statement" for SAS/SHARE in *SAS/SHARE User's Guide*

System Option:

"USER= System Option" on page 2042

LIBNAME Statement for WebDAV Server Access

Associates a libref with a SAS library and enables access to a WebDAV (Web-based Distributed Authoring And Versioning) server.

Valid: Anywhere

Category: Data Access

Restriction: Access to WebDAV servers is not supported on OpenVMS or z/OS.

See also: Base SAS LIBNAME Statement

Syntax

LIBNAME *libref* <*engine*> '*SAS-library*' < *options*> WEBDAV USER="*user-ID*" PASSWORD="*user-password*" *WEBDAV options*;

LIBNAME *libref* CLEAR | _ALL_ CLEAR;

LIBNAME *libref* LIST | _ALL_ LIST;

Arguments

libref
specifies a shortcut name for the aggregate storage location where your SAS files are stored.

Tip: The association between a libref and a SAS library lasts only for the duration of the SAS session or until you change it or discontinue it with another LIBNAME statement.

'*SAS-library*'
specifies the URL location (path) on a WebDAV server. The URL specifies either HTTP or HTTPS communication protocols.

Restriction: Only one data library is supported when using the WebDAV extension to the LIBNAME statement.

Requirement: When using the HTTPS communication protocol, you must use the SSL (Secure Sockets Layer) protocol that provides secure network communications. For more information, see *Encryption in SAS*.

engine
specifies the name of a valid SAS engine.

Restriction: REMOTE engines are not supported with the WebDAV options.

See: For a list of valid engines, see the SAS documentation for your operating environment.

CLEAR
disassociates one or more currently assigned librefs. When a libref using a WebDAV server is cleared, the cached files stored locally are deleted also.

Tip: Specify *libref* to disassociate a single libref. Specify _ALL_ to disassociate all currently assigned librefs.

LIST
writes the attributes of one or more SAS libraries to the SAS log.

Tip: Specify *libref* to list the attributes of a single SAS library. Specify _ALL_ to list the attributes of all SAS libraries that have librefs in your current session.

ALL
specifies that the CLEAR or LIST argument applies to all currently assigned librefs.

Options

For valid LIBNAME statement options, see "LIBNAME Statement" on page 1656.

WebDAV Specific Options

WEBDAV
specifies that the libref access a WebDAV server.

USER="*user-ID*"
specifies the user name for access to the WebDAV server. The user ID is case sensitive and it must be enclosed in single or double quotation marks.

Alias: UID

Tip: If PROMPT is specified, but USER= is not, then the user is prompted for an ID as well as a password.

PASSWORD="*user-password*"
specifies a password for the user to access the WebDAV server. The password is case sensitive and it must be enclosed in single or double quotation marks.

Alias: PWD=, PW=, PASS=

Tip: You can specify the PROMPT option instead of the PASSWORD= option.

PROMPT
specifies to prompt for the user login password, if necessary.

Interaction: If PROMPT is specified without USER=, then the user is prompted for an ID, as well as a password.

Tip: If you specify the PROMPT option, you do not need to specify the PASSWORD= option.

AUTHDOMAIN="*auth-domain*"
specifies the name of an authentication domain metadata object in order to connect to the WebDAV server. The authentication domain references credentials (user ID and password) without your having to explicitly specify the credentials. The *auth-domain* name is case sensitive, and it must be enclosed in double quotation marks.

An administrator creates authentication domain definitions while creating a user definition with the User Manager in SAS Management Console. The authentication domain is associated with one or more login metadata objects that provide access to the WebDAV server and is resolved by the BASE engine calling the SAS Metadata Server and returning the authentication credentials.

Requirement: The authentication domain and the associated login definition must be stored in a metadata repository, and the metadata server must be running in order to resolve the metadata object specification.

Interaction: If you specify AUTHDOMAIN=, you do not need to specify USER= and PASSWORD=.

See also: For complete information about creating and using authentication domains, see the discussion on credential management in the *SAS Intelligence Platform: Security Administration Guide*.

PROXY=*url*
specifies the Uniform Resource Locator (URL) for the proxy server in one of these forms:

"http://*hostname*"

"http://*hostname:port*"

LOCALCACHE="*directory name*"
specifies a directory where a temporary subdirectory is created to hold local copies of the server files. Each libref has its own unique subdirectory. If a directory is not specified, then the subdirectories are created in the SAS WORK directory. SAS deletes the temporary files when the SAS program completes.

Default: SAS WORK directory

LOCKDURATION=*n*
specifies the number of minutes that the files written through the WebDAV libref are locked. SAS unlocks the files when the SAS program successfully completes. If the SAS program fails, then the locks expire after the time allotted.

Default: 30

Data Set Options That Function Differently with a WebDAV Server

The following table lists the data set options that have different functionality when using a WebDAV server. All other data set options will function as described in the *SAS Language Reference: Dictionary*.

Table 6.8 Data Set Option Functionality with a WebDAV Server

Data Set Option	WebDAV Storage Functionality
CNTLLEV=	*LIB* locks all data sets in the library before writing the data into the local cache. All members are unlocked after the DATA step has completed and the data set has been written back to the WebDAV server. *MEM* locks the member before writing the data into the local cache. Member is unlocked after the DATA step has completed and the data has been written back to the WebDAV server. *REC* is not supported. WebDAV allows updates to the entire data set only.
FILECLOSE	The VxTAPE engine is not supported; therefore this option is ignored.
GENMAX=	This functionality is not supported because the maximum number of revisions to keep cannot be specified in the WebDAV server.
GENNUM=	This functionality is not supported in WebDAV.
IDXNAME=	Users can specify an index to use if one exists.

INDEX=	Indexes can be created in the local cache and saved on the WebDAV server.
TOBSNO=	Remote engines are not supported; therefore this option is ignored.

Details

WebDAV File Processing When accessing a WebDAV server, the file is pulled from the WebDAV server to your local disk storage for processing. When you complete the updating, the file is pushed back to the WebDAV server for storage. The file is removed from the local disk storage when it is pushed back.

Multiple Librefs to a WebDAV Library When you assign a libref to a file on a WebDAV server, the path (URL location), user ID, and password are associated with that libref. After the first libref has been assigned, the user ID and password will be validated on subsequent attempts to assign another libref to the same library.

Note: Lock errors that you typically would not see might occur if either a different user ID or the password, or both, are used in the subsequent attempt to assign a libref to the same library. △

Locked Files on a WebDAV Server In local libraries, SAS locks a file when you open it to prevent other users from altering the file while it is being read. WebDAV locks require write access to a library, and there is no concept of a read lock. In addition, WebDAV servers can go down, come back up, or go offline at any time. Consequently, SAS honors a lock request on a file on a WebDAV server only if the file is already locked by another user.

Example

The following example associates the libref **davdata** with the WebDAV directory **/users/mydir/datadir** on the WebDAV server **www.webserver.com**:

```
libname davdata v9 "https://www.webserver.com/users/mydir/datadir"
        webdav user="mydir" pw="12345";
```

See Also

Statements:

"FILENAME Statement, WebDAV Access Method" on page 1567
"LIBNAME Statement" on page 1656

LINK Statement

Directs program execution immediately to the statement label that is specified and, if followed by a RETURN statement, returns execution to the statement that follows the LINK statement.

Valid: in a DATA step

Category: Control

Type: Executable

Syntax

LINK *label*;

Arguments

label
specifies a statement label that identifies the LINK destination. You must specify the *label* argument.

Details

The LINK statement tells SAS to jump immediately to the statement label that is indicated in the LINK statement and to continue executing statements from that point until a RETURN statement is executed. The RETURN statement sends program control to the statement immediately following the LINK statement.

The LINK statement and the destination must be in the same DATA step. The destination is identified by a statement label in the LINK statement.

The LINK statement can branch to a group of statements that contain another LINK statement. This arrangement is known as nesting. To avoid infinite looping, SAS has set a default number of nested LINK statements. You can have up to 10 LINK statements with no intervening RETURN statements. When more than one LINK statement has been executed, a RETURN statement tells SAS to return to the statement that follows the last LINK statement that was executed. However, you can use the /STACK option in the DATA statement to increase the number of nested LINK statements.

Comparisons

The difference between the LINK statement and the GO TO statement is in the action of a subsequent RETURN statement. A RETURN statement after a LINK statement returns execution to the statement that follows LINK. A RETURN statement after a GO TO statement returns execution to the beginning of the DATA step, unless a LINK statement precedes GO TO. In that case, execution continues with the first statement after LINK. In addition, a LINK statement is usually used with an explicit RETURN statement, whereas a GO TO statement is often used without a RETURN statement.

When your program executes a group of statements at several points in the program, using the LINK statement simplifies coding and makes program logic easier to follow. If your program executes a group of statements at only one point in the program, using DO-group logic rather than LINK-RETURN logic is simpler.

Examples

In this example, when the value of variable TYPE is **aluv**, the LINK statement diverts program execution to the statements that are associated with the label CALCU. The program executes until it encounters the RETURN statement, which sends program execution back to the first statement that follows LINK. SAS executes the assignment statement, writes the observation, and then returns to the top of the DATA step to read the next record. When the value of TYPE is not **aluv**, SAS executes the assignment statement, writes the observation, and returns to the top of the DATA step.

```
data hydro;
   input type $ depth station $;
      /* link to label calcu: */
   if type ='aluv' then link calcu;
   date=today();
      /* return to top of step */
   return;
   calcu: if station='site_1'
      then elevatn=6650-depth;
   else if station='site_2'
      then elevatn=5500-depth;
         /* return to date=today(); */
   return;
   datalines;
aluv 523 site_1
uppa 234 site_2
aluv 666 site_2
...more data lines...
;
```

See Also

Statements:

"DATA Statement" on page 1465
"DO Statement" on page 1491
"GO TO Statement" on page 1579
"Labels, Statement" on page 1651
"RETURN Statement" on page 1752

LIST Statement

Writes to the SAS log the input data record for the observation that is being processed.

Valid: in a DATA step
Category: Action
Type: Executable

Syntax

LIST;

Without Arguments

The LIST statement causes the input data record for the observation being processed to be written to the SAS log.

Details

The LIST statement operates only on data that is read with an INPUT statement; it has no effect on data that is read with a SET, MERGE, MODIFY, or UPDATE statement.

In the SAS log, a ruler that indicates column positions appears before the first record listed.

For variable-length records (RECFM=V), SAS writes the record length at the end of the input line. SAS does not write the length for fixed-length records (RECFM=F), unless the amount of data read does not equal the record length (LRECL).

Comparisons

Action	LIST Statement	PUT Statement
Writes when	at the end of each iteration of the DATA step	immediately
Writes what	the input data records exactly as they appear	the variables or literals specified
Writes where	only to the SAS log	to the SAS log, the SAS output destination, or to any external file
Works with	INPUT statement only	any data-reading statement
Handles hexadecimal values	automatically prints a hexadecimal value if it encounters an unprintable character	represents characters in hexadecimal only when a hexadecimal format is given

Examples

Example 1: Listing Records That Contain Missing Data

This example uses the LIST statement to write to the SAS log any input records that contain missing data. Because of the #3 line pointer control in the INPUT statement, SAS reads three input records to create a single observation. Therefore, the LIST statement writes the three current input records to the SAS log each time a value for W2AMT is missing.

```
data employee;
   input ssn 1-9 #3 w2amt 1-6;
   if w2amt=. then list;
   datalines;
23456789
JAMES SMITH
356.79
345671234
Jeffrey Thomas
.
;
```

Output 6.17 Log Listing of Missing Data

```
RULE:----+----1----+----2----+----3----+----4----+----5----+----
9   345671234
10  Jeffrey Thomas
11  .
```

The numbers 9, 10, and 11 are line numbers in the SAS log.

Example 2: Listing the Record Length of Variable-Length Records

This example uses as input an external file that contains variable-length ID numbers. The RECFM=V option is specified in the INFILE statement, and the LIST statement writes the records to the SAS log. When the file has variable-length records, as indicated by the RECFM=V option in this example, SAS writes the record length at the end of each record that is listed in the SAS log.

```
data employee;
   infile 'your-external-file' recfm=v;
   input id $;
   list;
run;
```

Output 6.18 Log Listing of Variable-Length Records and Record Lengths

```
RULE:     ----+----1----+----2----+----3----+----4----+----5---
1         23456789 8
2         123456789 9
3         5555555555 10
4         345671234 9
5         2345678910 10
6         2345678 7
```

See Also

Statement:

"PUT Statement" on page 1708

%LIST Statement

Displays lines that are entered in the current session.

Valid: anywhere

Category: Program Control

Syntax

%LIST<*n* <:*m* | – *m*>>;

Without Arguments

In interactive line mode processing, if you use the %LIST statement without arguments, it displays all previously entered program lines.

Arguments

n
displays line *n*.

n–m
displays lines *n* through *m*.
Alias: *n:m*

Details

Where and When to Use The %LIST statement can be used anywhere in a SAS job except between a DATALINES or DATALINES4 statement and the matching semicolon (;) or semicolons (;;;;). This statement is useful mainly in interactive line mode sessions to display SAS program code on the monitor. It is also useful to determine lines to include when you use the %INCLUDE statement.

Interactions

CAUTION:
In all modes of execution, the SPOOL system option controls whether SAS statements are saved.
When the SPOOL system option is in effect in interactive line mode, all SAS statements and data lines are saved automatically when they are submitted. You can display them by using the %LIST statement. When NOSPOOL is in effect, %LIST cannot display previous lines. △

Examples

This %LIST statement displays lines 10 through 20:

```
%list 10-20;
```

See Also

Statement:
"%INCLUDE Statement" on page 1584
System Option:
"SPOOL System Option" on page 2004

LOCK Statement

Acquires and releases an exclusive lock on an existing SAS file.

Valid: Anywhere

Category: Program Control

Restriction: You cannot lock a SAS file that another SAS session is currently accessing (either from an exclusive lock or because the file is open).

Restriction: The LOCK statement syntax is the same whether you issue the statement in a single-user environment or in a client/server environment. However, some LOCK statement functionality applies only to a client/server environment.

Syntax

LOCK *libref*<*.member-name*<*.member-type* | *.entry-name.entry-type*>> <LIST | QUERY | SHOW | CLEAR> ;

Arguments

libref
is a name that is associated with a SAS library. The libref (library reference) must be a valid SAS name. If the libref is SASUSER or WORK, you must specify it.

Tip: In a single-user environment, you typically would not issue the LOCK statement to exclusively lock a library. To lock a library that is accessed via a multiuser SAS/SHARE server, see the LOCK statement in the *SAS/SHARE User's Guide*.

member-name
is a valid SAS name that specifies a member of the SAS library that is associated with the libref.

Restriction: The SAS file must be created before you can request a lock. For information about locking a member of a SAS library when the member does not exist, see the *SAS/SHARE User's Guide*.

member-type
is the type of SAS file to be locked. For example, valid values are DATA, VIEW, CATALOG, MDDB, and so on. The default is DATA.

entry-name
is the name of the catalog entry to be locked.

Tip: In a single-user environment, if you issue the LOCK statement to lock an individual catalog entry, the entire catalog is locked; you typically would not issue the LOCK statement to exclusively lock a catalog entry. To lock a catalog entry in a library that is accessed via a multiuser SAS/SHARE server, see the LOCK statement in the *SAS/SHARE User's Guide*.

entry-type
is the type of the catalog entry to be locked.

Tip: In a single-user environment, if you issue the LOCK statement to lock an individual catalog entry, the entire catalog is locked; you typically would not issue the LOCK statement to exclusively lock a catalog entry. To lock a catalog entry in a library that is accessed via a multiuser SAS/SHARE server, see the LOCK statement in the *SAS/SHARE User's Guide*.

LIST | QUERY | SHOW
writes to the SAS log whether you have an exclusive lock on the specified SAS file.

Tip: This option provides more information in a client/server environment. To use this option in a client/server environment, see the LOCK statement in the *SAS/SHARE User's Guide*.

CLEAR
releases a lock on the specified SAS file that was acquired by using the LOCK statement in your SAS session.

Details

General Information The LOCK statement enables you to acquire and release an exclusive lock on an existing SAS file. Once an exclusive lock is obtained, no other SAS session can read or write to the file until the lock is released. You release an exclusive lock by using the CLEAR option.

Acquiring Exclusive Access to a SAS File in a Single-User Environment Each time you issue a SAS statement or a procedure to process a SAS file, the file is opened for input, update, or output processing. At the end of the step, the file is closed. In a program with multiple tasks, a file could be opened and closed multiple times. Because multiple SAS sessions in a single-user environment can access the same SAS file, issuing the LOCK statement to acquire an exclusive lock on the file protects data while it is being updated in a multistep program.

For example, consider a nightly update process that consists of a DATA step to remove observations that are no longer useful, a SORT procedure to sort the file, and a DATASETS procedure to rebuild the file's indexes. If another SAS session accesses the file between any of the steps, the SORT and DATASETS procedures would fail, because they require member-level locking (exclusive) access to the file.

Including the LOCK statement before the DATA step provides the needed protection by acquiring exclusive access to the file. If the LOCK statement is successful, a SAS session that attempts to access the file between steps will be denied access, and the nightly update process runs uninterrupted. See Example 1 on page 1675.

Return Codes for the LOCK Statement The SAS macro variable SYSLCKRC contains the return code from the LOCK statement. The following actions result in a nonzero value in SYSLCKRC:

- You try to lock a file but cannot obtain the lock (for example, the file was in use or is locked by another SAS session).
- You use a LOCK statement with the LIST option to list a lock.
- You use a LOCK statement with the CLEAR option to release a lock that you do not have.

For more information about the SYSLCKRC SAS macro variable, see *SAS Macro Language: Reference*.

Comparisons

- With SAS/SHARE software, you can also use the LOCK statement. Some LOCK statement functionality applies only to a client/server environment.
- The CNTLLEV= data set option specifies the level at which shared update access to a SAS data set is denied.

Examples

Example 1: Locking a SAS File The following SAS program illustrates the process of locking a SAS data set. Including the LOCK statement provides protection for the

multistep program by acquiring exclusive access to the file. Any SAS session that attempts to access the file between steps will be denied access, which ensures that the program runs uninterrupted.

```
libname mydata 'SAS-library';

lock mydata.census; ❶

data mydata.census; ❷
   modify mydata.census;
   (statements to remove obsolete observations)
run;

proc sort force data=mydata.census; ❸
   by CrimeRate;
run;

proc datasets library=mydata; ❹
   modify census;
   index create CrimeRate;
quit;

lock mydata.census clear; ❺
```

1 Acquires exclusive access to the SAS data set MYDATA.CENSUS.

2 Opens MYDATA.CENSUS to remove observations that are no longer useful. At the end of the DATA step, the file is closed. However, because of the exclusive lock, any other SAS session that attempts to access the file is denied access.

3 Opens MYDATA.CENSUS to sort the file. At the end of the procedure, the file is closed but not available to another SAS session.

4 Opens MYDATA.CENSUS to rebuild the file's index. At the end of the procedure, the file is closed but still not available to another SAS session.

5 Releases the exclusive lock on MYDATA.CENSUS. The data set is now available to other SAS sessions.

See Also

Data Set Option:

"CNTLLEV= Data Set Option" on page 18

For information about locking a data object in a library that is accessed via a multiuser SAS/SHARE server, see the LOCK statement in the *SAS/SHARE User's Guide*.

LOSTCARD Statement

Resynchronizes the input data when SAS encounters a missing or invalid record in data that has multiple records per observation.

Valid: in a DATA step

Category: Action
Type: Executable

Syntax

LOSTCARD;

Without Arguments

The LOSTCARD statement prevents SAS from reading a record from the next group when the current group has a missing record.

Details

When to Use LOSTCARD When SAS reads multiple records to create a single observation, it does not discover that a record is missing until it reaches the end of the data. If there is a missing record in your data, the values for subsequent observations in the SAS data set might be incorrect. Using LOSTCARD prevents SAS from reading a record from the next group when the current group has fewer records than SAS expected.

LOSTCARD is most useful when the input data have a fixed number of records per observation and when each record for an observation contains an identification variable that has the same value. LOSTCARD usually appears in conditional processing such as in the THEN clause of an IF-THEN statement, or in a statement in a SELECT group.

When LOSTCARD Executes When LOSTCARD executes, SAS takes several steps:

1. Writes three items to the SAS log: a lost card message, a ruler, and all the records that it read in its attempt to build the current observation.
2. Discards the first record in the group of records being read, does not write an observation, and returns processing to the beginning of the DATA step.
3. Does not increment the automatic variable _N_ by 1. (Normally, SAS increments _N_ by 1 at the beginning of each DATA step iteration.)
4. Attempts to build an observation by beginning with the second record in the group, and reads the number of records that the INPUT statement specifies.
5. Repeats steps 1 through 4 when the IF condition for a lost card is still true. To make the log more readable, SAS prints the message and ruler only once for a given group of records. In addition, SAS prints each record only once, even if a record is used in successive attempts to build an observation.
6. Builds an observation and writes it to the SAS data set when the IF condition for a lost card is no longer true.

Examples

This example uses the LOSTCARD statement in a conditional construct to identify missing data records and to resynchronize the input data:

```
data inspect;
   input id 1-3 age 8-9 #2 id2 1-3 loc
         #3 id3 1-3 wt;
   if id ne id2 or id ne id3 then
    do;
      put 'DATA RECORD ERROR: ' id= id2= id3=;
```

```
      lostcard;
    end;
   datalines;
301    32
301    61432
301    127
302    61
302    83171
400    46
409    23145
400    197
411    53
411    99551
411    139
;
```

The DATA step reads three input records before writing an observation. If the identification number in record 1 (variable ID) does not match the identification number in the second record (ID2) or third record (ID3), a record is incorrectly entered or omitted. The IF-THEN DO statement specifies that if an identification number is invalid, SAS prints the message that is specified in the PUT statement message and executes the LOSTCARD statement.

In this example, the third record for the second observation (ID3=400) is missing. The second record for the third observation is incorrectly entered (ID=400 while ID2=409). Therefore, the data set contains two observations with ID values 301 and 411. There are no observations for ID=302 or ID=400. The PUT and LOSTCARD statements write these statements to the SAS log when the DATA step executes:

Output 6.19

```
DATA RECORD ERROR: id=302 id2=302 id3=400
NOTE: LOST CARD.
RULE:----+----1----+----2----+----3----+----4----+----5----+----
14   302    61
15   302    83171
16   400    46
DATA RECORD ERROR: id=302 id2=400 id3=409
NOTE: LOST CARD.
17   409    23145
DATA RECORD ERROR: id=400 id2=409 id3=400
NOTE: LOST CARD.
18   400    197
DATA RECORD ERROR: id=409 id2=400 id3=411
NOTE: LOST CARD.
19   411    53
DATA RECORD ERROR: id=400 id2=411 id3=411
NOTE: LOST CARD.
20   411    99551
```

The numbers 14, 15, 16, 17, 18, 19, and 20 are line numbers in the SAS log.

See Also

Statement:

"IF-THEN/ELSE Statement" on page 1582

MERGE Statement

Joins observations from two or more SAS data sets into a single observation.

Valid: in a DATA step
Category: File-handling
Type: Executable

Syntax

MERGE *SAS-data-set-1* <(*data-set-options*)>
SAS-data-set-2 <(*data-set-options*) >
<…*SAS-data-set-n*<(*data-set-options*)>>
<END=*variable*>;

Arguments

SAS-data-set
specifies at least two existing SAS data sets from which observations are read. You can specify individual data sets, data set lists, or a combination of both.

Tip: You can specify additional SAS data sets.

See: "Using Data Set Lists with MERGE" on page 1679

(*data-set-options*)
specifies one or more SAS data set options in parentheses after a SAS data set name.

Explanation: The data set options specify actions that SAS is to take when it reads observations into the DATA step for processing. For a list of data set options, see "Data Set Options by Category" on page 12.

Tip: Data set options that apply to a data set list apply to all of the data sets in the list.

END=*variable*
names and creates a temporary variable that contains an end-of-file indicator.

Explanation: The variable, which is initialized to 0, is set to 1 when the MERGE statement processes the last observation. If the input data sets have different numbers of observations, the END= variable is set to 1 when MERGE processes the last observation from all data sets.

Tip: The END= variable is not added to any SAS data set that is being created.

Details

Overview The MERGE statement is flexible and has a variety of uses in SAS programming. This section describes basic uses of MERGE. Other applications include using more than one BY variable, merging more than two data sets, and merging a few observations with all observations in another data set.

For more information, see "How to Prepare Your Data Sets" in *SAS Language Reference: Concepts*.

Using Data Set Lists with MERGE You can use data set lists with the MERGE statement. Data set lists provide a quick way to reference existing groups of data sets. These data set lists must be either name prefix lists or numbered range lists.

Name prefix lists refer to all data sets that begin with a specified character string. For example, **merge SALES1:;** tells SAS to merge all data sets starting with "SALES1" such as SALES1, SALES10, SALES11, and SALES12.

Numbered range lists require you to have a series of data sets with the same name, except for the last character or characters, which are consecutive numbers. In a numbered range list, you can begin with any number and end with any number. For example, these lists refer to the same data sets:

```
sales1 sales2 sales3 sales4

sales1-sales4
```

Note: If the numeric suffix of the first data set name contains leading zeros, the number of digits in the numeric suffix of the last data set name must be greater than or equal to the number of digits in the first data set name. Otherwise, an error will occur. For example, the data set lists sales001–sales99 and sales01–sales9 will cause an error. The data set list sales001–sales999 is valid. If the numeric suffix of the first data set name does not contain leading zeros, the number of digits in the numeric suffix of the first and last data set names do not have to be equal. For example, the data set list sales1–sales999 is valid. △

Some other rules to consider when using numbered data set lists are as follows:

- You can specify groups of ranges.

```
merge cost1-cost4 cost11-cost14 cost21-cost24;
```

- You can mix numbered range lists with name prefix lists.

```
merge cost1-cost4 cost2: cost33-37;
```

- You can mix single data sets with data set lists.

```
merge cost1 cost10-cost20 cost30;
```

- Quotation marks around data set lists are ignored.

```
/* these two lines are the same */
merge sales1-sales4;
merge 'sales1'n-'sales4'n;
```

- Spaces in data set names are invalid. If quotation marks are used, trailing blanks are ignored.

```
/* blanks in these statements will cause errors */
merge sales 1-sales 4;
merge 'sales 1'n - 'sales 4'n;

/* trailing blanks in this statement will be ignored */
merge 'sales1   'n - 'sales4   'n;
```

- The maximum numeric suffix is 2147483647.

```
/* this suffix will cause an error */
merge prod2000000000-prod2934850239;
```

- Physical pathnames are not allowed.

```
/* physical pathnames will cause an error */
&let work_path = %sysfunc(pathname(WORK));
merge "&work_path\dept.sas7bdat"-"&work_path\emp.sas7bdat" ;
```

One-to-One Merging One-to-one merging combines observations from two or more SAS data sets into a single observation in a new data set. To perform a one-to-one

merge, use the MERGE statement without a BY statement. SAS combines the first observation from all data sets that are named in the MERGE statement into the first observation in the new data set, the second observation from all data sets into the second observation in the new data set, and so on. In a one-to-one merge, the number of observations in the new data set is equal to the number of observations in the largest data set named in the MERGE statement. See Example 1 for an example of a one-to-one merge. For more information, see "Reading, Combining, and Modifying SAS Data Sets" in *SAS Language Reference: Concepts*.

CAUTION:
Use care when you combine data sets with a one-to-one merge. One-to-one merges can sometimes produce undesirable results. Test your program on representative samples of the data sets before you use this method. △

Match-Merging Match-merging combines observations from two or more SAS data sets into a single observation in a new data set according to the values of a common variable. The number of observations in the new data set is the sum of the largest number of observations in each BY group in all data sets. To perform a match-merge, use a BY statement immediately after the MERGE statement. The variables in the BY statement must be common to all data sets. Only one BY statement can accompany each MERGE statement in a DATA step. The data sets that are listed in the MERGE statement must be sorted in order of the values of the variables that are listed in the BY statement, or they must have an appropriate index. See Example 2 for an example of a match-merge. For more information, see "Reading, Combining, and Modifying SAS Data Sets" in *SAS Language Reference: Concepts*.

Note: The MERGE statement does not produce a Cartesian product on a many-to-many match-merge. Instead it performs a one-to-one merge while there are observations in the BY group in at least one data set. When all observations in the BY group have been read from one data set and there are still more observations in another data set, SAS performs a one-to-many merge until all observations have been read for the BY group. △

Comparisons

- MERGE combines observations from two or more SAS data sets. UPDATE combines observations from exactly two SAS data sets. UPDATE changes or updates the values of selected observations in a master data set as well. UPDATE also might add observations.
- Like UPDATE, MODIFY combines observations from two SAS data sets by changing or updating values of selected observations in a master data set.
- The results that are obtained by reading observations using two or more SET statements are similar to the results that are obtained by using the MERGE statement with no BY statement. However, with the SET statements, SAS stops processing before all observations are read from all data sets if the number of observations are not equal. In contrast, SAS continues processing all observations in all data sets named in the MERGE statement.

Examples

Example 1: One-to-One Merging This example shows how to combine observations from two data sets into a single observation in a new data set:

```
data benefits.qtr1;
   merge benefits.jan benefits.feb;
run;
```

Example 2: Match-Merging This example shows how to combine observations from two data sets into a single observation in a new data set according to the values of a variable that is specified in the BY statement:

```
data inventry;
   merge stock orders;
   by partnum;
run;
```

Example 3: Merging with a Data Set List This example uses a data list to define the data sets that are merged.

```
data d008; job=3; emp=19; run;
data d009; job=3; sal=50; run;
data d010; job=4; emp=97; run;
data d011; job=4; sal=15; run;
data comb;
merge d008-d011;
by job;
run;
proc print data=comb;
run;
```

See Also

Statements:

"BY Statement" on page 1452
"MODIFY Statement" on page 1684
"SET Statement" on page 1764
"UPDATE Statement" on page 1787

"Reading, Combining, and Modifying SAS Data Sets" in *SAS Language Reference: Concepts*

MISSING Statement

Assigns characters in your input data to represent special missing values for numeric data.

Valid: anywhere
Category: Information

Syntax

MISSING *character(s)*;

Arguments

character
is the value in your input data that represents a special missing value.

Range: Special missing values can be any of the 26 letters of the alphabet (uppercase or lowercase) or the underscore (_).

Tip: You can specify more than one character.

Details

The MISSING statement usually appears within a DATA step, but it is global in scope.

Comparisons

The MISSING= system option allows you to specify a character to be printed when numeric variables contain ordinary missing values (.). If your data contain characters that represent special missing values, such as **a** or **z**, do not use the MISSING= option to define them; simply define these values in a MISSING statement.

Examples

With survey data, you might want to identify certain types of missing data. For example, in the data, an **A** can mean that the respondent is not at home at the time of the survey; an **R** can mean that the respondent refused to answer. Use the MISSING statement to identify to SAS that the values **A** and **R** in the input data lines are to be considered special missing values rather than invalid numeric data values:

```
data survey;
   missing a r;
   input id answer;
   datalines;
001 2
002 R
003 1
004 A
005 2
;
```

The resulting data set SURVEY contains exactly the values that are coded in the input data.

See Also

Statement:

"UPDATE Statement" on page 1787

System Option:

"MISSING= System Option" on page 1944

MODIFY Statement

Replaces, deletes, and appends observations in an existing SAS data set in place but does not create an additional copy.

Valid: in a DATA step

Category: File-handling

Type: Executable

Restriction: Cannot modify the descriptor portion of a SAS data set, such as adding a variable

Syntax

❶ **MODIFY** *master-data-set* <(data-set-option(s))> *transaction-data-set*
<(*data-set-option(s)*)>
<NOBS=*variable*> <END=*variable*> <UPDATEMODE=MISSINGCHECK|
NOMISSINGCHECK>;
BY *by-variable*;

❷ **MODIFY** *master-data-set* <(*data-set-option(s)*)> KEY=*index* </ UNIQUE>
<NOBS=*variable*> <END=*variable*> ;

❸ **MODIFY** *master-data-set* <(*data-set-option(s)*)> <NOBS=*variable*> POINT=*variable*;

❹ **MODIFY** *master-data-set* <(*data-set-option(s)*)> <NOBS=*variable*> <END=*variable*>;

CAUTION:
Damage to the SAS data set can occur if the system terminates abnormally during a DATA step that contains the MODIFY statement. Observations in native SAS data files might have incorrect data values, or the data file might become unreadable. DBMS tables that are referenced by views are not affected. △

Note: If you modify a password-protected data set, specify the password with the appropriate data set option (ALTER= or PW=) within the MODIFY statement, and not in the DATA statement. △

Arguments

master-data-set
specifies the SAS data set that you want to modify.

Restriction: This data set must also appear in the DATA statement.

Restriction: The following restrictions apply:

- For sequential and matching access, the master data set can be a SAS data file, a SAS/ACCESS view, an SQL view, or a DBMS engine for the LIBNAME statement. It cannot be a DATA step view or a pass-through view.
- For random access using POINT=, the master data set must be a SAS data file or an SQL view that references a SAS data file.
- For direct access using KEY=, the master data set can be a SAS data file or the DBMS engine for the LIBNAME statement. If it is a SAS file, it must be indexed and the index name must be specified on the KEY= option.

- For a DBMS, the KEY= is set to the keyword DBKEY and the column names to use as an index must be specified on the DBKEY= data set option. These column names are used in constructing a WHERE expression that is passed to the DBMS.

transaction-data-set
specifies the SAS data set that provides the values for matching access. These values are the values that you want to use to update the master data set.

Restriction: Specify this data set *only* when the DATA step contains a BY statement.

by-variable
specifies one or more variables by which you identify corresponding observations.

END=*variable*
creates and names a temporary variable that contains an end-of-file indicator.

Explanation: The variable, which is initialized to zero, is set to 1 when the MODIFY statement reads the last observation of the data set being modified (for sequential access ❹) or the last observation of the transaction data set (for matching access ❶). It is also set to 1 when MODIFY cannot find a match for a KEY= value (random access ❷ ❸).

This variable is not added to any data set.

Restriction: Do not use this argument in the same MODIFY statement with the POINT= argument. POINT= indicates that MODIFY uses random access. The value of the END= variable is never set to 1 for random access.

KEY=*index*
specifies a simple or composite index of the SAS data file that is being modified. The KEY= argument retrieves observations from that SAS data file based on index values that are supplied by like-named variables in another source of information.

Default: If the KEY= value is not found, the automatic variable _ERROR_ is set to 1, and the automatic variable _IORC_ receives the value corresponding to the SYSRC autocall macro's mnemonic _DSENOM. See "Automatic Variable _IORC_ and the SYSRC Autocall Macro" on page 1688 .

Restriction: KEY= processing is different for SAS/ACCESS engines. See the SAS/ACCESS documentation for more information.

Tip: Examples of sources for index values include a separate SAS data set named in a SET statement and an external file that is read by an INPUT statement.

Tip: If duplicates exist in the master file, only the first occurrence is updated unless you use a DO-LOOP to execute a SET statement for the data set that is listed on the KEY=option for all duplicates in the master data set.

If duplicates exist in the transaction data set, and they are consecutive, use the UNIQUE option to force the search for a match in the master data set to begin at the top of the index. Write an accumulation statement to add each duplicate transaction to the observation in master. Without the UNIQUE option, only the first duplicate transaction observation updates the master.

If the duplicates in the transaction data set are not consecutive, the search begins at the beginning of the index each time, so that each duplicate is applied to the master. Write an accumulation statement to add each duplicate to the master.

See Also: UNIQUE on page 1686

Featured in: Example 4 on page 1696, Example 5 on page 1697, and Example 6 on page 1698

NOBS=*variable*
creates and names a temporary variable whose value is usually the total number of observations in the input data set. For certain SAS views, SAS cannot determine the

number of observations. In these cases, SAS sets the value of the NOBS= variable to the largest positive integer value available in the operating environment.

Explanation: At compilation time, SAS reads the descriptor portion of the data set and assigns the value of the NOBS= variable automatically. Thus, you can refer to the NOBS= variable before the MODIFY statement. The variable is available in the DATA step but is not added to the new data set.

Tip: The NOBS= and POINT= options are independent of each other.

Featured in: Example 3 on page 1694

POINT=*variable*
reads SAS data sets using random (direct) access by observation number. *variable* names a variable whose value is the number of the observation to read. The POINT= variable is available anywhere in the DATA step, but it is not added to any SAS data set.

Requirement: When using the POINT= argument, include one or both of the following programming constructs:

- a STOP statement
- programming logic that checks for an invalid value of the POINT= variable

Because POINT= reads only the specified observations, SAS cannot detect an end-of-file condition as it would if the file were being read sequentially. Because detecting an end-of-file condition terminates a DATA step automatically, failure to substitute another means of terminating the DATA step when you use POINT= can cause the DATA step to go into a continuous loop.

Restriction: You cannot use the POINT= option with any of the following:

- BY statement
- WHERE statement
- WHERE= data set option
- transport format data sets
- sequential data sets (on tape or disk)
- a table from another vendor's relational database management system.

Restriction: You can use POINT= with compressed data sets only if the data set was created with the POINTOBS= data set option set to YES, the default value.

Restriction: You can use the random access method on compressed files only with SAS version 7 and beyond.

Tip: If the POINT= value does not match an observation number, SAS sets the automatic variable _ERROR_ to 1.

Featured in: Example 3 on page 1694

UNIQUE
causes a KEY= search always to begin at the top of the index for the data file being modified.

Restriction: UNIQUE can appear only with the KEY= option.

Tip: Use UNIQUE when there are consecutive duplicate KEY= values in the transaction data set, so that the search for a match in the master data set begins at the top of the index file for each duplicate transaction. You must include an accumulation statement or the duplicate values overwrite each other causing only the last transaction value to be the result in the master observation.

Featured in: Example 5 on page 1697

UPDATEMODE=MISSINGCHECK | NOMISSINGCHECK
specifies whether missing variable values in a transaction data set are to be allowed to replace existing variable values in a master data set.

MISSINGCHECK
prevents missing variable values in a transaction data set from replacing values in a master data set.

NOMISSINGCHECK
allows missing variable values in a transaction data set to replace values in a master data set by preventing the check from being performed.

Default: MISSINGCHECK

Requirement: The UPDATEMODE argument must be accompanied by a BY statement that specifies the variables by which observations are matched.

Tip: However, special missing values are the exception and they replace values in the master data set even when MISSINGCHECK is in effect.

Details

❶Matching Access The matching access method uses the BY statement to match observations from the transaction data set with observations in the master data set. The BY statement specifies a variable that is in the transaction data set and the master data set.

When the MODIFY statement reads an observation from the transaction data set, it uses dynamic WHERE processing to locate the matching observation in the master data set. The observation in the master data set can be either

- replaced in the master data set with the value from the transaction data set
- deleted from the master data set
- appended to the master data set.

Example 2 on page 1693 shows the matching access method.

❶Duplicate BY Values Duplicates in the master and transaction data sets affect processing.

- If duplicates exist in the master data set, only the first occurrence is updated because the generated WHERE statement always finds the first occurrence in the master.
- If duplicates exist in the transaction data set, the duplicates are applied one on top of another unless you write an accumulation statement to add all of them to the master observation. Without the accumulation statement, the values in the duplicates overwrite each other so that only the value in the last transaction is the result in the master observation.

❷Direct Access by Indexed Values This method requires that you use the KEY= option in the MODIFY statement to name an indexed variable from the data set that is being modified. Use another data source (typically a SAS data set named in a SET statement or an external file read by an INPUT statement) to provide a like-named variable whose values are supplied to the index. MODIFY uses the index to locate observations in the data set that is being modified.

Example 4 on page 1696 shows the direct-access-by-indexed-values method.

❷Duplicate Index Values

- If there are duplicate values of the indexed variable in the master data set, only the first occurrence is retrieved, modified, or replaced. Use a DO LOOP to execute

a SET statement with the KEY= option multiple times to update all duplicates with the transaction value.

- If there are duplicate, *nonconsecutive* values in the like-named variable in the data source, MODIFY applies each transaction cumulatively to the first observation in the master data set whose index value matches the values from the data source. Therefore, only the value in the last duplicate transaction is the result in the master observation unless you write an accumulation statement to accumulate each duplicate transaction value in the master observation.
- If there are duplicate, *consecutive* values in the variable in the data source, the values from the first observation in the data source are applied to the master data set, but the DATA step terminates with an error when it tries to locate an observation in the master data set for the second duplicate from the data source. To avoid this error, use the UNIQUE option in the MODIFY statement. The UNIQUE option causes SAS to return to the top of the master data set before retrieving a match for the index value. You must write an accumulation statement to accumulate the values from all the duplicates. If you do not, only the last one applied is the result in the master observation.

  Example 5 on page 1697 shows how to handle duplicate index values.
- If there are duplicate index values in both data sets, you can use SQL to apply the duplicates in the transaction data set to the duplicates in the master data set in a one-to-one correspondence.

❸Direct (Random) Access by Observation Number You can use the POINT= option in the MODIFY statement to name a variable from another data source (not the master data set), whose value is the number of an observation that you want to modify in the master data set. MODIFY uses the values of the POINT= variable to retrieve observations in the data set that you are modifying. (You can use POINT= on a compressed data set only if the data set was created with the POINTOBS= data set option.)

It is good programming practice to validate the value of the POINT= variable and to check the status of the automatic variable _ERROR_.

Example 3 on page 1694 shows the direct (random) access by observation number method.

CAUTION:

POINT= can result in infinite looping. Be careful when you use POINT=, as failure to terminate the DATA step can cause the DATA step to go into a continuous loop. Use a STOP statement, programming logic that checks for an invalid value of the POINT= variable, or both. △

❹Sequential Access The sequential access method is the simplest form of the MODIFY statement, but it provides less control than the direct access methods. With the sequential access method, you can use the NOBS= and END= options to modify a data set; you do not use the POINT= or KEY= options.

Preparing Your Data Sets before Using MODIFY There are a number of things you can do to improve performance and get the results you want when using the MODIFY statement. For more information, see “Combining SAS Data Sets: Basic Concepts” in *SAS Language Reference: Concepts*.

Automatic Variable _IORC_ and the SYSRC Autocall Macro The automatic variable _IORC_ contains the return code for each I/O operation that the MODIFY statement attempts to perform. The best way to test for values of _IORC_ is with the mnemonic codes that are provided by the SYSRC autocall macro. Each mnemonic code describes

one condition. The mnemonics provide an easy method for testing problems in a DATA step program. These codes are useful:

_DSENMR
specifies that the transaction data set observation does not exist on the master data set (used only with MODIFY and BY statements). If consecutive observations with different BY values do not find a match in the master data set, both of them return _DSENMR.

_DSEMTR
specifies that multiple transaction data set observations with a given BY value do not exist on the master data set (used only with MODIFY and BY statements). If consecutive observations with the same BY values do not find a match in the master data set, the first observation returns _DSENMR and the subsequent observations return _DSEMTR.

_DSENOM
specifies that the data set being modified does not contain the observation that is requested by the KEY= option or the POINT= option.

_SENOCHN
specifies that SAS is attempting to execute an OUTPUT or REPLACE statement on an observation that contains a key value which duplicates one already existing on an indexed data set that requires unique key values.

_SOK
specifies that the observation was located.

Note: The IORCMSG function returns a formatted error message associated with the current value of _IORC_. △

Example 6 on page 1698 shows how to use the automatic variable _IORC_ and the SYSRC autocall macro.

Writing Observations When MODIFY Is Used in a DATA Step The way SAS writes observations to a SAS data set when the DATA step contains a MODIFY statement depends on whether certain other statements are present. The possibilities are

no explicit statement
writes the current observation to its original place in the SAS data set. The action occurs as the last action in the step (as if a REPLACE statement were the last statement in the step).

OUTPUT statement
if no data set is specified in the OUTPUT statement, writes the current observation to the end of all data sets that are specified in the DATA step. If a data set is specified, the statement writes the current observation to the end of the data set that is indicated. The action occurs at the point in the DATA step where the OUTPUT statement appears.

REPLACE <*data-set-name*> statement
rewrites the current observation in the specified data set or data sets, or, if no argument is specified, rewrites the current observation in each data set specified in the DATA statement. The action occurs at the point of the REPLACE statement.

REMOVE <*data-set-name*> statement
deletes the current observation in the specified data set or data sets, or, if no argument is specified, deletes the current observation in each data set specified in

the DATA statement. The deletion can be a physical one or a logical one, depending on the characteristics of the engine that maintains the data set.

Remember the following as you work with these statements:

- When no OUTPUT, REPLACE, or REMOVE statement is specified, the default action is REPLACE.
- The OUTPUT, REPLACE, and REMOVE statements are independent of each other. You can code multiple OUTPUT, REPLACE, and REMOVE statements to apply to one observation. However, once an OUTPUT, REPLACE, or REMOVE statement executes, the MODIFY statement must execute again before the next REPLACE or REMOVE statement executes.

  You can use OUTPUT and REPLACE in the following example of conditional logic because only one of the REPLACE or OUTPUT statements executes per observation:

```
data master;
   modify master trans; by key;
   if _iorc_=0 then replace;
   else
      output;
run;
```

  But you should not use multiple REPLACE operations on the same observation as in this example:

```
data master;
   modify master;
   x=1;
   replace;
   replace;
run;
```

  You can code multiple OUTPUT statements per observation. However, be careful when you use multiple OUTPUT statements. It is possible to go into an infinite loop with just one OUTPUT statement.

```
data master;
   modify master;
   output;
run;
```

- Using OUTPUT, REPLACE, or REMOVE in a DATA step overrides the default replacement of observations. If you use any one of these statements in a DATA step, you must explicitly program each action that you want to take.
- If both an OUTPUT statement and a REPLACE or REMOVE statement execute on a given observation, perform the OUTPUT action last to keep the position of the observation pointer correct.

Example 7 on page 1700 shows how to use the OUTPUT, REMOVE, and REPLACE statements to write observations.

Missing Values and the MODIFY Statement By default, the UPDATEMODE=MISSINGCHECK option is in effect, so missing values in the transaction data set do *not* replace existing values in the master data set. Therefore, if you want to update some but not all variables and if the variables that you want to update differ from one observation to the next, set to missing those variables that are not changing. If you want missing values in the transaction data set to replace existing values in the master data set, use UPDATEMODE=NOMISSINGCHECK.

Even when UPDATEMODE=MISSINGCHECK is in effect, you can replace existing values with missing values by using special missing value characters in the transaction data set. To create the transaction data set, use the MISSING statement in the DATA step. If you define one of the special missing values **A** through **Z** for the transaction data set, SAS updates numeric variables in the master data set to that value.

If you want the resulting value in the master data set to be a regular missing value, use a single underscore (_) to represent missing values in the transaction data set. The resulting value in the master data set will be a period (.) for missing numeric values and a blank for missing character values.

For more information about defining and using special missing value characters, see "MISSING Statement" on page 1682.

Using MODIFY with Data Set Options If you use data set options (such as KEEP=) in your program, then use the options in the MODIFY statement for the master data set. Using data set options in the DATA statement might produce unexpected results.

Using MODIFY in a SAS/SHARE Environment In a SAS/SHARE environment, the MODIFY statement accesses an observation in update mode. That is, the observation is locked from the time MODIFY reads it until a REPLACE or REMOVE statement executes. At that point the observation is unlocked. It cannot be accessed until it is re-read with the MODIFY statement. The MODIFY statement opens the data set in update mode, but the control level is based on the statement used. For example, KEY= and POINT= are member-level locking. Refer to *SAS/SHARE User's Guide* for more information.

Comparisons

- When you use a MERGE, SET, or UPDATE statement in a DATA step, SAS creates a new SAS data set. The data set descriptor of the new copy can be different from the old one (variables added or deleted, labels changed, and so on). When you use a MODIFY statement in a DATA step, however, SAS does not create a new copy of the data set. As a result, the data set descriptor cannot change.

  For information about DBMS replacement rules, see the SAS/ACCESS documentation.

- If you use a BY statement with a MODIFY statement, MODIFY works much like the UPDATE statement, except that

  - neither the master data set nor the transaction data set needs to be sorted or indexed. (The BY statement that is used with MODIFY triggers dynamic WHERE processing.)

    Note: Dynamic WHERE processing can be costly if the MODIFY statement modifies a SAS data set that is not in sorted order or has not been indexed. Having the master data set in sorted order or indexed and having the transaction data set in sorted order reduces processing overhead, especially for large files. △

  - both the master data set and the transaction data set can have observations with duplicate values of the BY variables. MODIFY treats the duplicates as described in "❶Duplicate BY Values" on page 1687.

  - MODIFY cannot make any changes to the descriptor information of the data set as UPDATE can. Thus, it cannot add or delete variables, change variable labels, and so on.

Input Data Set for Examples

The examples modify the INVTY.STOCK data set. INVTY.STOCK contains these variables:

PARTNO
is a character variable with a unique value identifying each tool number.

DESC
is a character variable with the text description of each tool.

INSTOCK
is a numeric variable with a value describing how many units of each tool the company has in stock.

RECDATE
is a numeric variable containing the SAS date value that is the day for which INSTOCK values are current.

PRICE
is a numeric variable with a value that describes the unit price for each tool.

In addition, INVTY.STOCK contains a simple index on PARTNO. This DATA step creates INVTY.STOCK:

```
libname invty 'SAS-library';


options yearcutoff= 1920;

data invty.stock(index=(partno));
   input PARTNO $ DESC $ INSTOCK @17
         RECDATE date7. @25 PRICE;
   format  recdate date7.;
   datalines;
K89R seal   34  27jul95 245.00
M4J7 sander 98  20jun95 45.88
LK43 filter 121 19may96 10.99
MN21 brace 43   10aug96 27.87
BC85 clamp 80   16aug96 9.55
NCF3 valve 198  20mar96 24.50
KJ66 cutter 6   18jun96 19.77
UYN7 rod  211   09sep96 11.55
JD03 switch 383 09jan97 13.99
BV1E timer 26   03jan97 34.50
;
```

Examples

Example 1: Modifying All Observations This example replaces the date on all of the records in the data set INVTY.STOCK with the current date. It also replaces the value of the variable RECDATE with the current date for all observations in INVTY.STOCK:

```
data invty.stock;
   modify invty.stock;
   recdate=today();
run;

proc print data=invty.stock noobs;
```

```
   title 'INVTY.STOCK';
run;
```

Output 6.20 Results of Updating the RECDATE Field

```
                          INVTY.STOCK                          1

       PARTNO    DESC      INSTOCK    RECDATE     PRICE

        K89R     seal          34      14MAR97    245.00
        M4J7     sander        98      14MAR97     45.88
        LK43     filter       121      14MAR97     10.99
        MN21     brace         43      14MAR97     27.87
        BC85     clamp         80      14MAR97      9.55
        NCF3     valve        198      14MAR97     24.50
        KJ66     cutter         6      14MAR97     19.77
        UYN7     rod          211      14MAR97     11.55
        JD03     switch       383      14MAR97     13.99
        BV1E     timer         26      14MAR97     34.50
```

The MODIFY statement opens INVTY.STOCK for update processing. SAS reads one observation of INVTY.STOCK for each iteration of the DATA step and performs any operations that the code specifies. In this case, the code replaces the value of RECDATE with the result of the TODAY function for every iteration of the DATA step. An implicit REPLACE statement at the end of the step writes each observation to its previous location in INVTY.STOCK.

Example 2: Modifying Observations Using a Transaction Data Set This example adds the quantity of newly received stock to its data set INVTY.STOCK as well as updating the date on which stock was received. The transaction data set ADDINV in the WORK library contains the new data.

The ADDINV data set is the data set that contains the updated information. ADDINV contains these variables:

PARTNO
: is a character variable that corresponds to the indexed variable PARTNO in INVTY.STOCK.

NWSTOCK
: is a numeric variable that represents quantities of newly received stock for each tool.

ADDINV is the second data set in the MODIFY statement. SAS uses it as the transaction data set and reads each observation from ADDINV sequentially. Because the BY statement specifies the common variable PARTNO, MODIFY finds the first occurrence of the value of PARTNO in INVTY.STOCK that matches the value of PARTNO in ADDINV. For each observation with a matching value, the DATA step changes the value of RECDATE to today's date and replaces the value of INSTOCK with the sum of INSTOCK and NWSTOCK (from ADDINV). MODIFY does not add NWSTOCK to the INVTY.STOCK data set because that would modify the data set descriptor. Thus, it is not necessary to put NWSTOCK in a DROP statement.

This example specifies ADDINV as the transaction data set that contains information to modify INVTY.STOCK. A BY statement specifies the shared variable whose values locate the observations in INVTY.STOCK.

This DATA step creates ADDINV:

```
data addinv;
   input PARTNO $ NWSTOCK;
   datalines;
```

```
K89R 55
M4J7 21
LK43 43
MN21 73
BC85 57
NCF3 90
KJ66 2
UYN7 108
JD03 55
BV1E 27
;
```

This DATA step uses values from ADDINV to update INVTY.STOCK.

```
libname invty 'SAS-library';


data invty.stock;
   modify invty.stock addinv;
   by partno;
   RECDATE=today();
   INSTOCK=instock+nwstock;
   if _iorc_=0 then replace;
run;


proc print data=invty.stock noobs;
   title 'INVTY.STOCK';
run;
```

Output 6.21 Results of Updating the INSTOCK and RECDATE Fields

INVTY.STOCK 1

PARTNO	DESC	INSTOCK	RECDATE	PRICE
K89R	seal	89	14MAR97	245.00
M4J7	sander	119	14MAR97	45.88
LK43	filter	164	14MAR97	10.99
MN21	brace	116	14MAR97	27.87
BC85	clamp	137	14MAR97	9.55
NCF3	valve	288	14MAR97	24.50
KJ66	cutter	8	14MAR97	19.77
UYN7	rod	319	14MAR97	11.55
JD03	switch	438	14MAR97	13.99
BV1E	timer	53	14MAR97	34.50

Example 3: Modifying Observations Located by Observation Number This example reads the data set NEWP, determines which observation number in INVTY.STOCK to update based on the value of TOOL_OBS, and performs the update. This example explicitly specifies the update activity by using an assignment statement to replace the value of PRICE with the value of NEWP.

The data set NEWP contains two variables:

TOOL_OBS
contains the observation number of each tool in the tool company's master data set, INVTY.STOCK.

NEWP
contains the new price for each tool.

This DATA step creates NEWP:

```
data newp;
   input TOOL_OBS NEWP;
   datalines;
 1 251.00
 2 49.33
 3 12.32
 4 30.00
 5 15.00
 6 25.75
 7 22.00
 8 14.00
 9 14.32
10 35.00
;
```

This DATA step updates INVTY.STOCK:

```
libname invty 'SAS-library';


data invty.stock;
   set newp;
   modify invty.stock point=tool_obs
          nobs=max_obs;
   if _error_=1 then
      do;
        put 'ERROR occurred for TOOL_OBS=' tool_obs /
        'during DATA step iteration' _n_ /
        'TOOL_OBS value might be out of range.';
        _error_=0;
        stop;
      end;
   PRICE=newp;
   RECDATE=today();
run;


proc print data=invty.stock noobs;
   title 'INVTY.STOCK';
run;
```

Output 6.22 Results of Updating the RECDATE and PRICE Fields

INVTY.STOCK 1

PARTNO	DESC	INSTOCK	RECDATE	PRICE
K89R	seal	34	14MAR97	251.00
M4J7	sander	98	14MAR97	49.33
LK43	filter	121	14MAR97	12.32
MN21	brace	43	14MAR97	30.00
BC85	clamp	80	14MAR97	15.00
NCF3	valve	198	14MAR97	25.75
KJ66	cutter	6	14MAR97	22.00
UYN7	rod	211	14MAR97	14.00
JD03	switch	383	14MAR97	14.32
BV1E	timer	26	14MAR97	35.00

Example 4: Modifying Observations Located by an Index This example uses the KEY= option to identify observations to retrieve by matching the values of PARTNO from ADDINV with the indexed values of PARTNO in INVTY.STOCK. ADDINV is created in Example 2 on page 1693.

KEY= supplies index values that allow MODIFY to access directly the observations to update. No dynamic WHERE processing occurs. In this example, you specify that the value of INSTOCK in the master data set INVTY.STOCK increases by the value of the variable NWSTOCK from the transaction data set ADDINV.

```
libname invty 'SAS-library';

data invty.stock;
   set addinv;
   modify invty.stock key=partno;
   INSTOCK=instock+nwstock;
   RECDATE=today();
   if _iorc_=0 then replace;
run;

proc print data=invty.stock noobs;
   title 'INVTY.STOCK';
run;
```

Output 6.23 Results of Updating the INSTOCK and RECDATE Fields by Using an Index

INVTY.STOCK 1

PARTNO	DESC	INSTOCK	RECDATE	PRICE
K89R	seal	89	14MAR97	245.00
M4J7	sander	119	14MAR97	45.88
LK43	filter	164	14MAR97	10.99
MN21	brace	116	14MAR97	27.87
BC85	clamp	137	14MAR97	9.55
NCF3	valve	288	14MAR97	24.50
KJ66	cutter	8	14MAR97	19.77
UYN7	rod	319	14MAR97	11.55
JD03	switch	438	14MAR97	13.99
BV1E	timer	53	14MAR97	34.50

Example 5: Handling Duplicate Index Values This example shows how MODIFY handles duplicate values of the variable in the SET data set that is supplying values to the index on the master data set.

The NEWINV data set is the data set that contains the updated information. NEWINV contains these variables:

PARTNO
is a character variable that corresponds to the indexed variable PARTNO in INVTY.STOCK. The NEWINV data set contains duplicate values for PARTNO; **M4J7** appears twice.

NWSTOCK
is a numeric variable that represents quantities of newly received stock for each tool.

This DATA step creates NEWINV:

```
data newinv;
   input PARTNO $ NWSTOCK;
   datalines;
K89R 55
M4J7 21
M4J7 26
LK43 43
MN21 73
BC85 57
NCF3 90
KJ66 2
UYN7 108
JD03 55
BV1E 27
;
```

This DATA step terminates with an error when it tries to locate an observation in INVTY.STOCK to match with the second occurrence of **M4J7** in NEWINV:

```
libname invty 'SAS-library';


   /* This DATA step terminates with an error! */
data invty.stock;
   set newinv;
   modify invty.stock key=partno;
   INSTOCK=instock+nwstock;
   RECDATE=today();
run;
```

This message appears in the SAS log:

```
ERROR: No matching observation was found in MASTER data set.
PARTNO=K89R NWSTOCK=55 DESC=  INSTOCK=. RECDATE=14MAR97 PRICE=.
_ERROR_=1 _IORC_=1230015 _N_=1
NOTE: Missing values were generated as a result of performing
      an operation on missing values.
      Each place is given by:
      (Number of times) at (Line):(Column).
      1 at 689:19
NOTE: The SAS System stopped processing this step because of
      errors.
NOTE: The data set INVTY.STOCK has been updated.  There were 0
      observations rewritten, 0 observations added and 0
      observations deleted.
```

Adding the UNIQUE option to the MODIFY statement avoids the error in the previous DATA step. The UNIQUE option causes the DATA step to return to the top of the index each time it looks for a match for the value from the SET data set. Thus, it finds the **M4J7** in the MASTER data set for each occurrence of **M4J7** in the SET data set. The updated result for **M4J7** in the output shows that both values of NWSTOCK from NEWINV for **M4J7** are added to the value of INSTOCK for **M4J7** in INVTY.STOCK. An accumulation statement sums the values; without it, only the value of the last instance of **M4J7** would be the result in INVTY.STOCK.

```
data invty.stock;
   set newinv;
   modify invty.stock key=partno / unique;
   INSTOCK=instock+nwstock;
   RECDATE=today();
   if _iorc_=0 then replace;
run;

proc print data=invty.stock noobs;
   title 'Results of Using the UNIQUE Option';
run;
```

Output 6.24 Results of Updating the INSTOCK and RECDATE Fields by Using the UNIQUE Option

Results of Using the UNIQUE Option 1

PARTNO	DESC	INSTOCK	RECDATE	PRICE
K89R	seal	89	14MAR97	245.00
M4J7	sander	145	14MAR97	45.88
LK43	filter	164	14MAR97	10.99
MN21	brace	116	14MAR97	27.87
BC85	clamp	137	14MAR97	9.55
NCF3	valve	288	14MAR97	24.50
KJ66	cutter	8	14MAR97	19.77
UYN7	rod	319	14MAR97	11.55
JD03	switch	438	14MAR97	13.99
BV1E	timer	53	14MAR97	34.50

Example 6: Controlling I/O This example uses the SYSRC autocall macro and the _IORC_ automatic variable to control I/O condition. This technique helps to prevent unexpected results that could go undetected. This example uses the direct access

method with an index to update INVTY.STOCK. The data in the NEWSHIP data set updates INVTY.STOCK.

This DATA step creates NEWSHIP:

```
options yearcutoff= 1920;

data newship;
   input PARTNO $ DESC $ NWSTOCK @17
         SHPDATE date7. @25 NWPRICE;
   datalines;
K89R seal 14    14nov96 245.00
M4J7 sander 24  23aug96 47.98
LK43 filter 11  29jan97 14.99
MN21 brace 9    09jan97 27.87
BC85 clamp 12   09dec96 10.00
ME34 cutter 8   14nov96 14.50
;
```

Each WHEN clause in the SELECT statement specifies actions for each input/output return code that is returned by the SYSRC autocall macro:

- _SOK indicates that the MODIFY statement executed successfully.
- _DSENOM indicates that no matching observation was found in INVTY.STOCK. The OUTPUT statement specifies that the observation be appended to INVTY.STOCK. See the last observation in the output.
- If any other code is returned by SYSRC, the DATA step terminates and the PUT statement writes the message to the log.

```
libname invty 'SAS-library';

data invty.stock;
   set newship;
   modify invty.stock key=partno;
   select (_iorc_);
      when (%sysrc(_sok)) do;
         INSTOCK=instock+nwstock;
         RECDATE=shpdate;
         PRICE=nwprice;
         replace;
      end;
      when (%sysrc(_dsenom)) do;
         INSTOCK=nwstock;
         RECDATE=shpdate;
         PRICE=nwprice;
         output;
         _error_=0;
      end;
      otherwise do;
         put
         'An unexpected I/O error has occurred.'/
         'Check your data and your program';
         _error_=0;
         stop;
       end;
   end;
run;
```

```
proc print data=invty.stock noobs;
   title 'INVTY.STOCK Data Set';
run;
```

Output 6.25 The Updated INVTY.STOCK Data Set

INVTY.STOCK Data Set 1

PARTNO	DESC	INSTOCK	RECDATE	PRICE
K89R	seal	48	14NOV96	245.00
M4J7	sander	122	23AUG96	47.98
LK43	filter	132	29JAN97	14.99
MN21	brace	52	09JAN97	27.87
BC85	clamp	92	09DEC96	10.00
NCF3	valve	198	20MAR96	24.50
KJ66	cutter	6	18JUN96	19.77
UYN7	rod	211	09SEP96	11.55
JD03	switch	383	09JAN97	13.99
BV1E	timer	26	03JAN97	34.50
ME34	cutter	8	14NOV96	14.50

Example 7: Replacing and Removing Observations and Writing Observations to Different SAS Data Sets This example shows that you can replace and remove (delete) observations and write observations to different data sets. Further, this example shows that if an OUTPUT, REPLACE, or REMOVE statement is present, you must specify explicitly what action to take because no default statement is generated.

The parts that were received in 1997 are output to INVTY.STOCK97 and are removed from INVTY.STOCK. Likewise, the parts that were received in 1995 are output to INVTY.STOCK95 and are removed from INVTY.STOCK. Only the parts that were received in 1996 remain in INVTY.STOCK, and the PRICE is updated only in INVTY.STOCK.

```
libname invty 'SAS-library';

data invty.stock invty.stock95 invty.stock97;
   modify invty.stock;
   if recdate>'01jan97'd then do;
       output invty.stock97;
       remove invty.stock;
   end;
   else if recdate<'01jan96'd then do;
       output invty.stock95;
       remove invty.stock;
   end;
   else do;
      price=price*1.1;
      replace invty.stock;
   end;
run;

proc print data=invty.stock noobs;
   title 'New Prices for Stock Received in 1996';
run;
```

Output 6.26 Output from Writing Observations to a Specific SAS Data Set

```
            New Prices for Stock Received in 1996     1

 PARTNO     DESC       INSTOCK     RECDATE      PRICE

  LK43      filter       121       19MAY96     12.089
  MN21      brace         43       10AUG96     30.657
  BC85      clamp         80       16AUG96     10.505
  NCF3      valve        198       20MAR96     26.950
  KJ66      cutter         6       18JUN96     21.747
  UYN7      rod          211       09SEP96     12.705
```

See Also

Statements:

"MISSING Statement" on page 1682

"OUTPUT Statement" on page 1704

"REMOVE Statement" on page 1741

"REPLACE Statement" on page 1745

"UPDATE Statement" on page 1787

"Reading, Combining, and Modifying SAS Data Sets" in *SAS Language Reference: Concepts*

"The SQL Procedure" in the *Base SAS Procedures Guide*

Null Statement

Signals the end of data lines or acts as a placeholder.

Valid: Anywhere

Category: Action

Type: Executable

Syntax

;

or

;;;;

Without Arguments

The Null statement signals the end of the data lines that occur in your program.

Details

The primary use of the Null statement is to signal the end of data lines that follow a DATALINES or CARDS statement. In this case, the Null statement functions as a step

boundary. When your data lines contain semicolons, use the DATALINES4 or CARDS4 statement and a Null statement of four semicolons.

Although the Null statement performs no action, it is an executable statement. Therefore, a label can precede the Null statement, or you can use it in conditional processing.

Examples

- The Null statement in this program marks the end of data lines and functions as a step boundary.

```
data test;
   input score1 score2 score3;
   datalines;
55 135 177
44 132 169
;
```

- The input data records in this example contain semicolons. Use the Null statement following the DATALINES4 statement to signal the end of the data lines.

```
data test2;
   input code1 $ code2 $ code3 $;
   datalines4;
55;39;1  135;32;4  177;27;3
78;29;1  149;22;4  179;37;3
;;;;
```

- The Null statement is useful while you are developing a program. For example, use it after a statement label to test your program before you code the statements that follow the label.

```
data _null_;
   set dsn;
   file print header=header;
   put 'report text';
     ...more statements...
   return;
   header:;
run;
```

See Also

Statements:

"DATALINES Statement" on page 1474
"DATALINES4 Statement" on page 1475
"GO TO Statement" on page 1579
"LABEL Statement" on page 1650

OPTIONS Statement

Specifies or changes the value of one or more SAS system options.

Valid: anywhere

Category: Program Control

See: OPTIONS Statement in the documentation for your operating environment.

Syntax

OPTIONS *option(s)*;

Arguments

option
specifies one or more SAS system options to be changed.

Details

The change that is made by the OPTIONS statement remains in effect for the rest of the job, session, SAS process, or until you issue another OPTIONS statement to change the options again. You can specify SAS system options through the OPTIONS statement, through the OPTIONS window, at SAS invocation, and at the initiation of a SAS process.

Note: If you want a particular group of options to be in effect for all your SAS jobs or sessions, store an OPTIONS statement in an autoexec file or list the system options in a configuration file or custom_option_set. △

Note: For a system option with a null value, the GETOPTION function returns a value of "(single quotation marks with a blank space between them), for example `EMAILID=' '`. This GETOPTION value can then be used in the OPTIONS statement. △

An OPTIONS statement can appear at any place in a SAS program, except within data lines.

Operating Environment Information: The system options that are available depend on your operating environment. Also, the syntax that is used to specify a system option in the OPTIONS statement might be different from the syntax that is used at SAS invocation. For details, see the SAS documentation for your operating environment. △

Comparisons

The OPTIONS statement requires you to enter the complete statement including system option name and value, if necessary. The SAS OPTIONS window displays the options' names and settings in columns. To change a setting, type over the value that is displayed and press ENTER or RETURN.

Examples

This example suppresses the date that is normally written to SAS output and sets a line size of 72:

```
options nodate linesize=72;
```

See Also

"Definition of System Options" on page 1822

OUTPUT Statement

Writes the current observation to a SAS data set.

Valid: in a DATA step

Category: Action

Type: Executable

Syntax

OUTPUT<*data-set-name(s)*>;

Without Arguments

Using OUTPUT without arguments causes the current observation to be written to all data sets that are named in the DATA statement.

Note: If a MODIFY statement is present, OUTPUT with no arguments writes the current observation to the end of the data set that is specified in the MODIFY statement. △

Arguments

data-set-name
specifies the name of a data set to which SAS writes the observation.

Restriction: All names specified in the OUTPUT statement must also appear in the DATA statement.

Tip: You can specify up to as many data sets in the OUTPUT statement as you specified in the DATA statement for that DATA step.

Details

When and Where the OUTPUT Statement Writes Observations The OUTPUT statement tells SAS to write the current observation to a SAS data set immediately, not at the end of the DATA step. If no data set name is specified in the OUTPUT statement, the observation is written to the data set or data sets that are listed in the DATA statement.

Implicit versus Explicit Output By default, every DATA step contains an implicit OUTPUT statement at the end of each iteration that tells SAS to write observations to the data set or data sets that are being created. Placing an explicit OUTPUT statement in a DATA step overrides the automatic output, and SAS adds an observation to a data set only when an explicit OUTPUT statement is executed. Once you use an OUTPUT statement to write an observation to any one data set, however, there is no implicit OUTPUT statement at the end of the DATA step. In this situation, a DATA step writes an observation to a data set only when an explicit OUTPUT executes. You can use the OUTPUT statement alone or as part of an IF-THEN or SELECT statement or in DO-loop processing.

When Using the MODIFY Statement When you use the MODIFY statement with the OUTPUT statement, the REMOVE and REPLACE statements override the implicit write action at the end of each DATA step iteration. See "Comparisons" on page 1705 for more information. If both the OUTPUT statement and a REPLACE or REMOVE statement execute on a given observation, perform the output action last to keep the position of the observation pointer correct.

Comparisons

- OUTPUT writes observations to a SAS data set; PUT writes variable values or text strings to an external file or the SAS log.
- To control when an observation is written to a specified output data set, use the OUTPUT statement. To control which variables are written to a specified output data set, use the KEEP= or DROP= data set option in the DATA statement, or use the KEEP or DROP statement.
- When you use the OUTPUT statement with the MODIFY statement, the following items apply.
  - Using an OUTPUT, REPLACE, or REMOVE statement overrides the default write action at the end of a DATA step. (OUTPUT is the default action; REPLACE becomes the default action when a MODIFY statement is used.) If you use any of these statements in a DATA step, you must explicitly program output for the new observations that are added to the data set.
  - The OUTPUT, REPLACE, and REMOVE statements are independent of each other. More than one statement can apply to the same observation, as long as the sequence is logical.
  - If both an OUTPUT and a REPLACE or REMOVE statement execute on a given observation, perform the OUTPUT action last to keep the position of the observation pointer correct.

Examples

Example 1: Sample Uses of OUTPUT These examples show how you can use an OUTPUT statement:

- This line of code writes the current observation to a SAS data set. `output;`
- This line of code writes the current observation to a SAS data set when a specified condition is tr

□ This line of code writes an observation to the data set MARKUP when the PHONE value is missing

Example 2: Creating Multiple Observations from Each Line of Input You can create two or more observations from each line of input data. This SAS program creates three observations in the data set RESPONSE for each observation in the data set SULFA:

```
data response(drop=time1-time3);
   set sulfa;
   time=time1;
   output;
   time=time2;
   output;
   time=time3;
   output;
run;
```

Example 3: Creating Multiple Data Sets from a Single Input File You can create more than one SAS data set from one input file. In this example, OUTPUT writes observations to two data sets, OZONE and OXIDES:

```
options yearcutoff= 1920;

data ozone oxides;
   infile file-specification;
   input city $ 1-15 date date9.
      chemical $ 26-27 ppm 29-30;
   if chemical='O3' then output ozone;
   else output oxides;
run;
```

Example 4: Creating One Observation from Several Lines of Input You can combine several input observations into one observation. In this example, OUTPUT creates one observation that totals the values of DEFECTS in the first ten observations of the input data set:

```
data discards;
   set gadgets;
   drop defects;
   reps+1;
   if reps=1 then total=0;
   total+defects;
   if reps=10 then do;
      output;
      stop;
   end;
run;
```

See Also

Statements:

"DATA Statement" on page 1465
"MODIFY Statement" on page 1684
"PUT Statement" on page 1708
"REMOVE Statement" on page 1741
"REPLACE Statement" on page 1745

PAGE Statement

Skips to a new page in the SAS log.

Valid: Anywhere
Category: Log Control

Syntax

PAGE;

Without Arguments

The PAGE statement skips to a new page in the SAS log.

Details

You can use the PAGE statement when you run SAS in a windowing environment, batch, or noninteractive mode. The PAGE statement itself does not appear in the log. When you run SAS in interactive line mode, PAGE might print blank lines to the display monitor (or to the alternate log file).

See Also

Statement:

"LIST Statement" on page 1670

System Options:

"LINESIZE= System Option" on page 1936
"PAGESIZE= System Option" on page 1958

PUT Statement

Writes lines to the SAS log, to the SAS output window, or to an external location that is specified in the most recent FILE statement.

Valid: in a DATA step

Category: File-handling

Type: Executable

Syntax

PUT <*specification(s)*><_ODS_><@|@@>;

Without Arguments

The PUT statement without arguments is called a *null PUT statement*. The null PUT statement

- writes the current output line to the current location, even if the current output line is blank
- releases an output line that is being held with a trailing @ by a previous PUT statement.

For an example, see Example 5 on page 1722. For more information, see "Using Line-Hold Specifiers" on page 1716.

Arguments

specification(s)
specifies what is written, how it is written, and where it is written. The specification can include

variable
specifies the variable whose value is written.

Note: Beginning with Version 7, you can specify column-mapped Output Delivery System variables in the PUT statement. This functionality is described briefly here in _ODS_ on page 1710, but documented more completely in PUT Statement for ODS in *SAS Output Delivery System: User's Guide*. △

(*variable-list*)
specifies a list of variables whose values are written.

Requirement: The (*format-list*) must follow the (*variable-list*).

See: "PUT Statement, Formatted" on page 1727

'character-string'
specifies a string of text, enclosed in quotation marks, to write.

Tip: To write a hexadecimal string in EBCDIC or ASCII, follow the ending quotation mark with an **x**.

Tip: If you use single quotation marks (') or double quotation marks ("") together (with no space in between them) as the string of text, SAS will

output a single quotation mark (’) or double quotation mark ("), respectively.

See Also: “List Output” on page 1713

Example: This statement writes HELLO when the hexadecimal string is converted to ASCII characters:

```
put '68656C6C6F'x;
```

*n**

specifies to repeat *n* times the subsequent character string.

Example: This statement writes a line of 132 underscores.

```
put 132*'_';
```

Featured in: Example 4 on page 1722

pointer-control

moves the output pointer to a specified line or column in the output buffer.

See: “Column Pointer Controls” on page 1710 and “Line Pointer Controls” on page 1711

column-specifications

specifies which columns of the output line the values are written.

See: “Column Output” on page 1713

Featured in: Example 2 on page 1719

format.

specifies a format to use when the variable values are written.

See: “Formatted Output” on page 1714

Featured in: Example 1 on page 1718

(*format-list*)

specifies a list of formats to use when the values of the preceding list of variables are written.

Restriction: The (*format-list*) must follow the (*variable-list*).

See: “PUT Statement, Formatted” on page 1727

INFILE

writes the last input data record that is read either from the current input file or from the data lines that follow a DATELINES statement.

Tip: _INFILE_ is an automatic variable that references the current INPUT buffer. You can use this automatic variable in other SAS statements.

Tip: If the most recent INPUT statement uses line-pointer controls to read multiple input data records, PUT _INFILE_ writes only the record that the input pointer is positioned on.

Example: This PUT statement writes all the values of the first input data record:

```
input #3 score #1 name $ 6-23;
put _infile_;
```

Featured in: Example 6 on page 1723

ALL

writes the values of all variables, which includes automatic variables, that are defined in the current DATA step by using named output.

See: “Named Output” on page 1714

ODS
moves data values for all columns (as defined by the ODS option in the FILE statement) into a special buffer, from which it is eventually written to the data component. The ODS option in the FILE statement defines the structure of the data component that holds the results of the DATA step.

Restriction: Use _ODS_ only if you have previously specified the ODS option in the FILE statement.

Tip: You can use the _ODS_ specification in conjunction with variable specifications and column pointers, and it can appear anywhere in a PUT statement.

Interaction: _ODS_ writes data to a specific column only if a PUT statement has not already specified a variable for that column with a column pointer. That is, a variable specification for a column overrides the _ODS_ option.

See: "PUT Statement for ODS" in *SAS Output Delivery System: User's Guide*

@|@@
holds an output line for the execution of the next PUT statement even across iterations of the DATA step. These line-hold specifiers are called *trailing* @ and *double trailing* @.

Restriction: The trailing @ or double trailing @ must be the last item in the PUT statement.

Tip: Use an @ or @@ to hold the pointer at its current location. The next PUT statement that executes writes to the same output line rather than to a new output line.

See: "Using Line-Hold Specifiers" on page 1716

Featured in: Example 5 on page 1722

Column Pointer Controls

@*n*
moves the pointer to column *n*.

Range: a positive integer

Example: @15 moves the pointer to column 15 before the value of NAME is written:

```
put @15 name $10.;
```

Featured in: Example 2 on page 1719 and Example 4 on page 1722

@*numeric-variable*
moves the pointer to the column given by the value of *numeric-variable*.

Range: a positive integer

Tip: If *n* is not an integer, SAS truncates the decimal portion and uses only the integer value. If *n* is zero or negative, the pointer moves to column 1.

Example: The value of the variable A moves the pointer to column 15 before the value of NAME is written:

```
a=15;
put @a name $10.;
```

Featured in: Example 2 on page 1719

@(*expression*)
moves the pointer to the column that is given by the value of *expression*.

Range: a positive integer

Tip: If the value of *expression* is not an integer, SAS truncates the decimal value and uses only the integer value. If it is zero, the pointer moves to column 1.

Example: The result of the expression moves the pointer to column 15 before the value of NAME is written:

```
b=5;
put @(b*3) name $10.;
```

+*n*
moves the pointer *n* columns.

Range: a positive integer or zero

Tip: If *n* is not an integer, SAS truncates the decimal portion and uses only the integer value.

Example: This statement moves the pointer to column 23, writes a value of LENGTH in columns 23 through 26, advances the pointer five columns, and writes the value of WIDTH in columns 32 through 35:

```
put @23 length 4. +5 width 4.;
```

+*numeric-variable*
moves the pointer the number of columns given by the value of *numeric-variable*.

Range: a positive or negative integer or zero

Tip: If *numeric-variable* is not an integer, SAS truncates the decimal value and uses only the integer value. If *numeric-variable* is negative, the pointer moves backward. If the current column position becomes less than 1, the pointer moves to column 1. If the value is zero, the pointer does not move. If the value is greater than the length of the output buffer, the current line is written out and the pointer moves to column 1 on the next line.

+(*expression*)
moves the pointer the number of columns given by *expression*.

Range: *expression* must result in an integer

Tip: If *expression* is not an integer, SAS truncates the decimal value and uses only the integer value. If *expression* is negative, the pointer moves backward. If the current column position becomes less than 1, the pointer moves to column 1. If the value is zero, the pointer does not move. If the value is greater than the length of the output buffer, the current line is written out and the pointer moves to column 1 on the next line.

Featured in: Example 2 on page 1719

Line Pointer Controls

#*n*
moves the pointer to line *n*.

Range: a positive integer

Example: The #2 moves the pointer to the second line before the value of ID is written in columns 3 and 4:

```
put @12 name $10. #2 id 3-4;
```

#*numeric-variable*
moves the pointer to the line given by the value of *numeric-variable*.

Range: a positive integer

Tip: If the value of *numeric-variable* is not an integer, SAS truncates the decimal value and uses only the integer value.

#(*expression*)
moves the pointer to the line that is given by the value of *expression*.

Range: *Expression* must result in a positive integer.

Tip: If the value of *expression* is not an integer, SAS truncates the decimal value and uses only the integer value.

/
advances the pointer to column 1 of the next line.

Example: The values for NAME and AGE are written on one line, and then the pointer moves to the second line to write the value of ID in columns 3 and 4:

```
put name age / id 3-4;
```

Featured in: Example 3 on page 1720

OVERPRINT
causes the values that follow the keyword OVERPRINT to print on the most recently written output line.

Requirement: You must direct the output to a file. Set the N= option in the FILE statement to 1 and direct the PUT statements to a file.

Tip: OVERPRINT has no effect on lines that are written to a display.

Tip: Use OVERPRINT in combination with column pointer and line pointer controls to overprint text.

Example: This statement overprints underscores, starting in column 15, which underlines the title:

```
put @15 'Report Title' overprint
    @15 '____________';
```

Featured in: Example 4 on page 1722

BLANKPAGE
advances the pointer to the first line of a new page, even when the pointer is positioned on the first line and the first column of a new page.

Tip: If the current output file contains carriage-control characters, _BLANKPAGE_ produces output lines that contain the appropriate carriage-control character.

Featured in: Example 3 on page 1720

PAGE
advances the pointer to the first line of a new page. SAS automatically begins a new page when a line exceeds the current PAGESIZE= value.

Tip: If the current output file is printed, _PAGE_ produces an output line that contains the appropriate carriage-control character. _PAGE_ has no effect on a file that is not printed.

Tip: If you specify FILE PRINT in an interactive SAS session, then the Output window interprets the form-feed control characters as page breaks, and they are removed from the output. The resulting file is a flat file without page break characters. If a file needs to contain the form-feed characters, then the FILE statement should include a physical file location and the PRINT option.

Featured in: Example 3 on page 1720

Details

When to Use PUT

Use the PUT statement to write lines to the SAS log, to the SAS output window, or to an external location. If you do not execute a FILE statement before the PUT statement

in the current iteration of a DATA step, SAS writes the lines to the SAS log. If you specify the PRINT option in the FILE statement, SAS writes the lines to the SAS output window.

The PUT statement can write lines that contain variable values, character strings, and hexadecimal character constants. With specifications in the PUT statement, you specify what to write, where to write it, and how to format it.

Output Styles

There are four ways to write variable values with the PUT statement:

- column
- list (simple and modified)
- formatted
- named

A single PUT statement might contain any or all of the available output styles, depending on how you want to write lines.

Column Output With *column output,* the column numbers follow the variable in the PUT statement. These numbers indicate where in the line to write the following value:

```
put name 6-15 age 17-19;
```

These lines are written to the SAS log.*

```
----+----1----+----2----+
     Peterson    21
     Morgan      17
```

The PUT statement writes values for NAME and AGE in the specified columns. See "PUT Statement, Column" on page 1724 for more information.

List Output With *list output,* list the variables and character strings in the PUT statement in the order in which you want to write them. For example, this PUT statement

```
put name age;
```

writes the values for NAME and AGE to the SAS log:*

```
----+----1----+----2----+
Peterson 21
Morgan 17
```

See "PUT Statement, List" on page 1731 for more information.

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

Formatted Output With *formatted output*, specify a SAS format or a user-written format after the variable name. The format gives instructions on how to write the variable value. Formats enable you to write in a non-standard form, such as packed decimal, or numbers that contain special characters such as commas. For example, this PUT statement

```
put name $char10. age 2. +1 date mmddyy10.;
```

writes the values for NAME, AGE, and DATE to the SAS log:*

```
----+----1----+----2----+
Peterson  21 07/18/1999
Morgan    17 11/12/1999
```

Using a pointer control of +1 inserts a blank space between the values of AGE and DATE. See "PUT Statement, Formatted" on page 1727 for more information.

Named Output With *named output*, list the variable name followed by an equal sign. For example, this PUT statement

```
put name= age=;
```

writes the values for NAME and AGE to the SAS log:*

```
----+----1----+----2----+
name=Peterson age=21
name=Morgan age=17
```

See "PUT Statement, Named" on page 1736 for more information.

Using Multiple Output Styles in a Single PUT Statement

A PUT statement can combine any or all of the different output styles. For example,

```
put name 'on ' date mmddyy8. ' weighs '
    startwght +(-1) '.' idno= 40-45;
```

See Example 1 on page 1718 for an explanation of the lines written to the SAS log.

When you combine different output styles, it is important to understand the location of the output pointer after each value is written. For more information about the pointer location, see "Pointer Location After a Value Is Written" on page 1716.

Avoiding a Common Error When Writing Both a Character Constant and a Variable

When using a PUT statement to write a character constant that is followed by a variable name, always put a blank space between the closing quotation mark and the variable name:

```
put 'Player:' name1 'Player:' name2 'Player:' name3;
```

Otherwise, SAS might interpret a character constant that is followed by a variable name as a special SAS constant as illustrated in this table.

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

Table 6.9 Characters That Cause Misinterpretation When They Follow a Character Constant

Starting Letter of Variable	Represents	Examples
b	bit testing constant	'00100000'b
d	date constant	'01jan04'd
dt	datetime constant	'18jan2003:9:27:05am'dt
n	name literal	'My Table'n
t	time constant	'9:25:19pm't
x	hexadecimal notation	'534153'x

Example 7 on page 1723 shows how to use character constants followed by variables. For more information about SAS name literals and SAS constants in expressions, see *SAS Language Reference: Concepts*.

Pointer Controls

As SAS writes values with the PUT statement, it keeps track of its position with a pointer. The PUT statement provides three ways to control the movement of the pointer:

column pointer controls
: reset the pointer's column position when the PUT statement starts to write the value to the output line.

line pointer controls
: reset the pointer's line position when the PUT statement writes the value to the output line.

line-hold specifiers
: hold a line in the output buffer so that another PUT statement can write to it. By default, the PUT statement releases the previous line and writes to a new line.

With column and line pointer controls, you can specify an absolute line number or column number to move the pointer or you can specify a column or line location that is relative to the current pointer position. The following table lists all pointer controls that are available in the PUT statement.

Table 6.10 Pointer Controls Available in the PUT Statement

Pointer Controls	Relative	Absolute
column pointer controls	+*n*	@*n*
	+*numeric-variable*	@*numeric-variable*
	+(*expression*)	@(*expression*)
line pointer controls	/ , _PAGE_ , _BLANKPAGE_	#*n*
		#*numeric-variable*
		#(*expression*)
	OVERPRINT	none

Pointer Controls	Relative	Absolute
line-hold specifiers	@	(not applicable)
	@@	(not applicable)

Note: Always specify pointer controls before the variable for which they apply. △

See "Pointer Location After a Value Is Written" on page 1716 for more information about how SAS determines the pointer position.

Using Line-Hold Specifiers

Line-hold specifiers keep the pointer on the current output line when

- more than one PUT statement writes to the same output line
- a PUT statement writes values from more than one observation to the same output line.

Without line-hold specifiers, each PUT statement in a DATA step writes a new output line.

In the PUT statement, trailing @ and double trailing @@ produce the same effect. Unlike the INPUT statement, the PUT statement does not automatically release a line that is held by a trailing @ when the DATA step begins a new iteration. SAS releases the current output line that is held by a trailing @ or double trailing @ when it encounters

- a PUT statement without a trailing @
- a PUT statement that uses _BLANKPAGE_ or _PAGE_
- the end of the current line (determined by the current value of the LRECL= or LINESIZE= option in the FILE statement, if specified, or the LINESIZE= system option)
- the end of the last iteration of the DATA step.

Using a trailing @ or double trailing @ can cause SAS to attempt to write past the current line length because the pointer value is unchanged when the next PUT statement executes. See "When the Pointer Goes Past the End of a Line" on page 1717.

Pointer Location After a Value Is Written

Understanding the location of the output pointer after a value is written is important, especially if you combine output styles in a single PUT statement. The pointer location after a value is written depends on which output style you use and whether a character string or a variable is written. With column or formatted output, the pointer is located in the first column after the end of the field that is specified in the PUT statement. These two styles write only variable values.

With list output or named output, the pointer is located in the second column after a variable value because PUT skips a column automatically after each value is written. However, when a PUT statement uses list output to write a character string, the pointer is located in the first column after the string. If you do not use a line pointer control or column output after a character string is written, add a blank space to the end of the character string to separate it from the next value.

After an _INFILE_ specification, the pointer is located in the first column after the record is written from the current input file.

When the output pointer is in the upper left corner of a page,

- PUT _BLANKPAGE_ writes a blank page and moves the pointer to the top of the next page.

- PUT _PAGE_ leaves the pointer in the same location.

You can determine the current location of the pointer by examining the variables that are specified with the COLUMN= option and the LINE= option in the FILE statement.

When the Pointer Goes Past the End of a Line

SAS does not write an output line that is longer than the current output line length. The line length of the current output file is determined by

- the value of the LINESIZE= option in the current FILE statement
- the value of the LINESIZE= system option (for the SAS output window)
- the LRECL= option in the current FILE statement (for external files).

You can inadvertently position the pointer beyond the current line length with one or more of these specifications:

- a + pointer control with a value that moves the pointer to a column beyond the current line length
- a column range that exceeds the current line length (for example, PUT X 90 – 100 when the current line length is 80)
- a variable value or character string that does not fit in the space that remains on the current output line.

By default, when PUT attempts to write past the end of the current line, SAS withholds the entire item that overflows the current line, writes the current line, and then writes the overflow item on a new line, starting in column 1. See the FLOWOVER, DROPOVER, and STOPOVER options in the statement “FILE Statement” on page 1503.

Arrays

You can use the PUT statement to write an array element. The subscript is any SAS expression that results in an integer when the PUT statement executes. You can use an array reference in a *numeric-variable* construction with a pointer control if you enclose the reference in parentheses, as shown here:

- @(*array-name*{i})
- +(*array-name*{i})
- #(*array-name*{i})

Use the array subscript asterisk (*) to write all elements of a previously defined array to an external location. SAS allows one-dimensional or multidimensional arrays, but it does not allow a _TEMPORARY_ array. Enclose the subscript in braces, brackets, or parentheses, and print the array using list, formatted, column, or named output. With list output, the form of this statement is

```
PUT array-name{*};
```

With formatted output, the form of this statement is

```
PUT array-name{*}(format|format.list)
```

The format in parentheses follows the array reference.

Comparisons

- The PUT statement writes variable values and character strings to the SAS log or to an external location while the INPUT statement reads raw data in external files or data lines entered instream.

- Both the INPUT and the PUT statements use the trailing @ and double trailing @ line-hold specifiers to hold the current line in the input or output buffer, respectively. In an INPUT statement, a double trailing @ holds a line in the input buffer from one iteration of the DATA step to the next. In a PUT statement, however, a trailing @ has the same effect as a double trailing @; both hold a line across iterations of the DATA step.
- Both the PUT and OUTPUT statements create output in a DATA step. The PUT statement uses an output buffer and writes output lines to an external location, the SAS log, or your monitor. The OUTPUT statement uses the program data vector and writes observations to a SAS data set.

Examples

Example 1: Using Multiple Output Styles in One PUT Statement This example uses several output styles in a single PUT statement:

```
options yearcutoff= 1920;

data club1;
   input idno name $ startwght date : date7.;
   put name 'on ' date mmddyy8. ' weighs '
       startwght +(-1) '.' idno= 32-40;
   datalines;
032 David 180 25nov99
049 Amelia 145 25nov99
219 Alan 210 12nov99
;
```

The following table shows the output style used for each variable in the example:

Variables	Output Style
NAME, STARTWGHT	list output
DATE	formatted output
IDNO	named output

The PUT statement also uses pointer controls and specifies both character strings and variable names.

The program writes the following lines to the SAS log:*

```
----+----1----+----2----+----3----+----4
David on 11/25/99 weighs 180.  idno=1032
Amelia on 11/25/99 weighs 145. idno=1049
Alan on 11/12/99 weighs 210.   idno=1219
```

Blank spaces are inserted at the beginning and the end of the character strings to change the pointer position. These spaces separate the value of a variable from the character string. The +(-1) pointer control moves the pointer backward to remove the unwanted blank that occurs between the value of STARTWGHT and the period. For more information about how to position the pointer, see "Pointer Location After a Value Is Written" on page 1716.

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

Example 2: Moving the Pointer within a Page These PUT statements show how to use column and line pointer controls to position the output pointer.

- To move the pointer to a specific column, use @ followed by the column number, variable, or expression whose value is that column number. For example, this statement moves the pointer to column 15 and writes the value of TOTAL SALES using list output:

```
put @15 totalsales;
```

  This PUT statement moves the pointer to the value that is specified in COLUMN and writes the value of TOTALSALES with the COMMA6 format:

```
data _null_;
   set carsales;
   column=15;
   put @column totalsales comma6.;
run;
```

- This program shows two techniques to move the pointer backward:

```
data carsales;
   input item $10. jan : comma5.
         feb : comma5. mar : comma5.;
   saleqtr1=sum(jan,feb,mar);
/* an expression moves pointer backward */
   put '1st qtr sales for ' item
       'is ' saleqtr1 : comma6. +(-1) '.';
/* a numeric variable with a negative
    value moves pointer backward.      */
   x=-1;
   put '1st qtr sales for ' item
       'is ' saleqtr1 : comma5. +x '.';
   datalines;
trucks         1,382     2,789     3,556
vans           1,265     2,543     3,987
sedans         2,391     3,011     3,658
;
```

  Because the value of SALEQTR1 is written with modified list output, the pointer moves automatically two spaces. For more information, see "How Modified List Output and Formatted Output Differ" on page 1733. To remove the unwanted blank that occurs between the value and the period, move the pointer backward by one space.

  The program writes the following lines to the SAS log:*

```
----+----1----+----2----+----3----+----4
st qtr sales for trucks is 7,727.
st qtr sales for trucks is 7,727.
st qtr sales for vans is 7,795.
st qtr sales for vans is 7,795.
st qtr sales for sedans is 9,060.
st qtr sales for sedans is 9,060.
```

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

□ This program uses a PUT statement with the / line pointer control to advance to the next output line:

```
data _null_;
   set carsales end=lastrec;
   totalsales+saleqtr1;
   if lastrec then
     put @2 'Total Sales for 1st Qtr'
         / totalsales 10-15;
run;
```

After the DATA step calculates TOTALSALES using all the observations in the CARSALES data set, the PUT statement executes. It writes a character string beginning in column 2 and moves to the next line to write the value of TOTALSALES in columns 10 through 15:*

```
----+----1----+----2----+----3
 Total Sales for 1st Qtr
          24582
```

Example 3: Moving the Pointer to a New Page This example creates a data set called STATEPOP, which contains information from the 1990 U.S. census about the population of metropolitan and non-metropolitan areas. It executes the FORMAT procedure to group the 50 states and the District of Columbia into four regions. It then uses the IF and PUT statements to control the printed output.

```
options pagesize=24 linesize=64 nodate pageno=1;

title1;

data statepop;
   input state $ cityp90 ncityp90 region @@;
   label cityp90= '1990 metropolitan population
                   (million)'
         ncityp90='1990 nonmetropolitan population
                   (million)'
         region=  'Geographic region';
   datalines;
ME    .443    .785  1   NH    .659    .450  1
VT    .152    .411  1   MA   5.788    .229  1
RI    .938    .065  1   CT   3.148    .140  1
NY  16.515   1.475  1   NJ   7.730     .A   1
PA  10.083   1.799  1   DE    .553    .113  2
MD   4.439    .343  2   DC    .607      .   2
VA   4.773   1.414  2   WV    .748   1.045  2
NC   4.376   2.253  2   SC   2.423   1.064  2
GA   4.352   2.127  2   FL  12.023    .915  2
KY   1.780   1.906  2   TN   3.298   1.579  2
AL   2.710   1.331  2   MS    .776   1.798  2
AR   1.040   1.311  2   LA   3.160   1.060  2
OK   1.870   1.276  2   TX  14.166   2.821  2
OH   8.826   2.021  3   IN   3.962   1.582  3
IL   9.574   1.857  3   MI   7.698   1.598  3
WI   3.331   1.561  3   MN   3.011   1.364  3
IA   1.200   1.577  3   MO   3.491   1.626  3
```

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

```
ND    .257    .381  3   SD    .221    .475  3
NE    .787    .791  3   KS   1.333   1.145  3
MT    .191    .608  4   ID    .296    .711  4
WY    .134    .319  4   CO   2.686    .608  4
NM    .842    .673  4   AZ   3.106    .559  4
UT   1.336    .387  4   NV   1.014    .183  4
WA   4.036    .830  4   OR   1.985    .858  4
CA  28.799    .961  4   AK    .226    .324  4
HI    .836    .272  4
;

proc format;
   value regfmt 1='Northeast'
                2='South'
                3='Midwest'
                4='West';
run;

data _null_;
   set statepop;
   by region;
   pop90=sum(cityp90,ncityp90);
   file print;
   put state 1-2 @5 pop90 7.3 ' million';
   if first.region then
      regioncitypop=0;      /* new region */
   regioncitypop+cityp90;
   if last.region then
      do;
         put // '1990 US CENSUS for ' region regfmt.
              / 'Total Urban Population: '
                 regioncitypop' million' _page_;
      end;
run;
```

Output 6.27 PUT Statement Output for the Northeast Region

```
                                                                    1
ME    1.228 million
NH    1.109 million
VT    0.563 million
MA    6.017 million
RI    1.003 million
CT    3.288 million
NY   17.990 million
NJ    7.730 million
PA   11.882 million

1990 US CENSUS for Northeast
Total Urban Population: 45.456  million
```

PUT _PAGE_ advances the pointer to line 1 of the new page when the value of LAST.REGION is 1. The example prints a footer message before exiting the page.

Example 4: Underlining Text This example uses OVERPRINT to underscore a value written by a previous PUT statement:

```
data _null_;
   input idno name $ startwght;
   file file-specification print;
   put name 1-10 @15 startwght 3.;
   if startwght > 200 then
     put overprint @15 '___';
   datalines;
032 David 180
049 Amelia 145
219 Alan 210
;
```

The second PUT statement underlines weights above 200 on the output line the first PUT statement prints.

This PUT statement uses OVERPRINT with both a column pointer control and a line pointer control:

```
   put @5 name $8. overprint @5 8*'_'
     / @20 address;
```

The PUT statement writes a NAME value, underlines it by overprinting eight underscores, and moves the output pointer to the next line to write an ADDRESS value.

Example 5: Holding and Releasing Output Lines This DATA step demonstrates how to hold and release an output line with a PUT statement:

```
data _null_;
   input idno name $ startwght 3.;
   put name @;
   if startwght ne . then
     put @15 startwght;
   else put;
   datalines;
032 David 180
049 Amelia 145
126 Monica
219 Alan 210
;
```

In this example,

- the trailing @ in the first PUT statement holds the current output line after the value of NAME is written
- if the condition is met in the IF-THEN statement, the second PUT statement writes the value of STARTWGHT and releases the current output line
- if the condition is not met, the second PUT never executes. Instead, the ELSE PUT statement executes. The ELSE PUT statement releases the output line and positions the output pointer at column 1 in the output buffer.

The program writes the following lines to the SAS log:*

```
----+----1----+----2
David         180
```

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

```
Amelia        145
Monica
Alan          210
```

Example 6: Writing the Current Input Record to the Log When a value for ID is less than 1000, PUT _INFILE_ executes and writes the current input record to the SAS log. The DELETE statement prevents the DATA step from writing the observation to the TEAM data set.

```
data team;
   input id team $ score1 score2;
   if id le 1000 then
     do;
       put _infile_;
       delete;
     end;
   datalines;
032 red 180 165
049 yellow 145 124
219 red 210 192
;
```

The program writes the following line to the SAS log:*

```
----+----1----+----2
219 red 210 192
```

Example 7: Avoiding a Common Error When Writing a Character Constant Followed by a Variable This example illustrates how to use a PUT statement to write character constants and variable values without causing them to be misinterpreted as SAS name literals. A SAS name literal is a name token that is expressed as a string within quotation marks, followed by the letter n. For more information about SAS name literals, see *SAS Language Reference: Concepts*.

In the program below, the PUT statement writes the constant 'n' followed by the value of the variable NVAR1, and then writes another constant 'n':

```
data _null_;
   n=5;
   nvar1=1;
   var1=7;
   put @1 'n' nvar1 'n';
run;
```

This program writes the following line to the SAS log:*

```
----+----1----+----2
n1 n
```

If all the spaces between the constants and the variables are removed from the previous PUT statement, SAS interprets 'n' as a name literal instead of reading 'n' as a constant. The next variable is read as VAR1 instead of NVAR1. The final 'n' constant is interpreted correctly.

```
put @1 'n'nvar1'n';
```

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

This PUT statement writes the following line to the SAS log:*

```
----+----1----+----2
5 7 n
```

To print character constants and variable values without intervening spaces, and without potential misinterpretation, you can add spaces between them and use pointer controls where necessary. For example, the following PUT statement uses a pointer control to write the correct character constants and variable values but does not insert blank spaces. Note that +(-1) moves the PUT statement pointer backwards by one space.

```
put @1 'n' nvar1 +(-1) 'n';
```

This PUT statement writes the following line to the SAS log:*

```
----+----1----+----2
n1n
```

See Also

Statements:

"FILE Statement" on page 1503
"PUT Statement, Column" on page 1724
"PUT Statement, Formatted" on page 1727
"PUT Statement, List" on page 1731
"PUT Statement, Named" on page 1736
PUT Statement for ODS

System Options:

"LINESIZE= System Option" on page 1936
"PAGESIZE= System Option" on page 1958

PUT Statement, Column

Writes variable values in the specified columns in the output line.

Valid: in a DATA step

Category: File-handling

Type: Executable

Syntax

PUT *variable start-column* <— *end-column*>
<.*decimal–places*> <@ | @@>;

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

Arguments

variable
specifies the variable whose value is written.

start-column
specifies the first column of the field where the value is written in the output line.

— *end-column*
specifies the last column of the field for the value.

Tip: If the value occupies only one column in the output line, omit *end-column*.

Example: Because *end-column* is omitted, the values for the character variable GENDER occupy only column 16:

```
put name 1-10 gender 16;
```

.decimal-places
specifies the number of digits to the right of the decimal point in a numeric value.

Range: positive integer

Tip: If you specify 0 for *d* or omit *d*, the value is written without a decimal point.

Featured in: "Examples" on page 1726

@| @@
holds an output line for the execution of the next PUT statement even across iterations of the DATA step. These line-hold specifiers are called *trailing* @ and *double trailing* @.

Requirement: The trailing @ or double trailing @ must be the last item in the PUT statement.

See: "Using Line-Hold Specifiers" on page 1716

Details

With column output, the column numbers indicate the position that each variable value will occupy in the output line. If a value requires fewer columns than specified, a character variable is left-aligned in the specified columns, and a numeric variable is right-aligned in the specified columns.

There is no limit to the number of column specifications you can make in a single PUT statement. You can write anywhere in the output line, even if a value overwrites columns that were written earlier in the same statement. You can combine column output with any of the other output styles in a single PUT statement. For more information, see "Using Multiple Output Styles in a Single PUT Statement" on page 1714.

Examples

Use column output in the PUT statement as shown here.

□ This PUT statement uses column output:

```
data _null_;
   input name $ 1-18 score1 score2 score3;
   put name 1-20 score1 23-25 score2 28-30
       score3 33-35;
   datalines;
Joseph                  11   32   76
Mitchel                 13   29   82
Sue Ellen               14   27   74
;
```

The program writes the following lines to the SAS log:*

```
----+----1----+----2----+----3----+----4
Joseph                 11   32   76
Mitchel                13   29   82
Sue Ellen              14   27   74
```

The values for the character variable NAME begin in column 1, the left boundary of the specified field (columns 1 through 20). The values for the numeric variables SCORE1 through SCORE3 appear flush with the right boundary of their field.

□ This statement produces the same output lines, but writes the SCORE1 value first and the NAME value last:

```
put score1 23-25 score2 28-30
    score3 33-35 name $ 1-20;
```

□ This DATA step specifies decimal points with column output:

```
data _null_;
    x=11;
    y=15;
    put x 10-18 .1 y 20-28 .1;
run;
```

This program writes the following line to the SAS log:*

```
----+----1----+----2----+----3----+----4
              11.0      15.0
```

See Also

Statement:

"PUT Statement" on page 1708

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

PUT Statement, Formatted

Writes variable values with the specified format in the output line.

Valid: in a DATA step
Category: File-handling
Type: Executable

Syntax

PUT <*pointer-control*> *variable format.* <@ | @@>;

PUT <*pointer-control*> (*variable-list*) (*format-list*)
<@ | @@>;

Arguments

pointer-control
moves the output pointer to a specified line or column.

See: "Column Pointer Controls" on page 1710 and "Line Pointer Controls" on page 1711

Featured in: Example 1 on page 1730

variable
specifies the variable whose value is written.

(*variable-list*)
specifies a list of variables whose values are written.

Requirement: The (*format-list*) must follow the (*variable-list*).

See: "How to Group Variables and Formats" on page 1729

Featured in: Example 1 on page 1730

format.
specifies a format to use when the variable values are written. To override the default alignment, you can add an alignment specification to a format:

-L left aligns the value.

-C centers the value.

-R right aligns the value.

Tip: Ensure that the format width provides enough space to write the value and any commas, dollar signs, decimal points, or other special characters that the format includes.

Example: This PUT statement uses the format dollar7.2 to write the value of X:

```
put x dollar7.2;
```

When X is 100, the formatted value uses seven columns:

```
$100.00
```

Featured in: Example 2 on page 1730

(*format-list*)
specifies a list of formats to use when the values of the preceding list of variables are written. In a PUT statement, a *format-list* can include

format.
specifies the format to use to write the variable values.

Tip: You can specify either a SAS format or a user-written format. See Chapter 3, "Formats," on page 81.

pointer-control
specifies one of these pointer controls to use to position a value: @, #, /, +, and OVERPRINT.

Example: Example 1 on page 1730

character-string
specifies one or more characters to place between formatted values.

Example: This statement places a hyphen between the formatted values of CODE1, CODE2, and CODE3:

```
put bldg $ (code1 code2 code3) (3. '-');
```

See: Example 1 on page 1730

*n**
specifies to repeat *n* times the next format in a format list.

Example: This statement uses the 7.2 format to write GRADES1, GRADES2, and GRADES3 and the 5.2 format to write GRADES4 and GRADES5:

```
put (grades1-grades5) (3*7.2, 2*5.2);
```

Restriction: The (*format-list*) must follow (*variable-list*).

See Also: "How to Group Variables and Formats" on page 1729

@| @@
holds an output line for the execution of the next PUT statement even across iterations of the DATA step. These line-hold specifiers are called *trailing* @ and *double trailing* @.

Restriction: The trailing @ or double trailing @ must be the last item in the PUT statement.

See: "Using Line-Hold Specifiers" on page 1716

Details

Using Formatted Output The Formatted output describes the output lines by listing the variable names and the formats to use to write the values. You can use a SAS format or a user-written format to control how SAS prints the variable values. For a complete description of the SAS formats, see "Definition of Formats" on page 84.

With formatted output, the PUT statement uses the format that follows the variable name to write each value. SAS does not automatically add blanks between values. If the value uses fewer columns than specified, character values are left-aligned and numeric values are right-aligned in the field that is specified by the format width.

Formatted output, combined with pointer controls, makes it possible to specify the exact line and column location to write each variable. For example, this PUT statement uses the dollar7.2 format and centers the value of X starting at column 12:

```
put @12 x dollar7.2-c;
```

How to Group Variables and Formats When you want to write values in a pattern on the output lines, use format lists to shorten your coding time. A format list consists of the corresponding formats separated by either blanks or commas and enclosed in parentheses. It must follow the names of the variables enclosed in parentheses.

For example, this statement uses a format list to write the five variables SCORE1 through SCORE5, one after another, using four columns for each value with no blanks in between:

```
put (score1-score5) (4. 4. 4. 4. 4.);
```

A shorter version of the previous statement is

```
put (score1-score5) (4.);
```

You can include any of the pointer controls (@, #, /, +, and OVERPRINT) in the list of formats, as well as *n**, and a character string. You can use as many format lists as necessary in a PUT statement, but do not nest the format lists. After all the values in the variable list are written, the PUT statement ignores any directions that remain in the format list. For an example, see Example 3 on page 1730.

You can also specify a reference to all elements in an array as (*array-name* {*}), followed by a list of formats. You cannot, however, specify the elements in a _TEMPORARY_ array in this way. This PUT statement specifies an array name and a format list:

```
put (array1{*}) (4.);
```

For more information about how to reference an array, see "Arrays" on page 1717.

Examples

Example 1: Writing a Character between Formatted Values This example formats some values and writes a - (hyphen) between the values of variables BLDG and ROOM:

```
data _null_;
   input name & $15. bldg $ room;
   put name @20 (bldg room) ($1. "-" 3.);
   datalines;
Bill Perkins  J 126
Sydney Riley  C 219
;
```

These lines are written to the SAS log:

```
Bill Perkins       J-126
Sydney Riley       C-219
```

Example 2: Overriding the Default Alignment of Formatted Values This example includes an alignment specification in the format:

```
data _null_;
   input name $ 1-12 score1 score2 score3;
   put name $12.-r +3 score1 3. score2 3.
       score3 4.;
   datalines;
Joseph                  11   32   76
Mitchel                 13   29   82
Sue Ellen               14   27   74
;
```

These lines are written to the log:

```
----+----1----+----2----+----3----+----4
      Joseph    11 32  76
     Mitchel    13 29  82
   Sue Ellen    14 27  74
```

The value of the character variable NAME is right-aligned in the formatted field. (Left alignment is the default for character variables.)

Example 3: Including More Format Specifications Than Necessary This format list includes more specifications than are necessary when the PUT statement executes:

```
data _null_;
   input x y z;
   put (x y z) (2.,+1);
   datalines;
2 24 36
0 20 30
;
```

The PUT statement writes the value of X using the 2. format. Then, the +1 column pointer control moves the pointer forward one column. Next, the value of Y is written with the 2. format. Again, the +1 column pointer moves the pointer forward one column. Then, the value of Z is written with the 2. format. For the third iteration, the PUT statement ignores the +1 pointer control.

These lines are written to the SAS log: *

```
----+----1----+
2 24 36
0 20 30
```

See Also

Statement:

"PUT Statement" on page 1708

PUT Statement, List

Writes variable values and the specified character strings in the output line.

Valid: in a DATA step

Category: File-handling

Type: Executable

Syntax

PUT <*pointer-control*> *variable* <@ | @@>;

PUT <*pointer-control*> <*n**>*'character-string'* <@ | @@>;

PUT <*pointer-control*> *variable* <: | ~> *format.*<@ | @@>;

Arguments

pointer-control

moves the output pointer to a specified line or column.

See: "Column Pointer Controls" on page 1710 and "Line Pointer Controls" on page 1711

Featured in: Example 2 on page 1734

variable

specifies the variable whose value is written.

Featured in: Example 1 on page 1734

* The ruled line is for illustrative purposes only; the PUT statement does not generate it.

n*

specifies to repeat *n* times the subsequent character string.

Example: This statement writes a line of 132 underscores:

```
put 132*'_';
```

'character-string'

specifies a string of text, enclosed in quotation marks, to write.

Interaction: When insufficient space remains on the current line to write the entire text string, SAS withholds the entire string and writes the current line. Then it writes the text string on a new line, starting in column 1. For more information, see "When the Pointer Goes Past the End of a Line" on page 1717.

Tip: To avoid misinterpretation, always put a space after a closing quotation mark in a PUT statement.

Tip: If you follow a quotation mark with X, SAS interprets the text string as a hexadecimal constant.

Tip: If you use single quotation (‘) or double quotes (“) together (with no space in between them) as the string of text, SAS will output a single quotation mark (’)or double quotation mark (“), respectively.

See Also: "How List Output Is Spaced" on page 1733

Featured in: Example 2 on page 1734

:

enables you to specify a format that the PUT statement uses to write the variable value. All leading and trailing blanks are deleted, and each value is followed by a single blank.

Requirement: You must specify a format.

See: "How Modified List Output and Formatted Output Differ" on page 1733

Featured in: Example 3 on page 1735

~

enables you to specify a format that the PUT statement uses to write the variable value. SAS displays the formatted value in quotation marks even if the formatted value does not contain the delimiter. SAS deletes all leading and trailing blanks, and each value is followed by a single blank. Missing values for character variables are written as a blank (" ") and, by default, missing values for numeric variables are written as a period (".").

Requirement: You must specify the DSD option in the FILE statement.

Featured in: Example 4 on page 1735

format.

specifies a format to use when the data values are written.

Tip: You can specify either a SAS format or a user-written format. See Chapter 3, "Formats," on page 81.

Featured in: Example 3 on page 1735

@ | @@

holds an output line for the execution of the next PUT statement even across iterations of the DATA step. These line-hold specifiers are called *trailing* @ and *double trailing* @.

Restriction: The trailing @ or double-trailing @ must be the last item in the PUT statement.

See: "Using Line-Hold Specifiers" on page 1716

Details

Using List Output With list output, you list the names of the variables whose values you want written, or you specify a character string in quotation marks. The PUT statement writes a variable value, inserts a single blank, and then writes the next value. Missing values for numeric variables are written as a single period. Character values are left-aligned in the field; leading and trailing blanks are removed. To include blanks (in addition to the blank inserted after each value), use formatted or column output instead of list output.

There are two types of list output:

- simple list output
- modified list output.

Modified list output increases the versatility of the PUT statement because you can specify a format to control how the variable values are written. See Example 3 on page 1735.

How List Output Is Spaced List output uses different spacing methods when it writes variable values and character strings. When a variable is written with list output, SAS automatically inserts a blank space. The output pointer stops at the second column that follows the variable value. However, when a character string is written, SAS does not automatically insert a blank space. The output pointer stops at the column that immediately follows the last character in the string.

To avoid spacing problems when both character strings and variable values are written, you might want to use a blank space as the last character in a character string. When a character string that provides punctuation follows a variable value, you need to move the output pointer backward. Moving the output pointer backward prevents an unwanted space from appearing in the output line. See Example 2 on page 1734.

Comparisons

How Modified List Output and Formatted Output Differ List output and formatted output use different methods to determine how far to move the pointer after a variable value is written. Therefore, modified list output, which uses formats, and formatted output produce different results in the output lines. Modified list output writes the value, inserts a blank space, and moves the pointer to the next column. Formatted output moves the pointer the length of the format, even if the value does not fill that length. The pointer moves to the next column; an intervening blank is not inserted.

The following DATA step uses modified list output to write each output line:

```
data _null_;
   input x y;
   put x : comma10.2 y : 7.2;
   datalines;
2353.20 7.10
6231 121
;
```

These lines are written to the SAS log:

```
----+----1----+----2
2,353.20 7.10
6,231.00 121.00
```

In comparison, the following example uses formatted output:

```
put x comma10.2 y 7.2;
```

These lines are written to the SAS log, with the values aligned in columns:

```
----+----1----+----2
  2,353.20   7.10
  6,231.00 121.00
```

Examples

Example 1: Writing Values with List Output

This DATA step uses a PUT statement with list output to write variable values to the SAS log:

```
data _null_;
   input name $ 1-10 sex $ 12 age 15-16;
   put name sex age;
   datalines;
Joseph     M  13
Mitchel    M  14
Sue Ellen  F  11
;
```

These lines are written to the SAS log:

```
----+----1----+----2----+----3----+----4
Joseph M 13
Mitchel M 14
Sue Ellen F 11
```

By default, the values of the character variable NAME are left-aligned in the field.

Example 2: Writing Character Strings and Variable Values This PUT statement adds a space to the end of a character string and moves the output pointer backward to prevent an unwanted space from appearing in the output line after the variable STARTWGHT:

```
data _null_;
   input idno name $ startwght;
   put name 'weighs ' startwght +(-1) '.';
   datalines;
032 David 180
049 Amelia 145
219 Alan 210
;
```

These lines are written to the SAS log:

```
David weighs 180.
Amelia weighs 145.
Alan weighs 210.
```

The blank space at the end of the character string changes the pointer position. This space separates the character string from the value of the variable that follows. The +(-1) pointer control moves the pointer backward to remove the unwanted blank that occurs between the value of STARTWGHT and the period.

Example 3: Writing Values with Modified List Output (:) This DATA step uses modified list output to write several variable values in the output line using the : argument:

```
data _null_;
   input salesrep : $10. tot : comma6. date : date9.;
   put 'Week of ' date : worddate15.
       salesrep : $12. 'sales were '
       tot : dollar9.  + (-1) '.';
   datalines;
Wong 15,300 12OCT2004
Hoffman 9,600 12OCT2004
;
```

These lines are written to the SAS log:

```
Week of Oct 12, 2004 Wong sales were $15,300.
Week of Oct 12, 2004 Hoffman sales were $9,600.
```

Example 4: Writing Values with Modified List Output and ~ This DATA step uses modified list output to write several variable values in the output line using the ~ argument:

```
data _null_;
   input salesrep : $10. tot : comma6. date : date9.;
   file log delimiter=" " dsd;
   put 'Week of ' date ~ worddate15.
       salesrep ~ $12. 'sales were '
       tot ~ dollar9.  + (-1) '.';
   datalines;
Wong 15,300 12OCT2004
Hoffman 9,600 12OCT2004
;
```

These lines are written to the SAS log:

```
Week of "Oct 12, 2004" "Wong" sales were "$15,300".
Week of "Oct 12, 2004" "Hoffman" sales were "$9,600".
```

See Also

Statements:

"PUT Statement" on page 1708

"PUT Statement, Formatted" on page 1727

PUT Statement, Named

Writes variable values after the variable name and an equal sign.

Valid: in a DATA step
Category: File-handling
Type: Executable

Syntax

PUT <*pointer-control*> *variable*= <*format.*> <@ | @@>;

PUT *variable*= *start-column* <— *end-column*>
<.*decimal-places*> <@ | @@>;

Arguments

pointer-control
moves the output pointer to a specified line or column in the output buffer.

See: "Column Pointer Controls" on page 1710 and "Line Pointer Controls" on page 1711

variable=
specifies the variable whose value is written by the PUT statement in the form

```
variable=value
```

format.
specifies a format to use when the variable values are written.

Tip: Ensure that the format width provides enough space to write the value and any commas, dollar signs, decimal points, or other special characters that the format includes.

Example: This PUT statement uses the format DOLLAR7.2 to write the value of X:

```
put x= dollar7.2;
```

When X=100, the formatted value uses seven columns:

```
X=$100.00
```

See: "Formatting Named Output" on page 1737

start-column
specifies the first column of the field where the variable name, equal sign, and value are to be written in the output line.

— end-column
determines the last column of the field for the value.

Tip: If the variable name, equal sign, and value require more space than the columns specified, PUT will write past the end column rather than truncate the value. You must leave enough space before beginning the next value.

.decimal-places
specifies the number of digits to the right of the decimal point in a numeric value. If you specify 0 for *d* or omit *d*, the value is written without a decimal point.

Range: positive integer

@ | @@
holds an output line for the execution of the next PUT statement even across iterations of the DATA step. These line-hold specifiers are called *trailing* @ and *double trailing* @.

Restriction: The trailing @ or double trailing @ must be the last item in the PUT statement.

See: "Using Line-Hold Specifiers" on page 1716

Details

Using Named Output With named output, follow the variable name with an equal sign in the PUT statement. You can use either list output, column output, or formatted output specifications to indicate how to position the variable name and values. To insert a blank space between each variable value automatically, use list output. To align the output in columns, use pointer controls or column specifications.

Formatting Named Output You can specify either a SAS format or a user-written format to control how SAS prints the variable values. The width of the format does *not* include the columns required by the variable name and equal sign. To align a formatted value, SAS deletes leading blanks and writes the variable value immediately after the equal sign. SAS does not align on the right side of the formatted length, as in unnamed formatted output.

For a complete description of the SAS formats, see "Definition of Formats" on page 84.

Examples

Use named output in the PUT statement as shown here.

□ This PUT combines named output with column pointer controls to align the output:

```
data _null_;
   input name $ 1-18 score1 score2 score3;
   put name = @20 score1= score3= ;
   datalines;
Joseph                  11   32   76
Mitchel                 13   29   82
Sue Ellen               14   27   74
;
```

The program writes the following lines to the SAS log:

```
----+----1----+----2----+----3----+----4
NAME=Joseph         SCORE1=11 SCORE3=76
NAME=Mitchel        SCORE1=13 SCORE3=82
NAME=Sue Ellen      SCORE1=14 SCORE3=74
```

□ This example specifies an output format for the variable AMOUNT:

```
put item= @25 amount= dollar12.2;
```

When the value of ITEM is binders and the value of AMOUNT is 153.25, this output line is produced:

```
----+----1----+----2----+----3----+----4
ITEM=binders            AMOUNT=$153.25
```

See Also

Statement:

"PUT Statement" on page 1708

PUTLOG Statement

Writes a message to the SAS log.

Valid: in a DATA step

Category: Action

Type: Executable

Syntax

PUTLOG '*message*';

Arguments

message

specifies the message that you want to write to the SAS log. *Message* can include character literals (enclosed in quotation marks), variable names, formats, and pointer controls.

Tip: You can precede your message text with WARNING, MESSAGE, or NOTE to better identify the output in the log.

Details

The PUTLOG statement writes a message that you specify to the SAS log. The PUTLOG statement is also helpful when you use macro-generated code because you can send output to the SAS log without affecting the current file destination.

Comparisons

The PUTLOG statement is similar to the ERROR statement except that PUTLOG does not set _ERROR_ to 1.

Examples

Example 1: Writing Messages to the SAS Log Using the PUTLOG Statement The following program creates the computeAverage92 macro, which computes the average score, validates input data, and uses the PUTLOG statement to write error messages to the SAS log. The DATA step uses the PUTLOG statement to write a warning message to the log.

```
data ExamScores;
   input Name $ 1-16 Score1 Score2 Score3;
   datalines;
Sullivan, James   86 92 88
```

```
Martinez, Maria   95 91 92
Guzik, Eugene     99 98 .
Schultz, John     90 87 93
van Dyke, Sylvia  98 . 91
Tan, Carol        93 85 85
;

options pageno=1 nodate linesize=80 pagesize=60;
filename outfile 'your-output-file';

   /* Create a macro that computes the average score, validates  */
   /* input data, and uses PUTLOG to write error messages to the */
   /* SAS log.                                                   */
%macro computeAverage92(s1, s2, s3, avg);
   if &s1 < 0 or &s2 < 0 or &s3 < 0 then
      do;
         putlog 'ERROR: Invalid score data ' &s1= &s2= &s3=;
         &avg = .;
      end;
   else
      &avg = mean(&s1, &s2, &s3);
%mend;

data _null_;
set ExamScores;
   file outfile;
   %computeAverage92(Score1, Score2, Score3, AverageScore);
   put name Score1 Score2 Score3 AverageScore;

      /* Use PUTLOG to write a warning message to the SAS log. */
   if AverageScore < 92 then
       putlog 'WARNING: Score below the minimum ' name= AverageScore= 5.2;
run;

proc print;
run;
```

The following lines are written to the SAS log.

Output 6.28 SAS Log Results from the PUTLOG Statement

```
WARNING: Score below the minimum Name=Sullivan, James AverageScore=88.67
ERROR: Invalid score data Score1=99 Score2=98 Score3=.
WARNING: Score below the minimum Name=Guzik, Eugene AverageScore=.
WARNING: Score below the minimum Name=Schultz, John AverageScore=90.00
ERROR: Invalid score data Score1=98 Score2=. Score3=91
WARNING: Score below the minimum Name=van Dyke, Sylvia AverageScore=.
WARNING: Score below the minimum Name=Tan, Carol AverageScore=87.67
```

SAS creates the following output file.

Output 6.29 Individual Examination Scores

Exam Scores 1

Obs	Name	Score1	Score2	Score3
1	Sullivan, James	86	92	88
2	Martinez, Maria	95	91	92
3	Guzik, Eugene	99	98	.
4	Schultz, John	90	87	93
5	van Dyke, Sylvia	98	.	91
6	Tan, Carol	93	85	85

See Also

Statement:

"ERROR Statement" on page 1502

REDIRECT Statement

Points to different input or output SAS data sets when you execute a stored program.

Valid: in a DATA step

Category: Action

Type: Executable

Requirement: You must specify the PGM= option in the DATA statement.

Syntax

REDIRECT INPUT | OUTPUT *old-name-1* = *new-name-1*<. . . *old-name-n* = *new-name-n*>;

Arguments

INPUT | OUTPUT
specifies whether to redirect input or output data sets. When you specify INPUT, the REDIRECT statement associates the name of the input data set in the source program with the name of another SAS data set. When you specify OUTPUT, the REDIRECT statement associates the name of the output data set with the name of another SAS data set.

old-name
specifies the name of the input or output data set in the source program.

new-name
specifies the name of the input or output data set that you want SAS to process for the current execution.

Details

The REDIRECT statement is available only when you execute a stored program. For more information about stored programs, see "Stored Compiled DATA Step Programs" in *SAS Language Reference: Concepts*.

CAUTION:

Use care when you redirect input data sets. The number and attributes of variables in the input data sets that you read with the REDIRECT statement should match the number and attributes of variables in the input data sets in the MERGE, SET, MODIFY, or UPDATE statements of the source code. If the variable type attributes differ, the stored program stops processing and an appropriate error message is written to the SAS log. If the variable length attributes differ, the length of the variable in the source code data set determines the length of the variable in the redirected data set. Extra variables in the redirected data sets cause the stored program to stop processing and an error message is written to the SAS log. △

Tip: The DROP or KEEP data set options can be added in the stored program if the input data set that you read with the REDIRECT statement has more variables than are in the data set used to compile the program.

Comparison

The REDIRECT statement applies only to SAS data sets. To redirect input and output stored in external files, include a FILENAME statement to associate the fileref in the source program with different external files.

Examples

This example executes the stored program called STORED.SAMPLE. The REDIRECT statement specifies the source of the input data as BASE.SAMPLE. The output data set from this execution of the program is redirected and stored in a data set named SUMS.SAMPLE.

```
libname stored 'SAS-library';
libname base 'SAS-library';
libname sums 'SAS-library';

data pgm=stored.sample;
   redirect input in.sample=base.sample;
   redirect output out.sample=sums.sample;
run;
```

See Also

Statement:

"DATA Statement" on page 1465

"Stored Compiled DATA Step Programs" in *SAS Language Reference: Concepts*.

REMOVE Statement

Deletes an observation from a SAS data set.

Valid: in a DATA step
Category: Action
Type: Executable
Restriction: Use only with a MODIFY statement.

Syntax

REMOVE <*data-set-name(s)*>;

Without Arguments

If you specify no argument, the REMOVE statement deletes the current observation from all data sets that are named in the DATA statement.

Arguments

data-set-name
specifies the data set in which the observation is deleted.

Restriction: The data set name must also appear in the DATA statement and in one or more MODIFY statements.

Details

The deletion of an observation can be physical or logical, depending on the engine that maintains the data set. Using REMOVE overrides the default replacement of observations. If a DATA step contains a REMOVE statement, you must explicitly program all output for the step.

Comparisons

- Using an OUTPUT, REPLACE, or REMOVE statement overrides the default write action at the end of a DATA step. (OUTPUT is the default action; REPLACE becomes the default action when a MODIFY statement is used.) If you use any of these statements in a DATA step, you must explicitly program all output for new observations.
- The OUTPUT, REPLACE, and REMOVE statements are independent of each other. More than one statement can apply to the same observation, as long as the sequence is logical.
- If both an OUTPUT and a REPLACE or REMOVE statement execute on a given observation, perform the OUTPUT action last to keep the position of the observation pointer correct.
- Because the REMOVE statement can perform a physical or a logical deletion, REMOVE is available with the MODIFY statement for all SAS data set engines. Both the DELETE and subsetting IF statements perform only physical deletions. Therefore, they are not available with the MODIFY statement for certain engines.

Examples

This example removes one observation from a SAS data set.

```
libname perm 'SAS-library';
```

```
data perm.accounts;
   input AcctNumber Credit;
   datalines;
1001 1500
1002 4900
1003 3000
;

data perm.accounts;
   modify perm.accounts;
   if AcctNumber=1002 then remove;
run;

proc print data=perm.accounts;
   title 'Edited Data Set';
run;
```

Here are the results of the PROC PRINT statement:

```
                        Edited Data Set                          1

                             Acct
                   OBS     Number    Credit

                    1       1001      1500
                    3       1003      3000
```

See Also

Statements:

"DELETE Statement" on page 1486
"IF Statement, Subsetting" on page 1581
"MODIFY Statement" on page 1684
"OUTPUT Statement" on page 1704
"REPLACE Statement" on page 1745

RENAME Statement

Specifies new names for variables in output SAS data sets.

Valid: in a DATA step
Category: Information
Type: Declarative

Syntax

RENAME *old-name-1=new-name-1 . . . <old-name-n=new-name-n>*;

Arguments

old-name
specifies the name of a variable or variable list as it appears in the input data set, or in the current DATA step for newly created variables.

new-name
specifies the name or list to use in the output data set.

Details

The RENAME statement allows you to change the names of one or more variables, variables in a list, or a combination of variables and variable lists. The new variable names are written to the output data set only. Use the old variable names in programming statements for the current DATA step. RENAME applies to all output data sets.

Note: The RENAME statement has an effect on data sets opened in output mode only. △

Comparisons

- RENAME cannot be used in PROC steps, but the RENAME= data set option can.
- The RENAME= data set option allows you to specify the variables you want to rename for each input or output data set. Use it in input data sets to rename variables before processing.
- If you use the RENAME= data set option in an output data set, you must continue to use the old variable names in programming statements for the current DATA step. After your output data is created, you can use the new variable names.
- The RENAME= data set option in the SET statement renames variables in the input data set. You can use the new names in programming statements for the current DATA step.
- To rename variables as a file management task, use the DATASETS procedure or access the variables through the SAS windowing interface. These methods are simpler and do not require DATA step processing.

Examples

- These examples show the correct syntax for renaming variables using the RENAME statement:
  - **rename street=address;**
  - **rename time1=temp1 time2=temp2 time3=temp3;**
  - **rename name=Firstname score1-score3=Newscore1-Newscore3;**
- This example uses the old name of the variable in program statements. The variable Olddept is named Newdept in the output data set, and the variable Oldaccount is named Newaccount.

```
rename Olddept=Newdept Oldaccount=Newaccount;
if Oldaccount>5000;
keep Olddept Oldaccount items volume;
```

- This example uses the old name OLDACCNT in the program statements. However, the new name NEWACCNT is used in the DATA statement because SAS applies the RENAME statement before it applies the KEEP= data set option.

```
data market(keep=newdept newaccnt items
           volume);
   rename olddept=newdept
          oldaccnt=newaccnt;
   set sales;
   if oldaccnt>5000;
run;
```

- The following example uses both a variable and a variable list to rename variables. New variable names appear in the output data set.

```
data temp;
   input (score1-score3) (2.,+1) name $;
   rename name=Firstname
          score1-score3=Newscore1-Newscore3;
   datalines;
12 24 36 Lisa
22 44 66 Fran
;
```

See Also

Data Set Option:

"RENAME= Data Set Option" on page 52

REPLACE Statement

Replaces an observation in the same location.

Valid: in a DATA step

Category: Action

Type: Executable

Restriction: Use only with a MODIFY statement.

Syntax

REPLACE <*data-set-name-1*><. . .*data-set-name-n*>;

Without Arguments

If you specify no argument, the REPLACE statement writes the current observation to the same physical location from which it was read in all data sets that are named in the DATA statement.

Arguments

data-set-name
specifies the data set to which the observation is written.

Requirement: The data set name must also appear in the DATA statement and in one or more MODIFY statements.

Details

Using an explicit REPLACE statement overrides the default replacement of observations. If a DATA step contains a REPLACE statement, explicitly program all output for the step.

Comparisons

- Using an OUTPUT, REPLACE, or REMOVE statement overrides the default write action at the end of a DATA step. (OUTPUT is the default action; REPLACE becomes the default action when a MODIFY statement is used.) If you use any of these statements in a DATA step, you must explicitly program output of a new observation for the step.
- The OUTPUT, REPLACE, and REMOVE statements are independent of each other. More than one statement can apply to the same observation, as long as the sequence is logical.
- If both an OUTPUT and a REPLACE or REMOVE statement execute on a given observation, perform the OUTPUT action last to keep the position of the observation pointer correct.
- REPLACE writes the observation to the same physical location, while OUTPUT writes a new observation to the end of the data set.
- REPLACE can appear only in a DATA step that contains a MODIFY statement. You can use OUTPUT with or without MODIFY.

Examples

This example updates phone numbers in data set MASTER with values in data set TRANS. It also adds one new observation at the end of data set MASTER. The SYSRC autocall macro tests the value of _IORC_ for each attempted retrieval from MASTER. (SYSRC is part of the SAS autocall macro library.) The resulting SAS data set appears after the code:

```
data master;
   input FirstName $ id $ PhoneNumber;
   datalines;
Kevin ABCjkh 904
Sandi defsns 905
Terry ghitDP 951
Jason jklJWM 962
;

data trans;
   input FirstName $ id $ PhoneNumber;
   datalines;
. ABCjkh 2904
. defsns 2905
Madeline mnombt 2983
;

data master;
   modify master trans;
   by id;
```

```
      /* obs found in master  */
      /* change info, replace */
   if _iorc_ = %sysrc(_sok) then replace;

      /* obs not in master    */
   else if _iorc_ = %sysrc(_dsenmr) then
      do;
         /* reset _error_        */
         _error_=0;
         /* reset _iorc_         */
         _iorc_=0;
         /* output obs to master */
      output;
      end;
run;

proc print data=master;
   title 'MASTER with New Phone Numbers';
run;
```

```
                MASTER with New Phone Numbers                3

                      First                    Phone
                OBS   Name          id        Number

                 1    Kevin       ABCjkh      2904
                 2    Sandi       defsns      2905
                 3    Terry       ghitDP       951
                 4    Jason       jklJWM       962
                 5    Madeline    mnombt      2983
```

See Also

Statements:

"MODIFY Statement" on page 1684
"OUTPUT Statement" on page 1704
"REMOVE Statement" on page 1741

RETAIN Statement

Causes a variable that is created by an INPUT or assignment statement to retain its value from one iteration of the DATA step to the next.

Valid: in a DATA step
Category: Information
Type: Declarative

Syntax

RETAIN <*element-list(s)* <*initial-value(s)* |

(*initial-value-1*) | (*initial-value-list-1*) >
< . . . *element-list-n* <*initial-value-n* |
(*initial-value-n*)| (*initial-value-list-n*)>>>;

Without Arguments

If you do not specify an argument, the RETAIN statement causes the values of all variables that are created with INPUT or assignment statements to be retained from one iteration of the DATA step to the next.

Arguments

element-list
specifies variable names, variable lists, or array names whose values you want retained.

Tip: If you specify _ALL_, _CHAR_, or _NUMERIC_, only the variables that are defined before the RETAIN statement are affected.

Tip: If a variable name is specified *only* in the RETAIN statement and you do not specify an initial value, the variable is *not* written to the data set, and a note stating that the variable is uninitialized is written to the SAS log. If you specify an initial value, the variable *is* written to the data set.

initial-value
specifies an initial value, numeric or character, for one or more of the preceding elements.

Tip: If you omit *initial-value*, the initial value is missing. *Initial-value* is assigned to all the elements that precede it in the list. All members of a variable list, therefore, are given the same initial value.

See Also: (*initial-value*) and (*initial-value-list*)

(*initial-value*)
specifies an initial value, numeric or character, for a single preceding element or for the first in a list of preceding elements.

(*initial-value-list*)
specifies an initial value, numeric or character, for individual elements in the preceding list. SAS matches the first value in the list with the first variable in the list of elements, the second value with the second variable, and so on.

Element values are enclosed in quotation marks. To specify one or more initial values directly, use the following format:

(*initial-value(s)*)

To specify an iteration factor and nested sublists for the initial values, use the following format:

<*constant-iter-value**> <(>*constant value* | *constant-sublist*<)>

Restriction: If you specify both an *initial-value-list* and an *element-list*, then *element-list* must be listed before *initial-value-list* in the RETAIN statement.

Tip: You can separate initial values by blank spaces or commas.

Tip: You can also use a shorthand notation for specifying a range of sequential integers. The increment is always +1.

Tip: You can assign initial values to both variables and temporary data elements.

Tip: If there are more variables than initial values, the remaining variables are assigned an initial value of missing and SAS issues a warning message.

Details

Default DATA Step Behavior Without a RETAIN statement, SAS automatically sets variables that are assigned values by an INPUT or assignment statement to missing before each iteration of the DATA step.

Assigning Initial Values Use a RETAIN statement to specify initial values for individual variables, a list of variables, or members of an array. If a value appears in a RETAIN statement, variables that appear before it in the list are set to that value initially. (If you assign different initial values to the same variable by naming it more than once in a RETAIN statement, SAS uses the last value.) You can also use RETAIN to assign an initial value other than the default value of 0 to a variable whose value is assigned by a sum statement.

Redundancy It is redundant to name any of these items in a RETAIN statement, because their values are automatically retained from one iteration of the DATA step to the next:

- variables that are read with a SET, MERGE, MODIFY or UPDATE statement
- a variable whose value is assigned in a sum statement
- the automatic variables _N_, _ERROR_, _I_, _CMD_, and _MSG_
- variables that are created by the END= or IN= option in the SET, MERGE, MODIFY, or UPDATE statement or by options that create variables in the FILE and INFILE statements
- data elements that are specified in a temporary array
- array elements that are initialized in the ARRAY statement
- elements of an array that have assigned initial values to any or all of the elements in the ARRAY statement.

You can, however, use a RETAIN statement to assign an initial value to any of the previous items, with the exception of _N_ and _ERROR_.

Comparisons

The RETAIN statement specifies variables whose values are *not set to missing* at the beginning of each iteration of the DATA step. The KEEP statement specifies variables that are to be included in any data set that is being created.

Examples

Example 1: Basic Usage

- This RETAIN statement retains the values of variables MONTH1 through MONTH5 from one iteration of the DATA step to the next:

```
retain month1-month5;
```

- This RETAIN statement retains the values of nine variables and sets their initial values:

```
retain month1-month5 1 year 0 a b c 'XYZ';
```

  The values of MONTH1 through MONTH5 are set initially to 1; YEAR is set to 0; variables A, B, and C are each set to the character value **XYZ**.

- This RETAIN statement assigns the initial value 1 to the variable MONTH1 only:

```
retain month1-month5 (1);
```

Variables MONTH2 through MONTH5 are set to missing initially.

- This RETAIN statement retains the values of all variables that are defined earlier in the DATA step but not the values that are defined *afterwards*:

```
retain _all_;
```

- All of these statements assign initial values of 1 through 4 to VAR1 through VAR4:
  - `retain var1-var4 (1 2 3 4);`
  - `retain var1-var4 (1,2,3,4);`
  - `retain var1-var4(1:4);`

Example 2: Overview of the RETAIN Operation This example shows how to use variable names and array names as elements in the RETAIN statement and shows assignment of initial values with and without parentheses:

```
data _null_;
   array City{3} $ City1-City3;
   array cp{3} Citypop1-Citypop3;
   retain Year Taxyear 1999 City ' '
          cp (10000,50000,100000);
   file file-specification print;
   put 'Values at beginning of DATA step:'
       / @3 _all_ /;
   input Gain;
   do i=1 to 3;
      cp{i}=cp{i}+Gain;
   end;
   put 'Values after adding Gain to city populations:'
       / @3 _all_;
   datalines;
5000
10000
;
```

Here are the initial values assigned by RETAIN:

- Year and Taxyear are assigned the initial value 1999.
- City1, City2, and City3 are assigned missing values.
- Citypop1 is assigned the value 10000.
- Citypop2 is assigned 50000.
- Citypop3 is assigned 100000.

Here are the lines written by the PUT statements:

```
Values at beginning of DATA step:
  City1=  City2=  City3=  Citypop1=10000
  Citypop2=50000 Citypop3=100000
Year=1999 Taxyear=1999 Gain=. i=.
_ERROR_=0 _N_=1

Values after adding GAIN to city populations:
  City1=  City2=  City3=  Citypop1=15000
  Citypop2=55000 Citypop3=105000
Year=1999 Taxyear=1999 Gain=5000 i=4
_ERROR_=0 _N_=1
Values at beginning of DATA step:
```

```
  City1=  City2=  City3=  Citypop1=15000
  Citypop2=55000 Citypop3=105000
Year=1999 Taxyear=1999 Gain=. i=.
_ERROR_=0 _N_=2

Values after adding GAIN to city populations:
  City1=  City2=  City3=  Citypop1=25000
  Citypop2=65000 Citypop3=115000
Year=1999 Taxyear=1999 Gain=10000 i=4
_ERROR_=0 _N_=2
Values at beginning of DATA step:
  City1=  City2=  City3=  Citypop1=25000
  Citypop2=65000 Citypop3=115000
Year=1999 Taxyear=1999 Gain=. i=.
_ERROR_=0 _N_=3
```

The first PUT statement is executed three times, while the second PUT statement is executed only twice. The DATA step ceases execution when the INPUT statement executes for the third time and reaches the end of the file.

Example 3: Selecting One Value from a Series of Observations In this example, the data set ALLSCORES contains several observations for each identification number and variable ID. Different observations for a particular ID value might have different values of the variable GRADE. This example creates a new data set, CLASS.BESTSCORES, which contains one observation for each ID value. The observation must have the highest GRADE value of all observations for that ID in BESTSCORES.

```
libname class 'SAS-library';

proc sort data=class.allscores;
   by id;
run;

data class.bestscores;
   drop grade;
   set class.allscores;
   by id;
      /* Prevents HIGHEST from being reset*/
      /* to missing for each iteration.   */
   retain highest;
      /* Sets HIGHEST to missing for each */
      /* different ID value.              */
   if first.id then highest=.;
      /* Compares HIGHEST to GRADE in     */
      /* current iteration and resets     */
      /* value if GRADE is higher.        */
   highest=max(highest,grade);
   if last.id then output;
run;
```

See Also

Statements:

"Assignment Statement" on page 1447

"BY Statement" on page 1452
"INPUT Statement" on page 1617

RETURN Statement

Stops executing statements at the current point in the DATA step and returns to a predetermined point in the step.

Valid: in a DATA step
Category: Control
Type: Executable

Syntax

RETURN;

Without Arguments

The RETURN statement causes execution to stop at the current point in the DATA step, and returns control to a previous DATA step statement.

Details

The point to which SAS returns depends on the order in which statements are executed in the DATA step.

The RETURN statement is often used with the

- GO TO statement
- HEADER= option in the FILE statement
- LINK statement.

When RETURN causes a return to the beginning of the DATA step, an implicit OUTPUT statement writes the current observation to any new data sets (unless the DATA step contains an explicit OUTPUT statement, or REMOVE or REPLACE statements with MODIFY statements). Every DATA step has an implied RETURN as its last executable statement.

Examples

In this example, when the values of X and Y are the same, SAS executes the RETURN statement and adds the observation to the data set. When the values of X and Y are not equal, SAS executes the remaining statements and then adds the observation to the data set.

```
data survey;
   input x y;
   if x=y then return;
   put x= y=;
   datalines;
21 25
20 20
```

```
7 17
;
```

See Also

Statements:

"FILE Statement" on page 1503

"GO TO Statement" on page 1579

"LINK Statement" on page 1669

RUN Statement

Executes the previously entered SAS statements.

Valid: anywhere

Category: Program Control

Syntax

RUN <CANCEL>;

Without Arguments

Without arguments, the RUN statement executes the previously entered SAS statements.

Arguments

CANCEL
terminates the current step without executing it. SAS prints a message that indicates that the step was not executed.

CAUTION:
The CANCEL option does not prevent execution of a DATA step that contains a DATALINES or DATALINES4 statement. △

CAUTION:
The CANCEL option has no effect when you use the KILL option with PROC DATASETS. △

Details

Although the RUN statement is not required between steps in a SAS program, using it creates a step boundary and can make the SAS log easier to read.

Examples

- This RUN statement marks a step boundary and executes this PROC PRINT step:

```
proc print data=report;
   title 'Status Report';
run;
```

- This example shows the usefulness of the CANCEL option in a line prompt mode session. The fourth statement in the DATA step contains an invalid value for PI (4.13 instead of 3.14). RUN with CANCEL ends the DATA step and prevents it from executing.

```
data circle;
   infile file-specification;
   input radius;
   c=2*4.13*radius;
run cancel;
```

The following message is written to the SAS log:

```
WARNING: DATA step not executed at user's request.
```

%RUN Statement

Ends source statements following a %INCLUDE * statement.

Valid: anywhere

Category: Program Control

Syntax

%RUN;

Without Arguments

The %RUN statement causes SAS to stop reading input from the keyboard (including subsequent SAS statements on the same line as %RUN) and resume reading from the previous input source.

Details

Using the %INCLUDE statement with an asterisk specifies that you enter source lines from the keyboard.

Note: The asterisk (*) cannot be used to specify keyboard entry if you use the Enhanced Editor in the Microsoft Windows operating environment. △

Comparisons

The RUN statement executes previously entered DATA or PROC steps. The %RUN statement ends the prompting for source statements and returns program control to the original source program, when you use the %INCLUDE statement to allow data to be entered from the keyboard.

The type of prompt that you use depends on how you run the SAS session. The include operation is most useful in interactive line and noninteractive modes, but it can

also be used in windowing and batch mode. When you are running SAS in batch mode, include the %RUN statement in the external file that is referenced by the SASTERM fileref.

Examples

- To request keyboard-entry source on a %INCLUDE statement, follow the statement with an asterisk:

```
%include *;
```

Note: The asterisk (*) cannot be used to specify keyboard entry if you use the Enhanced Editor in the Microsoft Windows operating environment. △

- When it executes this statement, SAS prompts you to enter source lines from the keyboard. When you finish entering code from the keyboard, type the following statement to return processing to the program that contains the %INCLUDE statement.

```
%run;
```

See Also

Statements:

"%INCLUDE Statement" on page 1584
"RUN Statement" on page 1753

SASFILE Statement

Opens a SAS data set and allocates enough buffers to hold the entire file in memory.

Valid: Anywhere

Category: Program Control

Restriction: A SAS data set opened by the SASFILE statement can be used for subsequent input (read) or update processing but not for output or utility processing.

See: SASFILE Statement in the documentation for your operating environment.

Syntax

SASFILE <*libref.*>*member-name*<*.member-type*> <(*password-option(s)*)> OPEN | LOAD | CLOSE ;

Arguments

libref

a name that is associated with a SAS library. The libref (library reference) must be a valid SAS name. The default libref is either USER (if assigned) or WORK (if USER not assigned).

Restriction: The libref cannot represent a concatenation of SAS libraries that contain a library in sequential format.

member-name
a valid SAS name that is a SAS data file (a SAS data set with the member type DATA) that is a member of the SAS library associated with the libref.

Restriction: The SAS data set must have been created with the V7, V8, or V9 Base SAS engine.

member-type
the type of SAS file to be opened. Valid value is DATA, which is the default.

password-option(s)
specifies one or more of the following password options:

READ=*password*
enables the SASFILE statement to open a read-protected file. The *password* must be a valid SAS name.

WRITE=*password*
enables the SASFILE statement to use the write password to open a file that is both read-protected and write-protected. The *password* must be a valid SAS name.

ALTER=*password*
enables the SASFILE statement to use the alter password to open a file that is both read-protected and alter-protected. The *password* must be a valid SAS name.

PW=*password*
enables the SASFILE statement to use the password to open a file that is assigned for all levels of protection. The *password* must be a valid SAS name.

Tip: When SASFILE is executed, SAS checks whether the file is read-protected. Therefore, if the file is read-protected, you must include the READ= password in the SASFILE statement. If the file is either write-protected or alter-protected, you can use a WRITE=, ALTER=, or PW= password. However, the file is opened only in input (read) mode. For subsequent processing, you must specify the necessary password or passwords. See Example 2 on page 1760.

OPEN
opens the file, allocates the buffers, but defers reading the data into memory until a procedure, statement, or application is executed.

LOAD
opens the file, allocates the buffers, and reads the data into memory.

Note: If the total number of allowed buffers is less than the number of buffers required for the file based on the number of data set pages and index file pages, SAS issues a warning to tell you how many pages are read into memory. △

CLOSE
frees the buffers and closes the file.

Details

General Information The SASFILE statement opens a SAS data set and allocates enough buffers to hold the entire file in memory. Once it is read, data is held in memory, available to subsequent DATA and PROC steps or applications, until either a second SASFILE statement closes the file and frees the buffers or the program ends, which automatically closes the file and frees the buffers.

Using the SASFILE statement can improve performance by

- reducing multiple open/close operations (including allocation and freeing of memory for buffers) to process a SAS data set to one open/close operation
- reducing I/O processing by holding the data in memory.

If your SAS program consists of steps that read a SAS data set multiple times and you have an adequate amount of memory so that the entire file can be held in real memory, the program should benefit from using the SASFILE statement. Also, SASFILE is especially useful as part of a program that starts a SAS server such as a SAS/SHARE server. However, as with most performance-improvement features, it is suggested that you set up a test in your environment to measure performance with and without the SASFILE statement.

Processing a SAS Data Set Opened with SASFILE When the SASFILE statement executes, SAS opens the specified file. Then when subsequent DATA and PROC steps execute, SAS does not have to open the file for each request; the file remains open until a second SASFILE statement closes it or the program or session ends.

When a SAS data set is opened by the SASFILE statement, the file is opened for input processing and can be used for subsequent input or update processing. However, the file cannot be used for subsequent utility or output processing, because utility and output processing requires exclusive access to the file (member-level locking). For example, you cannot replace the file or rename its variables.

Table 6.11 on page 1757 provides a list of some SAS procedures and statements and specifies whether they are allowed if the file is opened by the SASFILE statement:

Table 6.11 Processing Requests for a File Opened by SASFILE

Processing Request	Open Mode	Allowed
APPEND procedure	update	Yes
DATA step that creates or replaces the file	output	No
DATASETS procedure to rename or add a variable, add or change a label, or add or remove integrity constraints or indexes	utility	No
DATASETS procedure with AGE, CHANGE, or DELETE statements	does not open the file but requires exclusive access	No
FSEDIT procedure	update	Yes
PRINT procedure	input	Yes
SORT procedure that replaces original data set with sorted one	output	No
SQL procedure to modify, add, or delete observations	update	Yes

Processing Request	Open Mode	Allowed
SQL procedure with CREATE TABLE or CREATE VIEW statement	output	No
SQL procedure to create or remove integrity constraints or indexes	utility	No

Buffer Allocation A buffer is a reserved area of memory that holds a segment of data while it is processed. The number of allocated buffers determines how much data can be held in memory at one time.

The number of buffers is not a permanent attribute of a SAS file. That is, it is valid only for the current SAS session or job. When a SAS file is opened, a default number of buffers for processing the file is set. The default depends on the operating environment but typically is a small number such as one buffer. To specify a different number of buffers, you can use the BUFNO= data set option or system option.

When the SASFILE statement is executed, SAS automatically allocates the number of buffers based on the number of data set pages and index file pages (if an index file exists). For example:

- If the number of data set pages is five and there is not an index file, SAS allocates five buffers.
- If the number of data set pages is 500 and the number of index file pages is 200, SAS allocates 700 buffers.

If a file that is held in memory increases in size during processing, the number of allocated buffers increases to accommodate the file. Note that if SASFILE is executed for a SAS data set, the BUFNO= option is ignored.

I/O Processing An I/O (input/output) request reads a segment of data from a storage device (such as disk) and transfers the data to memory, or conversely transfers the data from memory and writes it to the storage device. When a SAS data set is opened by the SASFILE statement, data is read once and held in memory, which should reduce the number of I/O requests.

CAUTION:

I/O processing can be reduced only if there is sufficient *real memory*. If the SAS data set is very large, you might not have sufficient real memory to hold the entire file. If insufficient memory exists, your operating environment can simulate more memory than actually exists, which is virtual memory. If virtual memory occurs, data access I/O requests are replaced with swapping I/O requests, which could result in no performance improvement. In addition, both SAS and your operating environment have a maximum amount of memory that can be allocated, which could be exceeded by the needs of your program. If your program needs exceed the memory that is available, the number of allocated buffers might be decreased to the default allocation in order to free memory. △

Tip: To determine how much memory a SAS data set requires, execute the CONTENTS procedure for the file to list its page size, the number of data set pages, the index file size, and the number of index file pages.

Using the SASFILE Statement in a SAS/SHARE Environment The following are considerations for using the SASFILE statement with SAS/SHARE software:

- You must execute the SASFILE statement before you execute the PROC SERVER statement.
- If the client (the computer on which you use a SAS session to access a SAS/SHARE server) executes the SASFILE statement, it is rejected.
- Once the SASFILE statement is executed, all users who subsequently open the file will access the data held in memory instead of data that is stored on the disk.
- Once the SASFILE statement is executed, you cannot close the file and free the buffers until the SAS/SHARE server is terminated.
- You can use the ALLOCATE SASFILE command for the PROC SERVER statement as an alternative that brings part of the file into memory (controlled by the BUFNO= option).
- If the SASFILE statement is executed and you execute ALLOCATE SASFILE specifying a value for BUFNO= that is a larger number of buffers than allocated by SASFILE, performance will not be improved.

Comparisons

- Use the BUFNO= system option or data set option to specify a specific number of buffers.
- With SAS/SHARE software, you can use the ALLOCATE SASFILE command for the PROC SERVER statement to bring part of the file into memory (controlled by the BUFNO= option).

Examples

Example 1: Using SASFILE in a Program with Multiple Steps The following SAS program illustrates the process of opening a SAS data set, transferring its data to memory, and reading that data held in memory for multiple tasks. The program is consists of steps that read the file multiple times.

```
libname mydata 'SAS-library';

sasfile mydata.census.data open; ❶

data test1;
   set mydata.census; ❷
run;

data test2;
   set mydata.census; ❸
run;

proc summary data=mydata.census print; ❹
run;

data mydata.census; ❺
   modify mydata.census;
   .
   . (statements to modify data)
   .
run;

sasfile mydata.census close; ❻
```

1 Opens SAS data set MYDATA.CENSUS, and allocates the number of buffers based on the number of data set pages and index file pages.
2 Reads all pages of MYDATA.CENSUS, and transfers all data from disk to memory.
3 Reads MYDATA.CENSUS a second time, but this time from memory without additional I/O requests.
4 Reads MYDATA.CENSUS a third time, again from memory without additional I/O requests.
5 Reads MYDATA.CENSUS a fourth time, again from memory without additional I/O requests. If the MODIFY statement successfully changes data in memory, the changed data is transferred from memory to disk at the end of the DATA step.
6 Closes MYDATA.CENSUS, and frees allocated buffers.

Example 2: Specifying Passwords with the SASFILE Statement The following SAS program illustrates using the SASFILE statement and specifying passwords for a SAS data set that is both read-protected and alter-protected:

```
libname mydata 'SAS-data-data-library';

sasfile mydata.census (read=gizmo) open; ❶

proc print data=mydata.census (read=gizmo); ❷
run;

data mydata.census;
   modify mydata.census (alter=luke); ❸
   .
   . (statements to modify data)
   .
run;
```

1 The SASFILE statement specifies the read password, which is sufficient to open the file.
2 In the PRINT procedure, the read password must be specified again.
3 The alter password is used in the MODIFY statement, because the data set is being updated.

Note: It is acceptable to use the higher-level alter password instead of the read password in the above example. △

See Also

Data Set Option:
"BUFNO= Data Set Option" on page 15
System Option:
"BUFNO= System Option" on page 1851
"The SERVER Procedure" in *SAS/SHARE User's Guide*.

SELECT Statement

Executes one of several statements or groups of statements.

Valid: in a DATA step
Category: Control
Type: Executable

Syntax

SELECT <(*select-expression*)>;
  WHEN-1 (*when-expression-1* <..., *when-expression-n*>) *statement*;
  <... **WHEN**-*n* (*when-expression-1* <..., *when-expression-n*>) *statement*;>
  <**OTHERWISE** *statement*;>

END;

Arguments

(select-expression)
specifies any SAS expression that evaluates to a single value.

See: "Evaluating the *when-expression* When a *select-expression* Is Included" on page 1761

(when-expression)
specifies any SAS expression, including a compound expression. SELECT requires you to specify at least one *when-expression*.

Tip: Separating multiple *when-expressions* with a comma is equivalent to separating them with the logical operator OR.

Tip: The way a *when-expression* is used depends on whether a *select-expression* is present.

See: "Evaluating the *when-expression* When a *select-expression* Is Not Included" on page 1762

statement
can be any executable SAS statement, including DO, SELECT, and null statements. You must specify the *statement* argument.

Details

Using WHEN Statements in a SELECT Group The SELECT statement begins a SELECT group. SELECT groups contain WHEN statements that identify SAS statements that are executed when a particular condition is true. Use at least one WHEN statement in a SELECT group. An optional OTHERWISE statement specifies a statement to be executed if no WHEN condition is met. An END statement ends a SELECT group.

Null statements that are used in WHEN statements cause SAS to recognize a condition as true without taking further action. Null statements that are used in OTHERWISE statements prevent SAS from issuing an error message when all WHEN conditions are false.

Evaluating the *when-expression* When a *select-expression* Is Included If the *select-expression* is present, SAS evaluates the *select-expression* and *when-expression*. SAS compares the two for equality and returns a value of true or false. If the comparison is true, *statement* is executed. If the comparison is false, execution proceeds

either to the next *when-expression* in the current WHEN statement, or to the next WHEN statement if no more expressions are present. If no WHEN statements remain, execution proceeds to the OTHERWISE statement, if one is present. If the result of all SELECT-WHEN comparisons is false and no OTHERWISE statement is present, SAS issues an error message and stops executing the DATA step.

Evaluating the *when-expression* When a *select-expression* Is Not Included If no *select-expression* is present, the *when-expression* is evaluated to produce a result of true or false. If the result is true, *statement* is executed. If the result is false, SAS proceeds to the next *when-expression* in the current WHEN statement, or to the next WHEN statement if no more expressions are present, or to the OTHERWISE statement if one is present. (That is, SAS performs the action that is indicated in the first true WHEN statement.) If the result of all *when-expressions* is false and no OTHERWISE statement is present, SAS issues an error message. If more than one WHEN statement has a true *when-expression*, only the first WHEN statement is used. Once a *when-expression* is true, no other *when-expressions* are evaluated.

Processing Large Amounts of Data with %INCLUDE Files One way to process large amounts of data is to use %INCLUDE statements in your DATA step. Using %INCLUDE statements enables you to perform complex processing while keeping your main program manageable. The %INCLUDE files that you use in your main program can contain WHEN statements and other SAS statements to process your data. See Example 5 on page 1763 for an example.

Comparisons

Use IF-THEN/ELSE statements for programs with few statements. Use subsetting IF statements without a THEN clause to continue processing only those observations or records that meet the condition that is specified in the IF clause.

Examples

Example 1: Using Statements

```
select (a);
   when (1) x=x*10;
   when (2);
   when (3,4,5) x=x*100;
   otherwise;
end;
```

Example 2: Using DO Groups

```
select (payclass);
   when ('monthly') amt=salary;
   when ('hourly')
      do;
         amt=hrlywage*min(hrs,40);
         if hrs>40 then put 'CHECK TIMECARD';
      end;          /* end of do     */
   otherwise put 'PROBLEM OBSERVATION';
end;               /* end of select */
```

Example 3: Using a Compound Expression

```
select;
   when (mon in ('JUN', 'JUL', 'AUG')
   and temp>70) put 'SUMMER ' mon=;
   when (mon in ('MAR', 'APR', 'MAY'))
   put 'SPRING ' mon=;
   otherwise put 'FALL OR WINTER ' mon=;
end;
```

Example 4: Making Comparisons for Equality

```
   /* INCORRECT usage to select value of 2 */
select (x);
   /* evaluates T/F and compares for       */
   /* equality with x                      */
   when (x=2) put 'two';
end;

   /* correct usage */
select(x);
   /* compares 2 to x for equality */
   when (2) put 'two';
end;

   /* correct usage */
select;
   /* compares 2 to x for equality         */
   when (x=2) put 'two';
end;
```

Example 5: Processing Large Amounts of Data In the following example, the %INCLUDE statements contain code that includes WHEN statements to process new and old items in the inventory. The main program shows the overall logic of the DATA step.

```
data test (keep=ItemNumber);
   set ItemList;
   select;
      %include NewItems;
      %include OldItems;
      otherwise put 'Item ' ItemNumber ' is not in the inventory.';
   end;
run;
```

See Also

Statements:

"DO Statement" on page 1491

"IF Statement, Subsetting" on page 1581

"IF-THEN/ELSE Statement" on page 1582

SET Statement

Reads an observation from one or more SAS data sets.

Valid: in a DATA step
Category: File-handling
Type: Executable

Syntax

SET<*SAS-data-set(s)* <(*data-set-options(s)*)>>
<*options*>;

Without Arguments

When you do not specify an argument, the SET statement reads an observation from the most recently created data set.

Arguments

SAS-data-set (s)
specifies a one-level name, a two-level name, or one of the special SAS data set names.

Tip: You can specify data set lists. For more information, see "Using Data Set Lists with SET" on page 1768.

See Also: See "SAS Data Sets" in *SAS Language Reference: Concepts* for a description of the levels of SAS data set names and when to use each level.

Featured in: Example 13 on page 1772

(*data-set-options*)
specifies actions SAS is to take when it reads variables or observations into the program data vector for processing.

Tip: Data set options that apply to a data set list apply to all of the data sets in the list.

See: Refer to "Definition of Data Set Options" on page 10 for a list of the data set options to use with input data sets.

Options

END=*variable*
creates and names a temporary variable that contains an end-of-file indicator. The variable, which is initialized to zero, is set to 1 when SET reads the last observation of the last data set listed. This variable is not added to any new data set.

Restriction: END= cannot be used with POINT=. When random access is used, the END= variable is never set to 1.

Interaction: If you use a BY statement, END= is set to 1 when the SET statement reads the last observation of the interleaved data set. For more information, see "BY-Group Processing with SET" on page 1769.

Featured in: Example 11 on page 1772

KEY=*index*</UNIQUE>
provides nonsequential access to observations in a SAS data set, which are based on the value of an index variable or a key.

Range: Specify the name of a simple or a composite index of the data set that is being read.

Restriction: KEY= cannot be used with POINT=.

Tip: Using the _IORC_ automatic variable in conjunction with the SYSRC autocall macro provides you with more error-handling information than was previously available. When you use the SET statement with the KEY= option, the new automatic variable _IORC_ is created. This automatic variable is set to a return code that shows the status of the most recent I/O operation that is performed on an observation in a SAS data set. If the KEY= value is not found, the _IORC_ variable returns a value that corresponds to the SYSRC autocall macro's mnemonic _DSENOM and the automatic variable _ERROR_ is set to 1.

Featured in: Example 7 on page 1771 and Example 8 on page 1771.

See Also: For more information, see the description of the autocall macro SYSRC in *SAS Macro Language: Reference*.

See Also: UNIQUE option on page 1767

CAUTION:
Continuous loops can occur when you use the KEY= option. If you use the KEY= option without specifying the primary data set, you must include either a STOP statement to stop DATA step processing, or programming logic that uses the _IORC_ automatic variable in conjunction with the SYSRC autocall macro and checks for an invalid value of the _IORC_ variable, or both. △

INDSNAME=*variable*
creates and names a variable that stores the name of the SAS data set from which the current observation is read. The stored name can be a data set name or a physical name. The physical name is the name by which the operating environment recognizes the file.

Tip: For data set names, SAS will add the library name to the variable value (for example, WORK.PRICE) and convert the two-level name to uppercase.

Tip: Unless previously defined, the length of the variable is set to 41 characters. Use a LENGTH statement to make the variable length long enough to contain the value of the physical filename if it is longer than 41 characters.

If the variable is previously defined as a character variable with a specific length, that length is not changed. If the value placed into the INDSNAME variable is longer than that length, then the value is truncated.

If the variable is previously defined as a numeric variable, an error will occur.

Featured in: Example 12 on page 1772

NOBS=*variable*
creates and names a temporary variable whose value is usually the total number of observations in the input data set or data sets. If more than one data set is listed in the SET statement, NOBS= the total number of observations in the data sets that are listed. The number of observations includes those observations that are marked for deletion but are not yet deleted.

Restriction: For certain SAS views, SAS cannot determine the number of observations. In these cases, SAS sets the value of the NOBS= variable to the largest positive integer value that is available in your operating environment.

Tip: At compilation time, SAS reads the descriptor portion of each data set and assigns the value of the NOBS= variable automatically. Thus, you can refer to

the NOBS= variable before the SET statement. The variable is available in the DATA step but is not added to any output data set.

Interaction: The NOBS= and POINT= options are independent of each other.

Featured in: Example 10 on page 1771

OPEN=(IMMEDIATE | DEFER)
allows you to delay the opening of any concatenated SAS data sets until they are ready to be processed.

IMMEDIATE
during the compilation phase, opens all data sets that are listed in the SET statement.

Restriction: When you use the IMMEDIATE option KEY=, POINT=, and BY statement processing are mutually exclusive.

Tip: If a variable on a subsequent data set is of a different type (character versus numeric, for example) than the type of the same-named variable on the first data set, the DATA step will stop processing and produce an error message.

DEFER
opens the first data set during the compilation phase, and opens subsequent data sets during the execution phase. When the DATA step reads and processes all observations in a data set, it closes the data set and opens the next data set in the list.

Restriction: When you specify the DEFER option, you cannot use the KEY= statement option, the POINT= statement option, or the BY statement. These constructs imply either random processing or interleaving of observations from the data sets, which is not possible unless all data sets are open.

Requirement: You can use the DROP=, KEEP=, or RENAME= data set options to process a set of variables, but the set of variables that are processed for each data set must be identical. In most cases, if the set of variables defined by any subsequent data set differs from the variables defined by the first data set, SAS prints a warning message to the log but does not stop execution. SAS stops execution for some conditions:

1. If a variable on a subsequent data set is of a different type (character versus numeric, for example) than the type of the same-named variable on the first data set, the DATA step will stop processing and produce an error message.
2. If a variable on a subsequent data set was not defined by the first data set in the SET statement, but was defined previously in the DATA step program, the DATA step will stop processing and produce an error message. In this case, the value of the variable in previous iterations might be incorrect because the semantic behavior of SET requires this variable to be set to missing when processing the first observation of the first data set.

Default: IMMEDIATE

POINT=*variable*
specifies a temporary variable whose numeric value determines which observation is read. POINT= causes the SET statement to use random (direct) access to read a SAS data set.

Requirement: a STOP statement

Restriction: You cannot use POINT= with a BY statement, a WHERE statement, or a WHERE= data set option. In addition, you cannot use it with transport

format data sets, data sets in sequential format on tape or disk, and SAS/ACCESS views or the SQL procedure views that read data from external files.

Restriction: You cannot use POINT= with KEY=.

Tip: You must supply the values of the POINT= variable. For example, you can use the POINT= variable as the index variable in some form of the DO statement.

Tip: The POINT= variable is available anywhere in the DATA step, but it is not added to any new SAS data set.

Featured in: Example 6 on page 1771 and Example 9 on page 1771

CAUTION:

Continuous loops can occur when you use the POINT= option. When you use the POINT= option, you must include a STOP statement to stop DATA step processing, programming logic that checks for an invalid value of the POINT= variable, or both. Because POINT= reads only those observations that are specified in the DO statement, SAS cannot read an end-of-file indicator as it would if the file were being read sequentially. Because reading an end-of-file indicator ends a DATA step automatically, failure to substitute another means of ending the DATA step when you use POINT= can cause the DATA step to go into a continuous loop. If SAS reads an invalid value of the POINT= variable, it sets the automatic variable _ERROR_ to 1. Use this information to check for conditions that cause continuous DO-loop processing, or include a STOP statement at the end of the DATA step, or both. △

UNIQUE

causes a KEY= search always to begin at the top of the index for the data set that is being read.

Restriction: UNIQUE can appear only with the KEY= argument and must be preceded by a slash.

Explanation: By default, SET begins searching at the top of the index only when the KEY= value changes.

If the KEY= value does not change on successive executions of the SET statement, the search begins by following the most recently retrieved observation. In other words, when consecutive duplicate KEY= values appear, the SET statement attempts a one-to-one match with duplicate indexed values in the data set that is being read. If more consecutive duplicate KEY= values are specified than exist in the data set that is being read, the extra duplicates are treated as not found.

When KEY= is a unique value, only the first attempt to read an observation with that key value succeeds; subsequent attempts to read the observation with that value of the key will fail. The _IORC_ variable returns a value that corresponds to the SYSRC autocall macro's mnemonic _DSENOM. If you add the /UNIQUE option, subsequent attempts to read the observation with the unique KEY= value will succeed. The _IORC_ variable returns a 0.

Featured in: Example 8 on page 1771

See Also: For extensive examples, see *Combining and Modifying SAS Data Sets: Examples*.

Details

What SET Does Each time the SET statement is executed, SAS reads one observation into the program data vector. SET reads all variables and all observations from the

input data sets unless you tell SAS to do otherwise. A SET statement can contain multiple data sets; a DATA step can contain multiple SET statements. See *Combining and Modifying SAS Data Sets: Examples*.

Uses The SET statement is flexible and has a variety of uses in SAS programming. These uses are determined by the options and statements that you use with the SET statement:

- reading observations and variables from existing SAS data sets for further processing in the DATA step
- concatenating and interleaving data sets, and performing one-to-one reading of data sets
- reading SAS data sets by using direct access methods.

Using Data Set Lists with SET You can use data set lists with the SET statement. Data set lists provide a quick way to reference existing groups of data sets. These data set lists must be either name prefix lists or numbered range lists.

Name prefix lists refer to all data sets that begin with a specified character string. For example, **`set SALES1:;`** tells SAS to read all data sets starting with "SALES1" such as SALES1, SALES10, SALES11, and SALES12.

Numbered range lists require you to have a series of data sets with the same name, except for the last character or characters, which are consecutive numbers. In a numbered range list, you can begin with any number and end with any number. For example, these lists refer to the same data sets:

```
sales1 sales2 sales3 sales4


sales1-sales4
```

Note: If the numeric suffix of the first data set name contains leading zeros, the number of digits in the numeric suffix of the last data set name must be greater than or equal to the number of digits in the first data set name. Otherwise, an error will occur. For example, the data set lists sales001–sales99 and sales01–sales9 will cause an error. The data set list sales001–sales999 is valid. If the numeric suffix of the first data set name does not contain leading zeros, the number of digits in the numeric suffix of the first and last data set names do not have to be equal. For example, the data set list sales1–sales999 is valid. △

Some other rules to consider when using numbered data set lists are as follows:

- You can specify groups of ranges.

```
set cost1-cost4 cost11-cost14 cost21-cost24;
```

- You can mix numbered range lists with name prefix lists.

```
set cost1-cost4 cost2: cost33-37;
```

- You can mix single data sets with data set lists.

```
set cost1 cost10-cost20 cost30;
```

- Quotation marks around data set lists are ignored.

```
/* these two lines are the same */
set sales1 - sales4;
set 'sales1'n - 'sales4'n;
```

- Spaces in data set names are invalid. If quotation marks are used, trailing blanks are ignored.

```
/* blanks in these statements will cause errors */
set sales 1 - sales 4;
```

```
set 'sales 1'n - 'sales 4'n;

/* trailing blanks in this statement will be ignored */
set 'sales1   'n - 'sales4   'n;
```

- The maximum numeric suffix is 2147483647.

```
/* this suffix will cause an error */
set prod2000000000-prod2934850239;
```

- Physical pathnames are not allowed.

```
/* physical pathnames will cause an error */
&let work_path = %sysfunc(pathname(WORK));
set "&work_path\dept.sas7bdat";
```

BY-Group Processing with SET Only one BY statement can accompany each SET statement in a DATA step. The BY statement should immediately follow the SET statement to which it applies. The data sets that are listed in the SET statement must be sorted by the values of the variables that are listed in the BY statement, or they must have an appropriate index. SET, when it is used with a BY statement, interleaves data sets. The observations in the new data set are arranged by the values of the BY variable or variables, and within each BY group, by the order of the data sets in which they occur. See Example 2 on page 1770 for an example of BY-group processing with the SET statement.

Combining SAS Data Sets Use a single SET statement with multiple data sets to concatenate the specified data sets. That is, the number of observations in the new data set is the sum of the number of observations in the original data sets, and the order of the observations is all the observations from the first data set followed by all the observations from the second data set, and so on. See Example 1 on page 1770 for an example of concatenating data sets.

Use a single SET statement with a BY statement to interleave the specified data sets. The observations in the new data set are arranged by the values of the BY variable or variables, and within each BY group, by the order of the data sets in which they occur. See Example 2 on page 1770 for an example of interleaving data sets.

Use multiple SET statements to perform one-to-one reading (also called one-to-one matching) of the specified data sets. The new data set contains all the variables from all the input data sets. The number of observations in the new data set is the number of observations in the smallest original data set. If the data sets contain common variables, the values that are read in from the last data set replace the values that were read in from earlier ones. See Example 6 on page 1771, Example 7 on page 1771, and Example 8 on page 1771 for examples of one-to-one reading of data sets.

For extensive examples, see *Combining and Modifying SAS Data Sets: Examples*.

For more information about how to prepare your data sets, see "Combining SAS Data Sets: Basic Concepts" in *SAS Language Reference: Concepts*.

Comparisons

- SET reads an observation from an existing SAS data set. INPUT reads raw data from an external file or from in-stream data lines in order to create SAS variables and observations.
- Using the KEY= option with SET enables you to access observations nonsequentially in a SAS data set according to a value. Using the POINT= option with SET enables you to access observations nonsequentially in a SAS data set according to the observation number.

Examples

Example 1: Concatenating SAS Data Sets If more than one data set name appears in the SET statement, the resulting output data set is a concatenation of all the data sets that are listed. SAS reads all observations from the first data set, then all from the second data set, and so on, until all observations from all the data sets have been read. This example concatenates the three SAS data sets into one output data set named FITNESS:

```
data fitness;
   set health exercise well;
run;
```

Example 2: Interleaving SAS Data Sets To interleave two or more SAS data sets, use a BY statement after the SET statement:

```
data april;
   set payable recvable;
   by account;
run;
```

Example 3: Reading a SAS Data Set In this DATA step, each observation in the data set NC.MEMBERS is read into the program data vector. Only those observations whose value of CITY is **Raleigh** are output to the new data set RALEIGH.MEMBERS:

```
data raleigh.members;
   set nc.members;
   if city='Raleigh';
run;
```

Example 4: Merging a Single Observation with All Observations in a SAS Data Set An observation to be merged into an existing data set can be one that is created by a SAS procedure or another DATA step. In this example, the data set AVGSALES has only one observation:

```
data national;
      if _n_=1 then set avgsales;
      set totsales;
run;
```

Example 5: Reading from the Same Data Set More Than Once In this example, SAS treats each SET statement independently. That is, it reads from one data set as if it were reading from two separate data sets:

```
data drugxyz;
   set trial5(keep=sample);
   if sample>2;
   set trial5;
run;
```

For each iteration of the DATA step, the first SET statement reads one observation. The next time the first SET statement is executed, it reads the next observation. Each SET statement can read different observations with the same iteration of the DATA step.

Example 6: Combining One Observation with Many You can subset observations from one data set and combine them with observations from another data set by using direct access methods, as follows:

```
data south;
   set revenue;
   if region=4;
   set expense point=_n_;
run;
```

Example 7: Performing a Table Lookup This example illustrates using the KEY= option to perform a table lookup. The DATA step reads a primary data set that is named INVTORY and a lookup data set that is named PARTCODE. It uses the index PARTNO to read PARTCODE nonsequentially, by looking for a match between the PARTNO value in each data set. The purpose is to obtain the appropriate description, which is available only in the variable DESC in the lookup data set, for each part that is listed in the primary data set:

```
data combine;
   set invtory(keep=partno instock price);
   set partcode(keep=partno desc)  key=partno;
run;
```

Example 8: Performing a Table Lookup When the Master File Contains Duplicate Observations This example uses the KEY= option to perform a table lookup. The DATA step reads a primary data set that is named INVTORY, which is indexed on PARTNO, and a lookup data set named PARTCODE. PARTCODE contains quantities of new stock (variable NEW_STK). The UNIQUE option ensures that, if there are any duplicate observations in INVTORY, values of NEW_STK are added only to the first observation of the group:

```
data combine;
   set partcode(keep=partno new_stk);
   set invtory(keep=partno instock price)
   key=partno/unique;
   instock=instock+new_stk;
run;
```

Example 9: Reading a Subset by Using Direct Access These statements select a subset of 50 observations from the data set DRUGTEST by using the POINT= option to access observations directly by number:

```
data sample;
   do obsnum=1 to 100 by 2;
      set drugtest point=obsnum;
      if _error_ then abort;
      output;
   end;
   stop;
run;
```

Example 10: Performing a Function Until the Last Observation Is Reached These statements use NOBS= to set the termination value for DO-loop processing. The value of the temporary variable LAST is the sum of the observations in SURVEY1 and SURVEY2:

```
do obsnum=1 to last by 100;
   set survey1 survey2 point=obsnum nobs=last;
```

```
   output;
end;
stop;
```

Example 11: Writing an Observation Only After All Observations Have Been Read This example uses the END= variable LAST to tell SAS to assign a value to the variable REVENUE and write an observation only after the last observation of RENTAL has been read:

```
set rental end=last;
totdays + days;
if last then
   do;
      revenue=totdays*65.78;
      output;
   end;
```

Example 12: Retrieving the Name of the Data Set from Which the Current Observation Is Read This example creates three data sets and stores the data set name in a variable named *dsn*. The name is split into three parts and the example prints out the results.

```
/* Create some data sets to read */
data gas_price_option; value=395; run;
data gas_rbid_option; value=840; run;
data gas_price_forward; value=275; run;
/* Create a data set D */
data d;
   set gas_price_option gas_rbid_option gas_price_forward indsname=dsn;
   /* split the data set names into 3 parts */
   commodity = scan (dsn, 2, "._");
   type = scan (dsn, 3, "._");
   instrument = scan (dsn, 4, "._");
   run;
proc print data=d;
run;
```

Output 6.30 Data Set Name Split into Three Parts

```
                          The SAS System                                1

              Obs    value    commodity    type     instrument

               1      395        GAS       PRICE     OPTION
               2      840        GAS       RBID      OPTION
               3      275        GAS       PRICE     FORWARD
```

Example 13: Using Data Set Lists This example uses a numbered range list to input the data sets.

```
data dept008; emp=13; run;
data dept009; emp=9; run;
data dept010; emp=4; run;
data dept011; emp=33; run;
```

```
data _null_;
  set dept008-dept010;
  put _all_;
run;
```

The following lines are written to the SAS log.

Output 6.31 Using a Data Set List with the SET Statement

```
1    data dept008; emp=13; run;
NOTE: The data set WORK.DEPT008 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.06 seconds
      cpu time            0.03 seconds

2    data dept009; emp=9; run;
NOTE: The data set WORK.DEPT009 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

3    data dept010; emp=4; run;
NOTE: The data set WORK.DEPT010 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

4    data dept011; emp=33; run;
NOTE: The data set WORK.DEPT011 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

5
6   data _null_;
7     set dept008-dept010;
8     put _all_;
9   run;
emp=13 _ERROR_=0 _N_=1
emp=9 _ERROR_=0 _N_=2
emp=4 _ERROR_=0 _N_=3
NOTE: There were 1 observations read from the data set WORK.DEPT008.
NOTE: There were 1 observations read from the data set WORK.DEPT009.
NOTE: There were 1 observations read from the data set WORK.DEPT010.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds
```

In addition, you could use data set lists to find missing data sets. This example uses a numbered range list to locate the missing data sets. An error occurs for each data set that does not exist. Once you know which data sets are missing, you can correct the SET statement to reflect the data sets that actually exist.

```
data dept008; emp=13; run;
data dept009; emp=9; run;
data dept011; emp=4; run;
data dept014; emp=33; run;

data _null_;
  set dept008-dept014;
  put _all_;
run;
```

The following lines are written to the SAS log.

Output 6.32 Finding Missing Data Sets Using the SET Statement

```
1    data dept008; emp=13; run;
NOTE: The data set WORK.DEPT008 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.04 seconds

2    data dept009; emp=9; run;
NOTE: The data set WORK.DEPT009 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

3    data dept011; emp=4; run;
NOTE: The data set WORK.DEPT011 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.01 seconds

4    data dept014; emp=33; run;
NOTE: The data set WORK.DEPT014 has 1 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

5    data _null_;
6    set dept008-dept014;
ERROR: File WORK.DEPT010.DATA does not exist.
ERROR: File WORK.DEPT012.DATA does not exist.
ERROR: File WORK.DEPT013.DATA does not exist.
7    put _all_;
8    run;
NOTE: The SAS System stopped processing this step because of errors.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds
```

See Also

Statements:

"BY Statement" on page 1452

"DO Statement" on page 1491

"INPUT Statement" on page 1617

"MERGE Statement" on page 1679

"STOP Statement" on page 1775

"UPDATE Statement" on page 1787

"Rules for Words and Names" in *SAS Language Reference: Concepts*

"Reading, Modifying, and Combining SAS Data Sets" in *SAS Language Reference: Concepts*

"Definition of Data Set Options" on page 10

SAS Macro Language: Reference

Combining and Modifying SAS Data Sets: Examples

SKIP Statement

Creates a blank line in the SAS log.

Valid: Anywhere

Category: Log Control

Syntax

SKIP <*n*>;

Without Arguments

Using SKIP without arguments causes SAS to create one blank line in the log.

Arguments

n
specifies the number of blank lines that you want to create in the log.

Tip: If the number specified is greater than the number of lines that remain on the page, SAS goes to the top of the next page.

Details

The SKIP statement itself does not appear in the log. You can use this statement in all methods of operation.

See Also

Statement:

"PAGE Statement" on page 1707

System Options:

"LINESIZE= System Option" on page 1936

"PAGESIZE= System Option" on page 1958

STOP Statement

Stops execution of the current DATA step.

Valid: in a DATA step

Category: Action

Type: Executable

Syntax

STOP;

Without Arguments

The STOP statement causes SAS to stop processing the current DATA step immediately and resume processing statements after the end of the current DATA step.

Details

SAS outputs a data set for the current DATA step. However, the observation being processed when STOP executes is not added. The STOP statement can be used alone or in an IF-THEN statement or SELECT group.

Use STOP with any features that read SAS data sets using random access methods, such as the POINT= option in the SET statement. Because SAS does not detect an end-of-file with this access method, you must include program statements to prevent continuous processing of the DATA step.

Comparisons

- When you use a windowing environment or other interactive methods of operation, the ABORT statement and the STOP statement both stop processing. The ABORT statement sets the value of the automatic variable _ERROR_ to 1, but the STOP statement does not.
- In batch or noninteractive mode, the two statements also have different effects. Use the STOP statement in batch or noninteractive mode to continue processing with the next DATA or PROC step.

Examples

Example 1: Basic Usage

-
```
stop;
```
-
```
if idcode=9999 then stop;
```
-
```
select (a);
   when (0) output;
   otherwise stop;
end;
```

Example 2: Avoiding an Infinite Loop This example shows how to use STOP to avoid an infinite loop within a DATA step when you are using random access methods:

```
data sample;
   do sampleobs=1 to totalobs by 10;
      set master.research point=sampleobs
                          nobs=totalobs;
      output;
   end;
   stop;
run;
```

See Also

Statements:
"ABORT Statement" on page 1436
POINT= option in the SET statement on page 1766

Sum Statement

Adds the result of an expression to an accumulator variable.

Valid: in a DATA step
Category: Action
Type: Executable

Syntax

variable+expression;

Arguments

variable
specifies the name of the accumulator variable, which contains a numeric value.

Tip: The variable is automatically set to 0 before SAS reads the first observation. The variable's value is retained from one iteration to the next, as if it had appeared in a RETAIN statement.

Tip: To initialize a sum variable to a value other than 0, include it in a RETAIN statement with an initial value.

expression
is any SAS expression.

Tip: The expression is evaluated and the result added to the accumulator variable.

Tip: SAS treats an expression that produces a missing value as zero.

Comparisons

The sum statement is equivalent to using the SUM function and the RETAIN statement, as shown here:

```
retain variable 0;
variable=sum(variable,expression);
```

Examples

Here are examples of sum statements that illustrate various expressions:

```
balance+(-debit);
```

```
sumxsq+x*x;
```

```
nx+(x ne .);
```

```
if status='ready' then OK+1;
```

See Also

Function:

"SUM Function" on page 1148

Statement:

"RETAIN Statement" on page 1747

SYSECHO Statement

Fires a global statement complete event and passes a text string back to the IOM client.

Valid: anywhere

Category: Program Control

Restriction: Has an effect only in objectserver mode

Syntax

SYSECHO <"*text*">;

Without Arguments

Using SYSECHO without arguments sends a global statement complete event to the IOM client.

Arguments

"*text*"

specifies a text string that is passed back to the IOM client.

Range: 1–64 characters

Requirement: The text string must be enclosed in double quotation marks.

Details

The SYSECHO statement enables IOM clients to manually track the progress of a segment of a submitted SAS program.

When the SYSECHO statement is executed, a global statement complete event is generated and, if specified, the text string is passed back to the IOM client.

TITLE Statement

Specifies title lines for SAS output.

Valid: anywhere

Category: Output Control

See: TITLE Statement in the documentation for your operating environment.

Syntax

TITLE <*n*> <*ods-format-options*> <'*text*' | "*text*">;

Without Arguments

Using TITLE without arguments cancels all existing titles.

Arguments

n
specifies the relative line that contains the title line.

Range: 1 - 10

Tip: The title line with the highest number appears on the bottom line. If you omit *n*, SAS assumes a value of 1. Therefore, you can specify TITLE or TITLE1 for the first title line.

Tip: You can create titles that contain blank lines between the lines of text. For example, if you specify text with a TITLE statement and a TITLE3 statement, there will be a blank line between the two lines of text.

ods-format-options
specifies formatting options for the ODS HTML, RTF, and PRINTER destinations.

BOLD
specifies that the title text is bold font weight.

ODS Destinations: HTML, RTF, PRINTER

COLOR=*color*
specifies the title text color.

Alias: C

ODS Destinations: HTML, RTF, PRINTER

Featured in: Example 3 on page 1783

BCOLOR=*color*
specifies the background color of the title block.

ODS Destinations: HTML, RTF, PRINTER

FONT=*font-face*
specifies the font to use. If you supply multiple fonts, then the destination device uses the first one that is installed on your system.

Alias: F

ODS Destinations: HTML, RTF, PRINTER

HEIGHT=*size*
specifies the point size.

Alias: H

ODS Destinations: HTML, RTF, PRINTER

Featured in: Example 3 on page 1783

ITALIC
specifies that the title text is in italic style.

ODS Destinations: HTML, RTF, PRINTER

JUSTIFY= CENTER | LEFT | RIGHT
specifies justification.

CENTER
specifies center justification.

Alias: C

LEFT
specifies left justification.

Alias: L

RIGHT
specifies right justification.

Alias: R

Alias: J

ODS Destinations: HTML, RTF, PRINTER

Featured in: Example 3 on page 1783

LINK='*url*'
specifies a hyperlink.

Tip: The visual properties for LINK= always come from the current style.

ODS Destinations: HTML, RTF, PRINTER

UNDERLIN= 0 | 1 | 2 | 3
specifies whether the subsequent text is underlined. 0 indicates no underlining. 1, 2, and 3 indicates underlining.

Alias: U

Tip: ODS generates the same type of underline for values 1, 2, and 3. However, SAS/GRAPH uses values 1, 2, and 3 to generate increasingly thicker underlines.

ODS Destinations: HTML, RTF, PRINTER

Note: The defaults for how ODS renders the TITLE statement come from style elements relating to system titles in the current style. The TITLE statement syntax with *ods-format-options* is a way to override the settings provided by the current style.

The current style varies according to the ODS destination. For more information about how to determine the current style, see "What Are Style Definitions, Style Elements, and Style Attributes?" and "Concepts: Style Definitions and the TEMPLATE Procedure" in the *SAS Output Delivery System: User's Guide*. △

Tip: You can specify these options by letter, word, or words by preceding each letter or word of the *text* by the option.

For example, this code will make the title "Red, White, and Blue" appear in different colors.

```
title color=red "Red," color=white "White, and" color=blue "Blue";
```

'*text*' | "*text*"
specifies text that is enclosed in single or double quotation marks.

You can customize titles by inserting BY variable values (#BYVAL*n*), BY variable names (#BYVAR*n*), or BY lines (#BYLINE) in titles that are specified in PROC steps. Embed the items in the specified title text string at the position where you want the substitution text to appear.

#BYVAL*n* | #BYVAL(*variable-name*)
substitutes the current value of the specified BY variable for #BYVAL in the text string and displays the value in the title.

Follow these rules when you use #BYVAL in the TITLE statement of a PROC step:

- Specify the variable that is used by #BYVAL in the BY statement.
- Insert #BYVAL in the specified title text string at the position where you want the substitution text to appear.
- Follow #BYVAL with a delimiting character, either a space or other nonalphanumeric character (for example, a quotation mark) that ends the text string.
- If you want the #BYVAL substitution to be followed immediately by other text, with no delimiter, use a trailing dot (as with macro variables).

Specify the variable with one of the following:

n
specifies which variable in the BY statement #BYVAL should use. The value of *n* indicates the position of the variable in the BY statement.

Example: #BYVAL2 specifies the second variable in the BY statement.

variable-name
names the BY variable.

Example: #BYVAL(YEAR) specifies the BY variable, YEAR.

Tip: *Variable-name* is not case sensitive.

#BYVAR*n* | #BYVAR(*variable-name*)
substitutes the name of the BY variable or label that is associated with the variable (whatever the BY line would normally display) for #BYVAR in the text string and displays the name or label in the title.

Follow these rules when you use #BYVAR in the TITLE statement of a PROC step:

- Specify the variable that is used by #BYVAR in the BY statement.
- Insert #BYVAR in the specified title text string at the position where you want the substitution text to appear.
- Follow #BYVAR with a delimiting character, either a space or other nonalphanumeric character (for example, a quotation mark) that ends the text string.
- If you want the #BYVAR substitution to be followed immediately by other text, with no delimiter, use a trailing dot (as with macro variables).

Specify the variable with one of the following:

n
specifies which variable in the BY statement #BYVAR should use. The value of *n* indicates the position of the variable in the BY statement.

Example: #BYVAR2 specifies the second variable in the BY statement.

variable-name
names the BY variable.

Example: #BYVAR(SITES) specifies the BY variable SITES.

Tip: *variable-name* is not case sensitive.

#BYLINE
substitutes the entire BY line without leading or trailing blanks for #BYLINE in the text string and displays the BY line in the title.

Tip: #BYLINE produces output that contains a BY line at the top of the page unless you suppress it by using NOBYLINE in an OPTIONS statement.

See Also: For more information on NOBYLINE, see "BYLINE System Option" on page 1856.

Tip: For compatibility with previous releases, SAS accepts some text without quotation marks. When writing new programs or updating existing programs, always enclose text in quotation marks.

Tip: If you use single quotation marks (') or double quotation marks ("") together (with no space in between them) as the string of text, SAS will output a single quotation mark (') or double quotation marks (""), respectively.

Tip: If you use an automatic macro variable in the title text, you must enclose the title text in double quotation marks. The SAS macro facility will resolve the macro variable only if the text is in double quotation marks.

See Also: For more information about including quotation marks as part of the title, see "Expressions" in *SAS Language Reference: Concepts*.

Details

In a DATA Step or PROC Step A TITLE statement takes effect when the step or RUN group with which it is associated executes. Once you specify a title for a line, it is used for all subsequent output until you cancel the title or define another title for that line. A TITLE statement for a given line cancels the previous TITLE statement for that line and for all lines with larger *n* numbers.

Operating Environment Information: The maximum title length that is allowed depends on your operating environment and the value of the LINESIZE= system option. Refer to the SAS documentation for your operating environment for more information. △

Comparisons

You can also create titles with the TITLES window.

Examples

Example 1: Using the TITLE Statement The following examples show how you can use the TITLE statement:

- This statement suppresses a title on line *n* and all lines after it:

```
titlen;
```

- These code lines are examples of TITLE statements:

```
title 'First Draft';

title2 "Year's End Report";

title2 'Year''s End Report';
```

Example 2: Customizing Titles by Using BY Variable Values You can customize titles by inserting BY variable values in the titles that you specify in PROC steps. The following examples show how to use #BYVAL*n*, #BYVAR*n*, and #BYLINE:

- `title 'Quarterly Sales for #byval(site)';`
- `title 'Annual Costs for #byvar2';`
- `title 'Data Group #byline';`

Example 3: Customizing Titles and Footnotes by Using the Output Delivery System You can customize titles and footnotes with ODS. The following example shows you how to use PROC TEMPLATE to change the color, justification, and size of the text for the title and footnote.

```
/*******************************************
 *The following program creates the data set *
 *grain_production and the $cntry format.    *
 *******************************************/
data grain_production;
   length Country $ 3 Type $ 5;
   input Year country $ type $ Kilotons;
   datalines;

1995 BRZ  Wheat    1516
1995 BRZ  Rice     11236
1995 BRZ  Corn     36276
1995 CHN  Wheat    102207
1995 CHN  Rice     185226
1995 CHN  Corn     112331
1995 IND  Wheat    63007
1995 IND  Rice     122372
1995 IND  Corn     9800
1995 INS  Wheat    .
1995 INS  Rice     49860
1995 INS  Corn     8223
1995 USA  Wheat    59494
1995 USA  Rice     7888
1995 USA  Corn     187300
1996 BRZ  Wheat    3302
1996 BRZ  Rice     10035
1996 BRZ  Corn     31975
1996 CHN  Wheat    109000
1996 CHN  Rice     190100
1996 CHN  Corn     119350
1996 IND  Wheat    62620
1996 IND  Rice     120012
1996 IND  Corn     8660
1996 INS  Wheat    .
1996 INS  Rice     51165
1996 INS  Corn     8925
1996 USA Wheat    62099
1996 USA  Rice     7771
1996 USA  Corn     236064
;
run;
```

```
  proc format;
   value $cntry 'BRZ'='Brazil'
                'CHN'='China'
                'IND'='India'
                'INS'='Indonesia'
                'USA'='United States';
run;

/***************************************
 *This PROC TEMPLATE step creates the    *
 *table definition TABLE1 that is used   *
 *in the DATA step.                      *
 ***************************************/
proc template;
   define table table1;
      mvar sysdate9;
      dynamic colhd;
      classlevels=on;
   define column char_var;
      generic=on;
      blank_dups=on;
      header=colhd;
      style=cellcontents;
   end;

   define column num_var;
      generic=on;
      header=colhd;
      style=cellcontents;
   end;

   define footer table_footer;
   end;
 end;
run;

/*************************************************************************
 *The ODS LISTING CLOSE statement closes the Listing                     *
 *destination to conserve resources.                                     *
 *                                                                       *
 *The ODS HTML statement creates HTML output created with                *
 *the style defintion D3D.                                               *
 *                                                                       *
 *The TITLE statement specifies the text for the first title             *
 *and the attributes that ODS uses to modify it.                         *
 *The J= style attribute left-justifies the title.                       *
 *The COLOR= style attributes change the color of the title text         *
 *"Leading Grain" to blue and "Producers in" to green.                   *
 *                                                                       *
 *The TITLE2 statement specifies the text for the second title           *
 *and the attributes that ODS uses to modify it.                         *
 *The J= style attribute center justifies the title.                     *
 *The COLOR= attribute changes the color of the title text  "1996"       *
 *to red.                                                                *
 * The HEIGHT= attributes change the size of each                        *
```

```
 *individual number in "1996".                                          *
 *                                                                      *
 *The FOOTNOTE statement specifies the text for the first footnote     *
 *and the attributes that ODS uses to modify it.                       *
 *The J=left style attribute left-justifies the footnote.              *
 *The HEIGHT=20 style attribute changes the font size to 20pt.         *
 *The COLOR= style attributes change the color of the footnote text    *
 *"Prepared" to red and "on" to green.                                 *
 *                                                                      *
 *The FOOTNOTE2 statement specifies the text for the second footnote   *
 *and the attributes that ODS uses to modify it.                       *
 *The J= style attribute centers the footnote.                         *
 *The COLOR= attribute changes the color of the date                   *
 *to blue,                                                             *
 *The HEIGHT= attribute changes the font size                          *
 *of the date specified by the sysdate9 macro.                         *
 ************************************************************************/
ods listing close;

ods html body='newstyle-body.htm'
   style=d3d;

title j=left
      font= 'Times New Roman' color=blue bcolor=red "Leading Grain "
      c=green bold italic "Producers in";

title2 j=center color=red underlin=1
       height=28pt "1"
       height=24pt "9"
       height=20pt "9"
       height=16pt "6";

footnote j=left height=20pt
       color=red "Prepared "
       c='#FF9900' "on";

footnote2 j=center color=blue
          height=24pt "&sysdate9";
footnote3 link='http://support.sas.com' "SAS";
/*********************************************************
 *This step uses the DATA step and ODS to produce          *
 *an HTML report. It uses the default table definition     *
 *(template) for the DATA step and writes an output object *
 *to the HTML destination.                                 *
 *********************************************************/
 data _null_;
   set grain_production;
   where type  in ('Rice', 'Corn') and year=1996;
   file print ods=(
                template='table1'
                columns=(
                   char_var=country(generic=on format=$cntry.
                   dynamic=(colhd='Country'))
               char_var=type(generic dynamic=(colhd='Year'))
```

```
                num_var=kilotons(generic=on format=comma12.
                dynamic=(colhd='Kilotons'))
            )
        );

put _ods_;
run;

ods html close;
ods listing;
```

Display 6.1 Output with Customized Titles and Footnotes

See Also

Statement:

"FOOTNOTE Statement" on page 1573

System Option:

"LINESIZE= System Option" on page 1936

"The TEMPLATE Procedure" in the *SAS Output Delivery System: User's Guide*

UPDATE Statement

Updates a master file by applying transactions.

Valid: in a DATA step

Category: File-handling

Type: Executable

Syntax

UPDATE *master-data-set*<(*data-set-options*)> *transaction-data-set*<(*data-set-options*)>

<END=*variable*>

<UPDATEMODE=
MISSINGCHECK|NOMISSINGCHECK>;

BY *by-variable*;

Arguments

master-data-set
specifies the SAS data set used as the master file.

Range: The name can be a one-level name (for example, FITNESS), a two-level name (for example, IN.FITNESS), or one of the special SAS data set names.

See Also: "SAS Names and Words" in *SAS Language Reference: Concepts*.

(data-set-options)
specifies actions SAS is to take when it reads variables into the DATA step for processing.

Requirements: *Data-set-options* must appear within parentheses and follow a SAS data set name.

Tip: Dropping, keeping, and renaming variables is often useful when you update a data set. Renaming like-named variables prevents the second value that is read from over-writing the first one. By renaming one variable, you make the values of both of them available for processing, such as comparing.

Featured in: Example 2 on page 1789

See Also: A list of data set options to use with input data sets in "Data Set Options by Category" on page 12.

transaction-data-set
specifies the SAS data set that contains the changes to be applied to the master data set.

Range: The name can be a one-level name (for example, HEALTH), a two-level name (for example, IN.HEALTH), or one of the special SAS data set names.

END=*variable*
creates and names a temporary variable that contains an end-of-file indicator. This variable is initialized to 0 and is set to 1 when UPDATE processes the last observation. This variable is not added to any data set.

UPDATEMODE=MISSINGCHECK
UPDATEMODE=NOMISSINGCHECK
specifies whether missing variable values in a transaction data set are to be allowed to replace existing variable values in a master data set.

MISSINGCHECK
prevents missing variable values in a transaction data set from replacing values in a master data set.

NOMISSINGCHECK
allows missing variable values in a transaction data set to replace values in a master data set.

Default: MISSINGCHECK

Tip: Special missing values, however, are the exception and will replace values in the master data set even when MISSINGCHECK (the default) is in effect.

Details

Requirements

- The UPDATE statement must be accompanied by a BY statement that specifies the variables by which observations are matched.
- The BY statement should immediately follow the UPDATE statement to which it applies.
- The data sets listed in the UPDATE statement must be sorted by the values of the variables listed in the BY statement, or they must have an appropriate index.
- Each observation in the master data set should have a unique value of the BY variable or BY variables. If there are multiple values for the BY variable, only the first observation with that value is updated. The transaction data set can contain more than one observation with the same BY value. (Multiple transaction observations are all applied to the master observation before it is written to the output file.)

For more information, see "How to Prepare Your Data Sets" in *SAS Language Reference: Concepts*.

Transaction Data Sets Usually, the master data set and the transaction data set contain the same variables. However, to reduce processing time, you can create a transaction data set that contains only those variables that are being updated. The transaction data set can also contain new variables to be added to the output data set.

The output data set contains one observation for each observation in the master data set. If any transaction observations do not match master observations, they become new observations in the output data set. Observations that are not to be updated can be omitted from the transaction data set. See "Reading, Combining, and Modifying SAS Data Sets" in *SAS Language Reference: Concepts*.

Missing Values By default the UPDATEMODE=MISSINGCHECK option is in effect, so missing values in the transaction data set do *not* replace existing values in the master data set. Therefore, if you want to update some but not all variables and if the variables you want to update differ from one observation to the next, set to missing those variables that are not changing. If you want missing values in the transaction data set to replace existing values in the master data set, use UPDATEMODE=NOMISSINGCHECK.

Even when UPDATEMODE=MISSINGCHECK is in effect, you can replace existing values with missing values by using special missing value characters in the transaction

data set. To create the transaction data set, use the MISSING statement in the DATA step. If you define one of the special missing values **A** through **Z** for the transaction data set, SAS updates numeric variables in the master data set to that value.

If you want the resulting value in the master data set to be a regular missing value, use a single underscore (_) to represent missing values in the transaction data set. The resulting value in the master data set will be a period (.) for missing numeric values and a blank for missing character values.

For more information about defining and using special missing value characters, see "MISSING Statement" on page 1682.

Comparisons

- Both UPDATE and MERGE can update observations in a SAS data set.
- MERGE automatically replaces existing values in the first data set with missing values in the second data set. UPDATE, however, does not do so by default. To cause UPDATE to overwrite existing values in the master data set with missing ones in the transaction data set, you must use UPDATEMODE=NOMISSINGCHECK.
- UPDATE changes or updates the values of selected observations in a master file by applying transactions. UPDATE can also add new observations.

Examples

Example 1: Basic Updating These program statements create a new data set (OHIO.QTR1) by applying transactions to a master data set (OHIO.JAN). The BY variable STORE must appear in both OHIO.JAN and OHIO.WEEK4, and its values in the master data set should be unique:

```
data ohio.qtr1;
   update ohio.jan ohio.week4;
   by store;
run;
```

Example 2: Updating By Renaming Variables This example shows renaming a variable in the FITNESS data set so that it will not overwrite the value of the same variable in the program data vector. Also, the WEIGHT variable is renamed in each data set and a new WEIGHT variable is calculated. The master data set and the transaction data set are listed before the code that performs the update:

```
Master Data Set
            HEALTH

OBS    ID     NAME     TEAM    WEIGHT

 1    1114    sally    blue      125
 2    1441    sue      green     145
 3    1750    joey     red       189
 4    1994    mark     yellow    165
 5    2304    joe      red       170

Transaction Data Set
            FITNESS

OBS   ID     NAME     TEAM    WEIGHT

 1   1114    sally    blue      119
```

```
 2   1994    mark     yellow    174
 3   2304    joe      red       170

options nodate pageno=1 linesize=80 pagesize=60;

   /* Sort both data sets by ID */
proc sort data=health;
   by id;
run;
proc sort data=fitness;
   by id;
run;

   /* Update Master with Transaction */
data health2;
   length STATUS $11;
   update health(rename=(weight=ORIG) in=a)
          fitness(drop=name team in=b);
   by id ;
   if a and b then
      do;
         CHANGE=abs(orig - weight);
         if weight<orig then status='loss';
         else if weight>orig then status='gain';
         else status='same';
      end;
   else status='no weigh in';
run;

options nodate ls=78;

proc print data=health2;
   title 'Weekly Weigh-in Report';
run;
```

Output 6.33 Updating By Renaming Variables

```
                    Weekly Weigh-in Report              1

 OBS STATUS        ID   NAME   TEAM    ORIG  WEIGHT  CHANGE

  1  loss         1114  sally  blue     125    119       6
  2  no weigh in  1441  sue    green    145      .       .
  3  no weigh in  1750  joey   red      189      .       .
  4  gain         1994  mark   yellow   165    174       9
  5  same         2304  joe    red      170    170       0
```

Example 3: Updating with Missing Values This example illustrates the DATA steps used to create a master data set PAYROLL and a transaction data set INCREASE that contains regular and special missing values:

```
options nodate pageno=1 linesize=80 pagesize=60;

   /* Create the Master Data Set */
data payroll;
```

```
   input ID SALARY;
   datalines;
011 245
026 269
028 374
034 333
057 582
;

   /* Create the Transaction Data Set */
data increase;
   input ID SALARY;
   missing A _;
   datalines;
011 376
026 .
028 374
034 A
057 _
;

   /* Update Master with Transaction */
data newpay;
   update payroll increase;
   by id;
run;
proc print data=newpay;
   title 'Updating with Missing Values';
run;
```

Output 6.34 Updating With Missing Values

```
      Updating with Missing Values              1

          OBS     ID     SALARY

           1     1011      376
           2     1026      269    <=== value remains 269
           3     1028      374
           4     1034        A    <=== special missing value
           5     1057        .    <=== regular missing value
```

See Also

Statements:

"BY Statement" on page 1452
"MERGE Statement" on page 1679
"MISSING Statement" on page 1682
"MODIFY Statement" on page 1684
"SET Statement" on page 1764

System Option:

"MISSING= System Option" on page 1944

"Reading, Combining, and Modifying SAS Data Sets" in *SAS Language Reference: Concepts*

"Definition of Data Set Options" on page 10

WHERE Statement

Selects observations from SAS data sets that meet a particular condition.

Valid: in DATA and PROC steps

Category: Action

Type: Declarative

Syntax

WHERE *where-expression-1*
< *logical-operator where-expression-n*>;

Arguments

where-expression
is an arithmetic or logical expression that generally consists of a sequence of operands and operators.

Tip: The operands and operators described in the next several sections are also valid for the WHERE= data set option.

Tip: You can specify multiple where-expressions.

logical-operator
can be AND, AND NOT, OR, or OR NOT.

Details

The Basics Using the WHERE statement might improve the efficiency of your SAS programs because SAS is not required to read all observations from the input data set.

The WHERE statement cannot be executed conditionally. That is, you cannot use it as part of an IF-THEN statement.

WHERE statements can contain multiple WHERE expressions that are joined by logical operators.

Note: Using indexed SAS data sets can significantly improve performance when you use WHERE expressions to access a subset of the observations in a SAS data set. See "Understanding SAS Indexes" in the "SAS Data Files" section of *SAS Language Reference: Concepts* for a complete discussion of WHERE-expression processing with indexed data sets and a list of guidelines to consider before you index your SAS data sets. △

In DATA Steps The WHERE statement applies to all data sets in the preceding SET, MERGE, MODIFY, or UPDATE statement, and variables that are used in the WHERE statement must appear in all of those data sets. You cannot use the WHERE statement with the POINT= option in the SET and MODIFY statements.

You can apply OBS= and FIRSTOBS= processing to WHERE processing. For more information, see "Processing a Segment of Data That is Conditionally Selected" in the "WHERE-Expression Processing" section of *SAS Language Reference: Concepts*.

You cannot use the WHERE statement to select records from an external file that contains raw data, nor can you use the WHERE statement within the same DATA step in which you read in-stream data with a DATALINES statement.

For each iteration of the DATA step, the first operation SAS performs in each execution of a SET, MERGE, MODIFY, or UPDATE statement is to determine whether the observation in the input data set meets the condition of the WHERE statement. The WHERE statement takes effect immediately after the input data set options are applied and before any other statement in the DATA step is executed. If a DATA step combines observations using a WHERE statement with a MERGE, MODIFY, or UPDATE statement, SAS selects observations from each input data set before it combines them.

WHERE and BY in a DATA Step If a DATA step contains both a WHERE statement and a BY statement, the WHERE statement executes *before* BY groups are created. Therefore, BY groups reflect groups of observations in the subset of observations that are selected by the WHERE statement, not the actual BY groups of observations in the original input data set.

For a complete discussion of BY-group processing, see "BY-Group Processing in SAS Programs" in *SAS Language Reference: Concepts*.

In PROC Steps You can use the WHERE statement with any SAS procedure that reads a SAS data set. The WHERE statement is useful in order to subset the original data set for processing by the procedure. The *Base SAS Procedures Guide* documents the action of the WHERE statement only in those procedures for which you can specify more than one data set. In all other cases, the WHERE statement performs as documented here.

Use of Indexes A DATA or PROC step attempts to use an available index to optimize the selection of data when an indexed variable is used in combination with one of the following operators and functions:

- the BETWEEN-AND operator
- the comparison operators, with or without the colon modifier
- the CONTAINS operator
- the IS NULL and IS NOT NULL operators
- the LIKE operator
- the TRIM function
- the SUBSTR function, in some cases.

SUBSTR requires the following arguments:

```
where substr(variable,position,length)
 ='character-string';
```

An index is used in processing when the arguments of the SUBSTR function meet all of the following conditions:

- *position* is equal to 1
- *length* is less than or equal to the length of *variable*
- *length* is equal to the length of *character-string*.

Operands Used in WHERE Expressions Operands in WHERE expressions can contain the following values:

- constants
- time and date values

- values of variables that are obtained from the SAS data sets
- values created within the WHERE expression itself.

You cannot use variables that are created within the DATA step (for example, FIRST.*variable*, LAST.*variable*, _N_, or variables that are created in assignment statements) in a WHERE expression because the WHERE statement is executed before the SAS System brings observations into the DATA or PROC step. When WHERE expressions contain comparisons, the unformatted values of variables are compared.

The following are examples of using operands in WHERE expressions:

- where score>50;
- where date>='01jan1999'd and time>='9:00't;
- where state='Mississippi';

As in other SAS expressions, the names of numeric variables can stand alone. SAS treats values of 0 or missing as false; other values are true. These examples are WHERE expressions that contain the numeric variables EMPNUM and SSN:

- where empnum;
- where empnum and ssn;

Character literals or the names of character variables can also stand alone in WHERE expressions. If you use the name of a character variable by itself as a WHERE expression, SAS selects observations where the value of the character variable is not blank.

Operators Used in the WHERE Expression You can include both SAS operators and special WHERE-expression operators in the WHERE statement. For a complete list of the operators, see Table 6.12 on page 1794. For the rules SAS follows when it evaluates WHERE expressions, see "WHERE-Expression Processing" in *SAS Language Reference: Concepts*.

Table 6.12 WHERE Statement Operators

Operator Type	Symbol or Mnemonic	Description
Arithmetic		
	*	multiplication
	/	division
	+	addition
	–	subtraction
	**	exponentiation
Comparison [4]		
	= or EQ	equal to
	^=, ¬=, ~=, or NE[1]	not equal to
	> or GT	greater than
	< or LT	less than
	>= or GE	greater than or equal to
	<= or LE	less than or equal to
	IN	equal to one of a list
Logical (Boolean)		

Operator Type	Symbol or Mnemonic	Description
	& or AND	logical and
	| or OR[2]	logical or[1]
	~,^ , ¬, or NOT[1]	logical not
Other		
	||[3]	concatenation of character variables
	( )	indicate order of evaluation
	+ prefix	positive number
	– prefix	negative number
WHERE Expression Only		
	BETWEEN–AND	an inclusive range
	? or CONTAINS	a character string
	IS NULL or IS MISSING	missing values
	LIKE	match patterns
	=*	sounds-like
	SAME-AND	add clauses to an existing WHERE statement without retyping original one

1 The caret (^), tilde (~), and the not sign (¬)all indicate a logical not. Use the character available on your keyboard, or use the mnemonic equivalent.

2 The OR symbol (|), broken vertical bar (|), and exclamation point (!) all indicate a logical or. Use the character available on your keyboard, or use the mnemonic equivalent.

3 Two OR symbols (| |), two broken vertical bars (| |), or two exclamation points (!!) indicate concatenation. Use the character available on your keyboard.

4 You can use the colon modifier (:) with any of the comparison operators in order to compare only a specified prefix of a character string.

Comparisons

- You can use the WHERE command in SAS/FSP software to subset data for editing and browsing. You can use both the WHERE statement and WHERE= data set option in windowing procedures and in conjunction with the WHERE command.
- To select observations from individual data sets when a SET, MERGE, MODIFY, or UPDATE statement specifies more than one data set, apply a WHERE= data set option to each data set. In the DATA step, if a WHERE statement and a WHERE= data set option apply to the same data set, SAS uses the data set option and ignores the statement.
- The most important differences between the WHERE statement in the DATA step and the subsetting IF statement are as follows:
  - The WHERE statement selects observations *before* they are brought into the program data vector, making it a more efficient programming technique. The subsetting IF statement works on observations after they are read into the program data vector.
  - The WHERE statement can produce a different data set from the subsetting IF when a BY statement accompanies a SET, MERGE, or UPDATE statement. The different data set occurs because SAS creates BY groups before the subsetting IF statement selects but after the WHERE statement selects.

- The WHERE statement cannot be executed conditionally as part of an IF statement, but the subsetting IF statement can.
- The WHERE statement selects observations in SAS data sets only, whereas the subsetting IF statement selects observations from an existing SAS data set or from observations that are created with an INPUT statement.
- The subsetting IF statement cannot be used in SAS windowing procedures to subset observations for browsing or editing.

- Do not confuse the WHERE statement with the DROP or KEEP statement. The DROP and KEEP statements select variables for processing. The WHERE statement selects observations.

Examples

Example 1: Basic WHERE Statement Usage This DATA step produces a SAS data set that contains only observations from data set CUSTOMER in which the value for NAME begins with **Mac** and the value for CITY is **Charleston** or **Atlanta**.

```
data testmacs;
   set customer;
   where substr(name,1,3)='Mac' and
         (city='Charleston' or city='Atlanta');
run;
```

Example 2: Using Operators Available Only in the WHERE Statement

- Using BETWEEN-AND:

```
where empnum between 500 and 1000;
```

- Using CONTAINS:

```
where company ? 'bay';
where company contains 'bay';
```

- Using IS NULL and IS MISSING:

```
where name is null;
where name is missing;
```

- Using LIKE to select all names that start with the letter D:

```
where name like 'D%';
```

- Using LIKE to match patterns from a list of the following names:

  Diana
  Diane
  Dianna
  Dianthus
  Dyan

WHERE Statement	Name Selected
where name like 'D_an';	Dyan
where name like 'D_an_';	Diana, Diane
where name like 'D_an__';	Dianna
where name like 'D_an%';	all names from list

- Using the Sounds-like Operator to select names that sound like “Smith”:

```
where lastname=*'Smith';
```

- Using SAME-AND:

```
where year>1991;
...more SAS statements...
where same and year<1999;
```

In this example, the second WHERE statement is equivalent to the following WHERE statement:

```
where year>1991 and year<1999;
```

See Also

Data Set Option:

“WHERE= Data Set Option” on page 67

Statement:

“IF Statement, Subsetting” on page 1581

SAS SQL Query Window User's Guide

SAS/IML User's Guide

Base SAS Procedures Guide

“SAS Indexes” in *SAS Language Reference: Concepts*

“WHERE-Expression Processing” in *SAS Language Reference: Concepts*

“BY-Group Processing” in *SAS Language Reference: Concepts*

Beatrous, S. & Clifford, W. (1998), “Sometimes You Do Get What You Want: SAS I/O Enhancements in Version 7,” *Proceedings of the Twenty–third Annual SAS Users Group International Conference*, 23.

WINDOW Statement

Creates customized windows for your applications.

Valid: in a DATA step

Category: Window Display

Type: Declarative

Syntax

WINDOW *window* <*window-options*> *field-definition(s)*;

WINDOW *window* <*window-options*> *group-definition(s)*;

Arguments

window
specifies the window name.

Restriction: Window names must conform to SAS naming conventions.

window-options
specifies characteristics of the window as a whole. Specify these *window-options* before any field or GROUP= specifications:

COLOR=*color*
specifies the color of the window background for operating environments that have this capability. In other operating environments, this option affects the color of the window border. The following colors are available:

BLACK

BLUE

BROWN

CYAN

GRAY

GREEN

MAGENTA

ORANGE

PINK

RED

WHITE

YELLOW

Default: If you do not specify a color with the COLOR= option, the window's background color is device-dependent instead of black, and the color of a field is device-dependent instead of white.

Tip: The representation of colors might vary, depending on the monitor being used. COLOR= has no effect on monochrome monitors.

COLUMNS=*columns*
specifies the number of columns in the window.

Default: The window fills all remaining columns on the monitor; the number of columns that are available depends on the type of monitor that is being used.

ICOLUMN=*column*
specifies the initial column within the monitor at which the window is displayed.

Default: SAS displays the window at column 1.

IROW=*row*
specifies the initial row (or line) within the monitor at which the window is displayed.

Default: SAS displays the window at row 1.

KEYS=<<*libref.*>*catalog.*>*keys-entry*
specifies the name of a KEYS entry that contains the function key definitions for the window.

Default: SAS uses the current function key settings that are defined in the KEYS window.

Tip: If you specify only an entry name, SAS looks in the SASUSER.PROFILE catalog for a KEYS entry of the name that is specified. You can also specify the three-level name of a KEYS entry, in the form

```
libref.catalog.keys-entry
```

Tip: To create a set of function key definitions for a window, use the KEYS window. Define the keys as you want, and use the SAVE command to save the definitions in the SASUSER.PROFILE catalog or in a SAS library and catalog that you specify.

MENU=<<*libref.*>*catalog.*>*pmenu-entry*
specifies the name of a menu (pmenu) you have built with the PMENU procedure.

Tip: If you specify only an entry name, SAS looks in the SASUSER.PROFILE catalog for a PMENU entry of the name specified. You can also specify the three-level name of a PMENU entry in the form

```
libref.catalog.pmenu-entry
```

ROWS=*rows*
specifies the number of rows (or lines) in the window.

Default: The window fills all remaining rows on the monitor.

Tip: The number of rows that are available depends on the type of monitor that is being used.

field-definition(s)
specifies and describes a variable or character string to be displayed in a window or within a group of related fields.

Tip: A window or group can contain any number of fields, and you can define the same field in several groups or windows.

Tip: You can specify multiple *field-definitions*.

See Also: The form of *field-definition* is given in "Field Definitions" on page 1800.

group-definition(s)
specifies a group and defines all fields within a group. A group definition consists of two parts: the GROUP= option and one or more field definitions.

GROUP=*group*
specifies a group of related fields.

Restriction: *group* must be a SAS name.

Default: A window contains one unnamed group of fields.

Tip: When you refer to a group in a DISPLAY statement, write the name as *window.group*.

Tip: A group contains all fields in a window that you want to display at the same time. Display various groups of fields within the same window at different times by naming each group. Choose the group to appear by specifying *window.group* in the DISPLAY statement.

Tip: Specifying several groups within a window prevents repetition of window options that do not change and helps you to keep track of related displays. For example, if you are defining a window to check data values, arrange the display of variables and messages for most data values in the data set in a group that is named STANDARD. Arrange the display of different messages in a group that

is named CHECKIT that appears when data values meet the conditions that you want to check.

Details

Operating Environment Information: The WINDOW statement has some functionality that might be specific to your operating environment. For details, see the SAS documentation for your operating environment. △

You can use the WINDOW statement in the SAS windowing environment, in interactive line mode, or in noninteractive mode to create customized windows for your applications.* Windows that you create can display text and accept input; they have command and message lines. The window name appears at the top of the window. Use commands and function keys with windows that you create. A window definition remains in effect only for the DATA step that contains the WINDOW statement.

Define a window before you display it. Use the DISPLAY statement to display windows that are created with the WINDOW statement. For information about the DISPLAY statement, see "DISPLAY Statement" on page 1488.

Field Definitions Use a field definition to identify a variable or a character string to be displayed, its position, and its attributes. Enclose character strings in quotation marks. The position of an item is its beginning row (or line) and column. Attributes include color, whether you can enter a value into the field, and characteristics such as highlighting.

You can define a field to contain a variable value or a character string, but not both. The form of a field definition for a variable value is

<row column> variable <format> options

The form for a character string is

<row column> 'character-string' options

The elements of a field definition are described here.

row column
specifies the position of the variable or character string.

Default: If you omit *row* in the first field of a window or group, SAS uses the first row of the window. If you omit *row* in a later field specification, SAS continues on the row that contains the previous field. If you omit *column*, SAS uses column 1 (the left border of the window).

Tip: Although you can specify either *row* or *column* first, the examples in this documentation show the row first.

SAS keeps track of its position in the window with a pointer. For example, when you tell SAS to write a variable's value in the third column of the second row of a window, the pointer moves to row 2, column 3 to write the value. Use the pointer controls that are listed here to move the pointer to the appropriate position for a field.

In a field definition, *row* can be one of these row pointer controls:

#n
specifies row *n* within the window.

Range: *n* must be a positive integer.

* You cannot use the WINDOW statement in batch mode because no computer is connected to a batch executing process.

#*numeric-variable*
specifies the row within the window that is given by the value of *numeric-variable*.

Restriction: #*numeric-variable* must be a positive integer. If the value is not an integer, the decimal portion is truncated and only the integer is used.

#(*expression*)
specifies the row within the window that is given by the value of *expression*.

Restriction: *expression* can contain array references and must evaluate to a positive integer.

Restriction: Enclose *expression* in parentheses.

/
moves the pointer to column 1 of the next row.

In a field definition, *column* can be one of these column pointer controls:

@*n*
specifies column *n* within the window.

Restriction: *n* must be a positive integer.

@*numeric-variable*
specifies the column within the window that is given by the value of *numeric-variable*.

Restriction: *numeric-variable* must be a positive integer. If the value is not an integer, the decimal portion is truncated and only the integer is used.

@(*expression*)
specifies the column within the window that is given by the value of *expression*.

Restriction: *expression* can contain array references and must evaluate to a positive integer.

Restriction: Enclose *expression* in parentheses.

+*n*
moves the pointer *n* columns.

Range: *n* must be a positive integer.

+*numeric-variable*
moves the pointer the number of columns that is given by the *numeric-variable*.

Restriction: +*numeric-variable* must be a positive or negative integer. If the value is not an integer, the decimal portion is truncated and only the integer is used.

variable
specifies a variable to be displayed or to be assigned the value that you enter at that position when the window is displayed.

Tip: *variable* can be the name of a variable or of an array reference.

Tip: To allow a variable value in a field to be displayed but not changed by the user, use the PROTECT= option (described later in this section). You can also protect an entire window or group for the current execution of the DISPLAY statement by specifying the NOINPUT option in the DISPLAY statement.

Tip: If a field definition contains the name of a new variable, that variable is added to the data set that is being created (unless you use a KEEP or DROP specification).

format
gives the format for the variable.

Default: If you omit *format*, SAS uses an informat and format that are specified elsewhere (for example, in an ATTRIB, INFORMAT, or FORMAT statement or permanently stored with the data set) or a SAS default informat and format.

Tip: If a field displays a variable that cannot be changed (that is, you use the PROTECT=YES option), *format* can be any SAS format or a format that you define with the FORMAT procedure.

Tip: If a field can both display a variable and accept input, you must either specify the informat in an INFORMAT or ATTRIB statement or use a SAS format such as $CHAR. or TIME. that has a corresponding informat.

Tip: If a format is specified, the corresponding informat is assigned automatically to fields that can accept input.

Tip: A format and an informat in a WINDOW statement override an informat and a format that are specified elsewhere.

'*character-string*'
contains the text of a character string to be displayed.

Restriction: The character string must be enclosed in quotation marks.

Restriction: You cannot enter a value in a field that contains a character string.

options
Specify field definition attributes:

ATTR=*highlighting-attribute*
controls these highlighting attributes of the field:

BLINK
causes the field to blink.

HIGHLIGHT
displays the field at high intensity.

REV_VIDEO
displays the field in reverse video.

UNDERLINE
underlines the field.

Alias: A=

Tip: To specify more than one highlighting attribute, use the form

ATTR=(*highlighting-attribute-1*, . . .)

Tip: The highlighting attributes that are available depend on the type of monitor that you use.

AUTOSKIP=YES | NO
controls whether the cursor moves to the next unprotected field of the current window or group when you have entered data in all positions of a field.

YES	specifies that the cursor moves automatically to the next unprotected field.
NO	specifies that the cursor does not move automatically.

Alias: AUTO=

Default: NO

COLOR=*color*
specifies a color for the variable or character string. You can specify one of the following colors:

BLACK

BLUE

BROWN

CYAN

GRAY

GREEN

MAGENTA

ORANGE

PINK

RED

WHITE

YELLOW

Alias: C=

Default: WHITE

Tip: The representation of colors might vary, depending on the monitor you use.

Tip: COLOR= has no effect on monochrome monitors.

DISPLAY=YES | NO
controls whether the contents of a field are displayed.

YES — specifies that SAS displays characters in a field as you type them in.

NO — specifies that the entered characters are not displayed.

Default: YES

PERSIST=YES | NO
controls whether a field is displayed by all executions of a DISPLAY statement in the same iteration of the DATA step until the DISPLAY statement contains the BLANK option.

YES — specifies that each execution of the DISPLAY statement displays all previously displayed contents of the field as well as the contents that are scheduled for display by the current DISPLAY statement. If the new contents overlap persisting contents, the persisting contents are no longer displayed.

NO — specifies that each execution of a DISPLAY statement displays only the current contents of the field.

Default: NO

Tip: PERSIST= is most useful when the position of a field changes in each execution of a DISPLAY statement.

Featured in: Example 3 on page 1806

PROTECT=YES | NO
controls whether information can be entered into a field.

YES specifies that you cannot enter information.

NO specifies that you can enter information.

Alias: P=

Default: No

Tip: Use PROTECT= only for fields that contain variables; fields that contain text are automatically protected.

REQUIRED=YES | NO
controls whether a field can be left blank.

NO specifies that you can leave the field blank.

YES specifies that you must enter a value in the field.

Default: NO

Tip: If you try to leave a field blank that was defined with REQUIRED=YES, SAS does not allow you to input values in any subsequent fields in the window.

Automatic Variables The WINDOW statement creates two automatic SAS variables: _CMD_ and _MSG_.

CMD contains the last command from the window's command line that was not recognized by the window.

Tip: _CMD_ is a character variable of length 80; its value is set to "(blank) before each execution of a DISPLAY statement.

Featured in: Example 4 on page 1807

MSG contains a message that you specify to be displayed in the message area of the window.

Tip: _MSG_ is a character variable with length 80; its value is set to "(blank) after each execution of a DISPLAY statement.

Featured in: Example 4 on page 1807

Displaying Windows The DISPLAY statement enables you to display windows. Once you display a window, the window remains visible until you display another window over it or until the end of the DATA step. When you display a window that contains fields into which you can enter values, either enter a value or press ENTER at *each* unprotected field to cause SAS to proceed to the next display. While a window is being displayed, you can use commands and function keys to view other windows, change the size of the current window, and so on. The execution proceeds to the next display only after you have pressed ENTER in all unprotected fields.

A DATA step that contains a DISPLAY statement continues execution until

- the last observation that is read by a SET, MERGE, MODIFY, UPDATE, or INPUT statement has been processed
- a STOP or ABORT statement is executed
- an END command executes.

Comparisons

- The WINDOW statement creates a window, and the DISPLAY statement displays it.
- The %WINDOW and %DISPLAY statements in the macro language create and display windows that are controlled by the macro facility.

Examples

Example 1: Creating a Single Window This DATA step creates a window with a single group of fields:

```
data _null_;
   window start
          #9  @26 'WELCOME TO THE SAS SYSTEM'
                  color=black
          #12  @19 'THIS PROGRAM CREATES'
          #12  @40 'TWO SAS DATA SETS'
          #14 @26 'AND USES THREE PROCEDURES'
          #18 @27 'Press ENTER to continue';
    display start;
    stop;
run;
```

The START window fills the entire monitor. The first line of text is black. The other three lines are the default for your operating environment. The text begins in the column that you specified in your program. The START window does not require you to input any values. However, to exit the window do one of the following:

- Press ENTER to cause DATA step execution to proceed to the STOP statement.
- Issue the END command.

If you omit the STOP statement from this program, the DATA step executes endlessly until you execute END from the window, either with a function key or from the command line. (Because this DATA step does not read any observations, SAS cannot detect an end-of-file to end DATA step execution.)

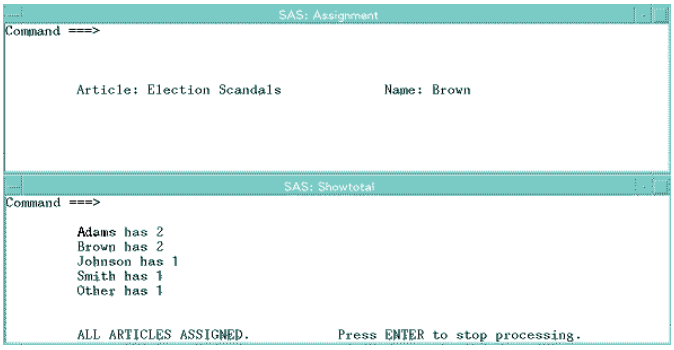
Example 2: Displaying Two Windows Simultaneously The following statements assign news articles to reporters. The list of article topics is stored as variable art in SAS data set category.article. This application allows you to assign each topic to a writer and to view the accumulating assignments. The program creates a new SAS data set named Assignment.

```
libname category 'SAS-library';

data Assignment;
   set category.article end=final;
   drop a b j s o;
   window Assignment irow=1 rows=12 color=white
          #3 @10 'Article:' +1 art protect=yes
          'Name:' +1 name $14.;
```

```
   window Showtotal irow=20 rows=12 color=white
          group=subtotal
          #1 @10 'Adams has' +1 a
          #2 @10 'Brown has' +1 b
          #3 @10 'Johnson has' +1 j
          #4 @10 'Smith has' +1 s
          #5 @10 'Other has' +1 o
          group=lastmessage
          #8 @10
          'ALL ARTICLES ASSIGNED.
          Press ENTER to stop processing.';
   display Assignment blank;
   if name='Adams' then a+1;
   else if name='Brown' then b+1;
   else if name='Johnson' then j+1;
   else if name='Smith' then s+1;
   else o+1;
   display Showtotal.subtotal blank noinput;
   if final then display Showtotal.lastmessage;
run;
```

When you execute the DATA step, the following windows appear.



In the Assignment window (located at the top of the monitor), you see the name of the article and a field into which you enter a reporter's name. After you type a name and press ENTER, SAS displays the Showtotal window (located at the bottom of the monitor) which shows the number of articles that are assigned to each reporter (including the assignment that you just made). As you continue to make assignments, the values in the Showtotal window are updated. During the last iteration of the DATA step, SAS displays the message that all articles are assigned, and instructs you to press ENTER to stop processing.

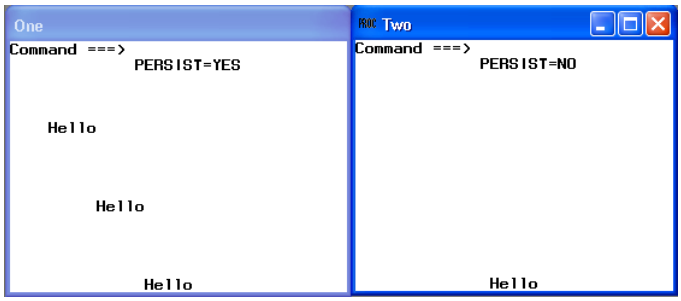
Example 3: Persisting and Nonpersisting Fields This example demonstrates the PERSIST= option. You move from one window to the other by positioning the cursor in the current window and pressing ENTER.

```
data _null_;
   array row{3} r1-r3;
   array col{3} c1-c3;
   input row{*} col{*};
   window One
          rows=20 columns=36
          #1 @14 'PERSIST=YES' color=black
          #(row{i}) @(col{i}) 'Hello'
```

```
          color=black persist=yes;

   window Two
          icolumn=43 rows=20 columns=36
          #1 @14 'PERSIST=NO' color=black
          #(row{i}) @(col{i}) 'Hello'
          color=black persist=no;
   do i=1 to 3;
      display One;
      display Two;
   end;
   datalines;
5 10 15 5 10 15
;
```

The following windows show the results of this DATA step after its third iteration.



Note that window One shows **Hello** in all three positions in which it was displayed. Window Two shows only the third and final position in which **Hello** was displayed.

Example 4: Sending a Message This example uses the _CMD_ and _MSG_ automatic variables to send a message when you execute an erroneous windowing command in a window that is defined with the WINDOW statement:

```
if _cmd_ ne ' ' then
   _msg_='CAUTION: UNRECOGNIZED COMMAND' || _cmd_;
```

When you enter a command that contains an error, SAS sets the value of _CMD_ to the text of the erroneous command. Because the value of _CMD_ is no longer blank, the IF statement is true. The THEN statement assigns to _MSG_ the value that is created by concatenating CAUTION: UNRECOGNIZED COMMAND and the value of _CMD_ (up to a total of 80 characters). The next time a DISPLAY statement displays that window, the message line of the window displays

```
CAUTION: UNRECOGNIZED COMMAND command
```

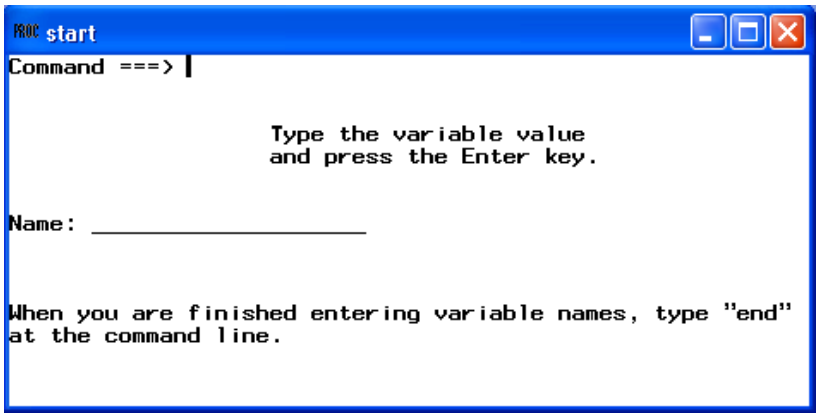
Command is the erroneous windowing command.

Example 5: Creating a SAS Data Set The following statements create a SAS data set by using input from the WINDOW statement.

```
data new;
   length name $20;
   window start
      #3  @20 'Type the variable name'
      #4  @20 'and press the Enter key.'
      #7 'Name:' +1 name attr=underline
```

```
      #11 'When you are finished entering variable names, type "end"'
      #12 'at the command line.';
   display start;
run;

proc print;
run;
```



See Also

Statements:

"DISPLAY Statement" on page 1488

"The PMENU Procedure" in *Base SAS Procedures Guide*

X Statement

Issues an operating-environment command from within a SAS session.

Valid: anywhere

Category: Operating Environment

See: X Statement in the documentation for your operating environment.

Syntax

X <*'operating-environment-command'*>;

Without Arguments

Using X without arguments places you in your operating environment, where you can issue commands that are specific to your environment.

Arguments

'operating-environment-command'
specifies an operating environment command that is enclosed in quotation marks.

Details

In all operating environments, you can use the X statement when you run SAS in windowing or interactive line mode. In some operating environments, you can use the X statement when you run SAS in batch or noninteractive mode.

Operating Environment Information: The X statement is dependent on your operating environment. See the SAS documentation for your operating environment to determine whether it is a valid statement on your system. Keep in mind:

- The way you return from operating environment mode to the SAS session is dependent on your operating environment.
- The commands that you use with the X statement are specific to your operating environment.

△

You can use the X statement with SAS macros to write a SAS program that can run in multiple operating environments. See *SAS Macro Language: Reference* for information.

Comparisons

In a windowing session, the X command works exactly like the X statement except that you issue the command from a command line. You submit the X statement from the Program Editor window.

The X statement is similar to the SYSTEM function, the X command, and the CALL SYSTEM routine. In most cases, the X statement, X command or %SYSEXEC macro statement are preferable because they require less overhead. However, the SYSTEM function can be executed conditionally. The X statement is a global statement and executes as a DATA step is being compiled.

See Also

CALL Routine:
"CALL SYSTEM Routine" on page 538
Function:
"SYSTEM Function" on page 1159

SAS Statements Documented in Other SAS Publications

In addition to system options documented in *SAS Language Reference: Dictionary*, statements are also documented in the following publications:

"*SAS Companion for Windows*" on page 1810
"*SAS Companion for OpenVMS on HP Integrity Servers*" on page 1810
"*SAS Companion for UNIX Environments*" on page 1810
"*SAS Companion for z/OS*" on page 1811
"*SAS Language Interfaces to Metadata*" on page 1811

SAS Companion for Windows

The statements listed here are documented only in *SAS Companion for Windows*. Other statements in *SAS Companion for Windows* contain information specific to the Windows operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter statements are not listed here.

Statement	Description
SYSTASK	Executes, lists, or terminates asynchronous tasks.
WAITFOR	Suspends execution of the current SAS session until the specified tasks finish executing.

SAS Companion for OpenVMS on HP Integrity Servers

The statements listed here are documented only in *SAS Companion for OpenVMS on HP Integrity Servers*. Other statements in *SAS Companion for OpenVMS on HP Integrity Servers* contain information specific to the OpenVMS operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter statements are not listed here.

Statement	**Description**
SYSTASK	Executes, lists, or kills asynchronous tasks.
WAITFOR	Suspends execution of the current SAS session until the specified tasks finish executing.

SAS Companion for UNIX Environments

The statements listed here are documented only in *SAS Companion for UNIX Environments*. Other statements in *SAS Companion for UNIX Environments* contain information specific to the UNIX operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter statements are not listed here.

Statement	Description
SYSTASK	Executes asynchronous tasks.
WAITFOR	Suspends execution of the current SAS session until the specified tasks finish executing.

SAS Companion for z/OS

The statements listed here are documented only in *SAS Companion for z/OS*. Other statements in *SAS Companion for z/OS* contain information specific to the z/OS operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter statements are not listed here.

Statement	Description
DSNEXST	Checks to see whether the specified physical file exists and is available.
SYSTASK LIST	Lists asynchronous tasks.
TSO	Issues a TSO command or invokes a CLIST or a REXX exec during a SAS session.
WAITFOR	Suspends execution of the current SAS session until the specified tasks finish executing.

SAS Language Interfaces to Metadata

Statement	Description
LIBNAME Statement for the Metadata Engine	Associates a SAS libref with the metadata that is in a SAS Metadata Repository on the SAS Metadata Server.

SAS Macro Language: Reference

Statement	Description
%ABORT	Stops the macro that is executing along with the current DATA step, SAS job, or SAS session.
%* Macro Comment	Designates comment text.
%COPY	Copies specified items from a SAS macro library.
%DISPLAY	Displays a macro window.
%DO	Begins a %DO group.

Statement	Description
%DO, Iterative	Executes a section of a macro repetitively based on the value of an index variable.
%DO %UNTIL	Executes a section of a macro repetitively until a condition is true.
%DO %WHILE	Executes a section of a macro repetitively while a condition is true.
%END	Ends a %DO group.
%GLOBAL	Creates macro variables that are available during the execution of an entire SAS session.
%GOTO	Branches macro processing to the specified label.
%IF-%THEN/%ELSE	Conditionally process a portion of a macro.
%INPUT	Supplies values to macro variables during macro execution.
%label	Identifies the destination of a %GOTO statement.
%LET	Creates a macro variable and assigns it a value.
%LOCAL	Creates macro variables that are available only during the execution of the macro where they are defined.
%MACRO	Begins a macro definition.
%MEND	Ends a macro definition.
%PUT	Writes text or macro variable information to the SAS log.
%RETURN	Execution causes normal termination of the currently executing macro.
%SYMDEL	Deletes the specified variable or variables from the macro global symbol table.
%SYSCALL	Invokes a SAS call routine.
%SYSEXEC	Issues operating environment commands.
%SYSLPUT	Creates a new macro variable or modifies the value of an existing macro variable on a remote host or server.
%SYSRPUT	Assigns the value of a macro variable on a remote host to a macro variable on the local host.
%WINDOW	Defines customized windows.

SAS Output Delivery System: User's Guide

Statement	Description
FILE, ODS	Creates an ODS output object by binding the data component to the table definition (template). Listing the variables to include in the ODS output, and specifying options that control the way that the variables are formatted is optional.
LIBNAME, SASEDOC	Uses the SASEDOC engine to associate a SAS libref (library reference) with one or more ODS output objects that are stored in an ODS document.

Statement	Description
ODS _ALL_CLOSE	Closes all open ODS output destinations.
ODS CHTML	Opens, manages, or closes the CHTML destination, which produces a compact, minimal HTML that does not use style information.
ODS CSVALL	Opens, manages, or closes the CSVALL destination, which produces HTML output containing columns of data values that are separated by commas, and produces tabular output with titles, notes, and by lines.
ODS DECIMAL_ALIGN	Controls the justification of numeric columns when no justification is specified.
ODS DOCBOOK	Opens, manages, or closes the DOCBOOK destination, which produces XML output that conforms to the DocBook DTD by OASIS.
ODS DOCUMENT	Opens, manages, or closes the DOCUMENT destination, which produces a hierarchy of output objects that enables you to produce multiple ODS output formats without rerunning a PROC or DATA step.
ODS ESCAPECHAR	Defines a representative character to be used in output strings.
ODS EXCLUDE	Specifies output objects to exclude from ODS destinations.
ODS GRAPHICS	Enables ODS automatic graphic capabilities.
ODS HTML	Opens, manages, or closes the HTML destination, which produces HTML 4.0 output that contains embedded style sheets.
ODS HTMLCSS	Opens, manages, or closes the HTMLCSS destination, which produces HTML output with cascading style sheets.
ODS HTML3	Opens, manages, or closes the HTML3 destination, which produces HTML 3.2 formatted output.
ODS IMODE	Opens, manages, or closes the IMODE destination, which produces HTML output as a column of output, separated by lines.
ODS LISTING	Opens, manages, or closes the LISTING destination.
ODS MARKUP	Opens, manages, or closes the MARKUP destination, which produces SAS output that is formatted using one of many different markup languages.
ODS OUTPUT	Produces a SAS data set from an output object and manages the selection and exclusion lists for the OUTPUT destination.
ODS PACKAGE	Opens, adds to, publishes, or closes one SAS ODS package object.
ODS PATH	Specifies locations to write to or read from when creating or using PROC TEMPLATE definitions and the order in which to search for them.
ODS PCL	Opens, manages, or closes the PCL destination, which produces printable output for PCL (HP LaserJet) files.
ODS PDF	Opens, manages, or closes the PDF destination, which produces PDF output, a form of output that is read by Adobe Acrobat and other applications.
ODS PHTML	Opens, manages, or closes the PHTML destination, which produces simple HTML output that uses 12 style elements and no class attributes.

Statement	Description
ODS PRINTER	Opens, manages, or closes the PRINTER destination, which produces printable output.
ODS PROCLABEL	Enables you to change a procedure label.
ODS PROCTITLE	Determines whether to write the title that identifies the procedure that produces the results in the output.
ODS PS	Opens, manages, or closes the PS destination, which produces PostScript (PS) output.
ODS RESULTS	Tracks ODS output in the Results window.
ODS RTF	Opens, manages, or closes the RTF destination, which produces output written in Rich Text Format for use with Microsoft Word 2002.
ODS SELECT	Specifies output objects for ODS destinations.
ODS SHOW	Writes the specified selection or exclusion list to the SAS log.
ODS TAGSETS.RTF	Opens, manages, or closes the RTF destination, which produces measured output that is written in Rich Text Format for use with Microsoft Word 2002.
ODS TEXT=	Inserts text into your ODS output.
ODS TRACE	Writes to the SAS log a record of each output object that is created, or else suppresses the writing of this record.
ODS USEGOPT	Determines whether ODS uses graphics option settings.
ODS VERIFY	Prints or suppresses a message indicating that a style definition or a table definition being used is not supplied by SAS.
ODS WML	Opens, manages, or closes the WML destination, which uses the Wireless Application Protocol (WAP) to produce a Wireless Markup Language (WML) DTD with a simple list for a table of contents.
PUT, ODS	Writes data values to a special buffer from which they can be written to the data component and then formatted by ODS.

SAS Scalable Performance Data Engine: Reference

Statement	Description
LIBNAME for the Scalable Performance Data Engine	Associates a SAS libref with a SAS library for rapid processing of very large data sets by multiple CPUs.

SAS XML LIBNAME Engine: User's Guide

Statement	Description
LIBNAME for the XML Engine	Associates a SAS libref with the physical location of an XML document.

SAS/ACCESS for Relational Databases: Reference

Statement	Description
LIBNAME for SAS/ ACCESS Relational Databases	Associates a SAS libref with a database management system (DBMS) database, schema, server, or group of tables or SAS views.

SAS/CONNECT User's Guide

Statement	Description
LIBNAME for SAS/ CONNECT Remote Library Services	Associates a libref with a SAS library that is located on the server for client access.
LIBNAME for SAS/ CONNECT TCP/IP Pipe	Associates a libref with a TCP/IP pipe (instead of a physical disk device) for processing input and output. The SASESOCK engine is required for SAS/CONNECT applications that implement MP CONNECT with piping.
RSUBMIT	Marks the beginning of a block of statements that a client session submits to a server session for execution.
SIGNON	Initiates a connection between a client session and a server session.

SAS/SHARE User's Guide

Statement	Description
LIBNAME for SAS/SHARE	In a client session, associates a libref with a SAS library that is located on the server for client access. In a server session, pre-defines a server library that clients are permitted to access.

CHAPTER 7

SAS System Options

Definition of System Options

System options are instructions that affect your SAS session. They control the way that SAS performs operations such as SAS System initialization, hardware and software interfacing, and the input, processing, and output of jobs and SAS files.

Syntax

Specifying System Options in an OPTIONS Statement

The syntax for specifying system options in an OPTIONS statement is

OPTIONS *option(s)*;

where

option
specifies one or more SAS system options that you want to change.

The following example shows how to use the system options NODATE and LINESIZE= in an OPTIONS statement:

```
options nodate linesize=72;
```

Operating Environment Information: On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Specifying Hexadecimal Values

Hexadecimal values for system options must begin with a number (0-9), followed by an X. For example, the following OPTIONS statement sets the line size to 160 using a hexadecimal number:

```
options linesize=0a0x;
```

Using SAS System Options

Default Settings

SAS system options are initialized with default settings when SAS is invoked. However, the default settings for some SAS system options vary both by operating environment and by site. Your on-site SAS support personnel might have created a default options table. The default options table is created by your site administrator and provides a global set of default values that are specific for your site. It reduces the need to duplicate options in every system configuration file at your site.

Information about creating and maintaining the default options table is provided in the configuration guide for SAS software for your operating environment.

Saving and Loading SAS System Options

SAS system options can be saved to either the SAS registry or a SAS data set by using the OPTSAVE procedure or by using the DMOPTSAVE command in the SAS windowing environment. Some system options cannot be saved. For a list of these options and additional information about saving options, see the OPTSAVE Procedure in *Base SAS Procedures Guide*.

To load a set of saved system options you use either the OPTLOAD procedure or the DMOPTLOAD command. For information about loading system option, see the OPTLOAD Procedure in *Base SAS Procedures Guide*.

For information about the DMOPTSAVE command and the DMOPTLOAD command, see the SAS Help and Documentation.

Determining Which Settings Are in Effect

To determine which settings are in effect for SAS system options, use one of the following:

OPLIST system option
: Writes to the SAS log the system options that were specified on the SAS invocation command line. (See the SAS documentation for your operating environment for more information.)

VERBOSE system option
Writes to the SAS log the system options that were specified in the configuration file and on the SAS invocation command line.

SAS System Options window
Lists all system option settings.

OPTIONS procedure
Writes system option settings to the SAS log. To display the settings of system options with a specific functionality, such as error handling, use the GROUP= option:

```
proc options GROUP=errorhandling;
run;
```

(See the *Base SAS Procedures Guide* for more information.)

GETOPTION function
Returns the value of a specified system option.

VOPTION Dictionary table
Located in the SASHELP library, VOPTION contains a list of all current system option settings. You can view this table with SAS Explorer, or you can extract information from the VOPTION table using PROC SQL.

dictionary.options SQL table
Accessed with the SQL procedure, this table lists the system options that are in effect.

Restricted Options

Restricted options are system options whose value are determined by the site administrator and cannot be overridden. The site administrator can create a restricted options table that specifies the option values that are restricted when SAS starts. Any attempt to modify a system option that is listed in the restricted options table is either ignored, or if you use the OPTIONS statement to set a restricted option, SAS issues a warning message to the log.

To determine which system options are restricted by your site administrator, use the RESTRICT option of the OPTIONS procedure. The RESTRICT option displays the option's value, scope, and setting. In the following example, the SAS log shows that only one option, CMPOPT, is restricted:

```
proc options restrict;
run;
```

Output 7.1 Restricted Option Information

```
1    proc options restrict;
2    run;
    SAS (r) Proprietary Software Release xxx  TS1B0

Option Value Information For SAS Option CMPOPT
    Option Value: (NOEXTRAMATH NOMISSCHECK NOPRECISE NOGUARDCHECK NOFUNCDIFFERENCING)
    Option Scope: SAS Session
    How option value set:  Site Administrator Restricted
```

The OPTIONS procedure displays this information for all options that are restricted. If your site administrator has not restricted any options, then the following message appears in the SAS log:

```
Your site administrator has not restricted any options.
```

The following table lists the system options that cannot be restricted:

Table 7.1 System Options That Cannot Be Restricted

Option	All Operating Environments	OpenVMS	UNIX	Windows	z/OS
ALTLOG	X				
ALTPRINT	X				
ASYNCHIO					X
AUTOEXEC	X				
BOMFILE	X				
BOTTOMMARGIN	X				
COMDEF			X	X	
CONFIG	X				
CPUCOUNT	X				
DATESTYLE	X				
DBCS			X		
DFLANG	X				
DLDMGACTION	X				
DMR	X				
DMS	X				
DMSEXP	X				
DMSPGMLINESIZE	X				
ENGINE	X				
EXPLORER	X				
FILELOCKWAITMAX			X	X	
INITCMD	X				
INITSTMT	X				
JREOPTIONS					X
LEFTMARGIN	X				
LINESIZE	X				
LAST	X				
LOG	X				
LOGAPPLNAME	X				
LOGPARM	X				

Option	All Operating Environments	OpenVMS	UNIX	Windows	z/OS
MEMCACHE				X	
MEMLIB				X	
METAPASS	X				
METAPROTOCOL	X				
METAREPOSITORY	X				
METASERVER	X				
METAUSER	X				
MSYMTABMAX	X				
MVARSIZE	X				
OBJECTSERVER	X				
ORIENTATION	X				
OVP	X				
PAGESIZE	X				
PAPERSIZE	X				
PATH				X	
PDFPASSWORD	X				
PRINT	X				
PRINTERPATH	X				
RESOURCESLOC				X	
RIGHTMARGIN	X				
SASCONTROL				X	
SASFRSCR	X				
SASUSER	X				
SGIO				X	
SOURCE	X				
SSLPKCS12LOC			X		
SSLPKSC12PASS			X		
SSPI	X				
STARTLIB	X				
SYSIN	X				
SYSPRINTFONT	X				
TERMINAL	X				
TOOLDEF				X	
TOPMARGIN	X				
TRANTAB	X				

Option	All Operating Environments	OpenVMS	UNIX	Windows	z/OS
USER	X				
WORK	X				

Determining How a SAS System Option Value Was Set

To determine how a system option value was set, use the OPTIONS procedure with the VALUE option specified in the OPTIONS statement. The VALUE option displays the specified option's value and scope. For example, the following statements write a message to the SAS log that tells you how the option value for the system option CENTER was set:

```
proc options option=center value;
run;
```

The following partial SAS log shows that the option value for CENTER was the shipped default.

Output 7.2 Option Value Information for the System Option CENTER

```
2    proc options option=center value;
3    run;


Option Value Information for SAS Option CENTER
    Option Value: CENTER
    Option Scope: Default
    How option value set:  Shipped Default
```

If no value is assigned to a character system option, then SAS assigns the option a value of `''`(a space between two single quotation marks) and **`Option Value`** displays a blank space.

Obtaining Descriptive Information about a System Option

You can quickly obtain basic descriptive information about a system option by specifying the DEFINE option in the PROC OPTIONS statement.

The DEFINE option writes the following descriptive information about a system option to the SAS log:

- a description of the option
- the name and description of each System option group that the option is a part of
- type information
- when in the SAS session it can be set
- if it can be restricted by the system administrator
- if the OPTSAVE procedure or the DMOPTSAVE command will save the option.

For example, the following statements write a message to the SAS log that contains descriptive information about the system option CENTER:

```
proc options option=center define;
run;
```

Output 7.3 Descriptive Information for the System Option CENTER

```
1    proc options option=center define;
      2   run;
 CENTER
 Option Definition Information for SAS Option CENTER
    Group= LISTCONTROL
    Group Description: Procedure output and display settings
    Description: Center SAS procedure output
    Type: The option value is of type BOOLEAN
    When Can Set: Startup or anytime during the SAS Session
    Restricted: Your Site Administrator can restrict modification of this option
    Optsave: Proc Optsave or command Dmoptsave will save this option.
```

Changing SAS System Option Settings

SAS provides default settings for SAS system options. You can override the default settings of any unrestricted system option in several ways, depending on the function of the system option:

- *On the command line or in a configuration file:* You can specify any unrestricted SAS system option setting either on the SAS command line or in a configuration file. If you use the same option settings frequently, then it is usually more convenient to specify the options in a configuration file, rather than on the command line. Either method sets your SAS system options during SAS invocation. Many SAS system option settings can be specified only during SAS invocation. Descriptions of the individual options provide details.
- *In an OPTIONS statement:* You can specify an OPTIONS statement at any time during a session except within data lines or parmcard lines. Settings remain in effect throughout the current program or process unless you reset them with another OPTIONS statement or change them in the SAS System Options window. You can also place an OPTIONS statement in an autoexec file.
- *In the OPTLOAD procedure or the DMOPTLOAD command:* You can use the OPTLOAD procedure or the DMOPTLOAD command to read option settings that were specified with the OPTSAVE procedure and saved to a SAS data set.
- *In a SAS System Options window:* If you are in a windowing environment, type **options** in the toolbar or on the command line to open the SAS System Options window. The SAS System Options window lists the names of the SAS system option groups. You can then expand the groups to see the option names and to change their current settings. Alternatively, you can use the Find Option command in the Options pop-up menu to go directly to an option. Changes take effect immediately and remain in effect throughout the session unless you reset them with an OPTIONS statement or change them in the SAS System Options window.

SAS system options can be restricted by a site administrator so that after they are set by the administrator, they cannot be changed by a user. Depending upon your operating environment, system options can be restricted globally, by group, or by user. You can use the OPTIONS procedure to determine which options are restricted. For

more information about how to restrict options, see the SAS configuration guide for your operating environment. For more information about the OPTIONS procedure, see "The Options Procedure" in *Base SAS Procedures Guide* and the SAS documentation for your operating environment.

How Long System Option Settings Are in Effect

When you specify a SAS system option setting, the setting applies to the next step and to *all subsequent steps* for the duration of the SAS session, or until you reset the system option setting, as shown:

```
data one;
   set items;
run;

   /* option applies to all subsequent steps */
options obs=5;

  /* printing ends with the fifth observation */
proc print data=one;
run;

   /* the SET statement stops reading
      after the fifth observation */
data two;
   set items;
run;
```

To read more than five observations, you must reset the OBS= system option. For more information, see "OBS= System Option" on page 1948.

Order of Precedence

If the same system option appears in more than one place, the order of precedence from highest to lowest is

1. restricted options table, if it exists
2. OPTIONS statement and SAS System Options window
3. autoexec file (that contains an OPTIONS statement)
4. command-line specification
5. configuration file specification
6. SAS system default settings.

Operating Environment Information: In some operating environments, you can specify system options in other places. See the SAS documentation for your operating environment. △

The following table shows the order of precedence that SAS uses for execution mode options. These options are a subset of the SAS invocation options and are specified on the command line during SAS invocation.

Table 7.2 Order of Precedence for SAS Execution Mode Options

Execution Mode Option	Precedence
OBJECTSERVER	Highest
DMR	2nd
SYSIN	3rd
INITCMD	4th
DMS	4th
DMSEXP	4th
EXPLORER	4th
none (default is interactive line mode under UNIX and OpenVMS; interactive full screen mode under z/OS)	5th

The order of precedence of SAS execution mode options consists of the following rules:

- SAS uses the execution mode option with the highest precedence.
- If you specify more than one execution mode option of equal precedence, SAS uses only the last option listed.

See the descriptions of the individual options for more details.

Interaction with Data Set Options

Many system options and data set options share the same name and have the same function. System options remain in effect for all DATA and PROC steps in a SAS job or session until their settings are changed. A data set option, however, overrides a system option only for the particular data set in the step in which it appears.

In this example, the OBS= system option in the OPTIONS statement specifies that only the first 100 observations will be read from any data set within the SAS job. The OBS= data set option in the SET statement, however, overrides the system option and specifies that only the first five observations will be read from data set TWO. The PROC PRINT step uses the system option setting and reads and prints the first 100 observations from data set THREE:

```
options obs=100;

data one;
   set two(obs=5);
run;

proc print data=three;
run;
```

Comparisons

Note the differences between system options, data set options, and statement options.

system options
: remain in effect for all DATA and PROC steps in a SAS job or current process unless they are respecified.

data set options
: apply to the processing of the SAS data set with which they appear. Some data set options have corresponding system options or LIBNAME statement options. For an individual data set, you can use the data set option to override the setting of these other options.

statement options
: control the action of the statement in which they appear. Options in global statements, such as in the LIBNAME statement, can have a broader impact.

SAS System Options by Category

Table 7.3 Categories and Descriptions of SAS System Options

Category	SAS System Options	Description
Communications: Email	"EMAILAUTHPROTOCOL= System Option" on page 1895	Specifies the authentication protocol for SMTP E-mail.
	"EMAILFROM System Option" on page 1896	When sending e-mail by using SMTP, specifies whether the e-mail option FROM is required in either the FILE or FILENAME statement.
	"EMAILHOST= System Option" on page 1897	Specifies one or more SMTP servers that support e-mail access.
	"EMAILID= System Option" on page 1898	Identifies an e-mail sender by specifying either a logon ID, an e-mail profile, or an e-mail address.
	"EMAILPORT System Option" on page 1899	Specifies the port that the SMTP server is attached to.
	"EMAILPW= System Option" on page 1900	Specifies an e-mail logon password.
Communications: Networking and encryption	"HTTPSERVERPORTMAX= System Option" on page 1922	Specifies the highest port number that can be used by the SAS HTTP server for remote browsing.
	"HTTPSERVERPORTMIN= System Option" on page 1923	Specifies the lowest port number that can be used by the SAS HTTP server for remote browsing.
Environment control: Display	"AUTOSAVELOC= System Option" on page 1848	Specifies the location of the Program Editor autosave file.
	"CHARCODE System Option" on page 1863	Specifies whether specific keyboard combinations are substituted for special characters that are not on the keyboard.

Category	SAS System Options	Description
	"DMSLOGSIZE= System Option" on page 1888	Specifies the maximum number of rows that the SAS Log window can display.
	"DMSOUTSIZE= System Option" on page 1889	Specifies the maximum number of rows that the SAS Output window can display.
	"DMSPGMLINESIZE= System Option" on page 1890	Specifies the maximum number of characters in a Program Editor line.
	"FONTSLOC= System Option" on page 1914	Specifies the location of the fonts that are supplied by SAS; names the default font file location for registering fonts that use the FONTREG procedure.
	"FORMS= System Option" on page 1916	If forms are used for printing, specifies the default form to use.
	"SOLUTIONS System Option" on page 1998	Specifies whether the SOLUTIONS menu is included in SAS windows.
	"TOOLSMENU System Option" on page 2038	Specifies whether the Tools menu is included in SAS windows.
	"VIEWMENU System Option" on page 2053	Specifies whether the View menu is included in SAS windows.
Environment control: Error handling	"BYERR System Option" on page 1855	Specifies whether SAS produces errors when the SORT procedure attempts to process a _NULL_ data set.
	"CLEANUP System Option" on page 1864	For an out-of-resource condition, specifies whether to perform an automatic cleanup or a user-specified cleanup.
	"DMSSYNCHK System Option" on page 1890	In the SAS windowing environment, specifies whether to enable syntax check mode for DATA step and PROC step processing.
	"DSNFERR System Option" on page 1892	When a SAS data set cannot be found, specifies whether SAS issues an error message.
	"ERRORABEND System Option" on page 1902	Specifies whether SAS responds to errors by terminating.
	"ERRORBYABEND System Option" on page 1903	Specifies whether SAS ends a program when an error occurs in BY-group processing.
	"ERRORCHECK= System Option" on page 1904	Specifies whether SAS enters syntax-check mode when errors are found in the LIBNAME, FILENAME, %INCLUDE, and LOCK statements.
	"ERRORS= System Option" on page 1905	Specifies the maximum number of observations for which SAS issues complete error messages.
	"FMTERR System Option" on page 1910	When a variable format cannot be found, specifies whether SAS generates an error or continues processing.
	"QUOTELENMAX System Option" on page 1981	If a quoted string exceeds the maximum length allowed, specifies whether SAS writes a warning message to the SAS log.
	"STEPCHKPT System Option" on page 2013	Specifies whether checkpoint-restart data is to be recorded for a batch program.

Category	SAS System Options	Description
	"STEPCHKPTLIB= System Option" on page 2014	Specifies the libref of the library where checkpoint-restart data is saved.
	"STEPRESTART System Option" on page 2016	Specifies whether to execute a batch program by using checkpoint-restart data.
	"SYNTAXCHECK System Option" on page 2031	In non-interactive or batch SAS sessions, specifies whether to enable syntax check mode for multiple steps.
	"VNFERR System Option" on page 2054	Specifies whether SAS issues an error or warning when a BY variable exists in one data set but not another data set when processing the SET, MERGE, UPDATE, or MODIFY statements.
Environment control: Files	"APPEND= System Option" on page 1844	Appends a value to the existing value of the specified system option.
	"APPLETLOC= System Option" on page 1845	Specifies the location of Java applets.
	"FMTSEARCH= System Option" on page 1910	Specifies the order in which format catalogs are searched.
	"HELPENCMD System Option" on page 1920	Specifies whether SAS uses the English version or the translated version of the keyword list for the command–line Help.
	"INSERT= System Option" on page 1929	Inserts the specified value as the first value of the specified system option.
	"NEWS= System Option" on page 1946	Specifies an external file that contains messages to be written to the SAS log, immediately after the header.
	"PARM= System Option" on page 1963	Specifies a parameter string that is passed to an external program.
	"PARMCARDS= System Option" on page 1963	Specifies the file reference to open when SAS encounters the PARMCARDS statement in a procedure.
	"RSASUSER System Option" on page 1986	Specifies whether to open the SASUSER library for read access or read-write access.
	"SASHELP= System Option" on page 1994	Specifies the location of the SASHELP library.
	"SASUSER= System Option" on page 1995	Specifies the SAS library to use as the SASUSER library.
	"TRAINLOC= System Option" on page 2040	Specifies the URL for SAS online training courses.
	"USER= System Option" on page 2042	Specifies the default permanent SAS library.
	"UUIDCOUNT= System Option" on page 2044	Specifies the number of UUIDs to acquire from the UUID Generator Daemon.
	"UUIDGENDHOST= System Option" on page 2045	Identifies the host and port or the LDAP URL that the UUID Generator Daemon runs on.

Category	SAS System Options	Description
	"V6CREATEUPDATE= System Option" on page 2047	Specifies the type of message to write to the SAS log when Version 6 data sets are created or updated.
	"WORK= System Option" on page 2055	Specifies the WORK data library.
	"WORKINIT System Option" on page 2056	Specifies whether to initialize the WORK library at SAS invocation.
	"WORKTERM System Option" on page 2057	Specifies whether to erase the WORK files when SAS terminates.
Environment control: Help	"HELPBROWSER= System Option" on page 1919	Specifies the browser to use for SAS Help and ODS output.
	"HELPHOST System Option" on page 1921	Specifies the name of the computer where the remote browser is to send Help and ODS output.
	"HELPPORT= System Option" on page 1922	Specifies the port number for the remote browser client.
Environment control: Initialization and operation	"AUTHPROVIDERDOMAIN System Option" on page 1846	Associates a domain suffix with an authentication provider.
	"DMR System Option" on page 1885	Specifies whether to enable SAS to invoke a server session for use with a SAS/CONNECT client.
	"DMS System Option" on page 1886	Specifies whether to invoke the SAS windowing environment and display the Log, Editor, and Output windows.
	"DMSEXP System Option" on page 1887	Specifies whether to invoke the SAS windowing environment and display the Explorer, Editor, Log, Output, and Results windows.
	"EXPLORER System Option" on page 1906	Specifies whether to invoke the SAS windowing environment and display only the Explorer window.
	"INITCMD System Option" on page 1926	Specifies an application invocation command and optional SAS windowing environment or text editor commands that SAS executes before processing AUTOEXEC file during SAS invocation.
	"INITSTMT= System Option" on page 1928	Specifies a SAS statement to execute after any statements in the autoexec file and before any statements from the SYSIN= file.
	"MULTENVAPPL System Option" on page 1946	Specifies whether the fonts available in a SAS application font selector window lists only the SAS fonts that are available in all operating environments.
	"PRIMARYPROVIDERDOMAIN= System Option" on page 1976	Specifies the domain name of the primary authentication provider.
	"TERMINAL System Option" on page 2035	Specifies whether to associate a terminal with a SAS session.

Category	SAS System Options	Description
	"TERMSTMT= System Option" on page 2035	Specifies the SAS statements to execute when SAS terminates.
Environment control: Language control	"DATESTYLE= System Option" on page 1879	Specifies the sequence of month, day, and year when ANYDTDTE, ANYDTDTM, or ANYDTTME informat data is ambiguous.
	"PAPERSIZE= System Option" on page 1960	Specifies the paper size to use for printing.
Files: External files	"LRECL= System Option" on page 1941	Specifies the default logical record length to use for reading and writing external files.
	"STARTLIB System Option" on page 2013	Specifies whether SAS assigns user-defined permanent librefs when SAS starts.
Files: SAS Files	"BUFNO= System Option" on page 1851	Specifies the number of buffers to be allocated for processing SAS data sets.
	"BUFSIZE= System Option" on page 1853	Specifies the permanent buffer page size for output SAS data sets.
	"CATCACHE= System Option" on page 1860	Specifies the number of SAS catalogs to keep open in cache memory.
	"CBUFNO= System Option" on page 1861	Specifies the number of extra page buffers to allocate for each open SAS catalog.
	"CMPLIB= System Option" on page 1866	Specifies one or more SAS data sets that contain compiler subroutines to include during program compilation.
	"COMPRESS= System Option" on page 1872	Specifies the type of compression of observations to use for output SAS data sets.
	"DATASTMTCHK= System Option" on page 1877	Specifies which SAS statement keywords are prohibited from being specified as a one-level DATA step name to protect against overwriting an input data set.
	"DKRICOND= System Option" on page 1883	Specifies the level of error detection to report when a variable is missing from an input data set during the processing of a DROP=, KEEP=, or RENAME= data set option.
	"DKROCOND= System Option" on page 1883	Specifies the level of error detection to report when a variable is missing for an output data set during the processing of a DROP=, KEEP=, or RENAME= data set option.
	"DLDMGACTION= System Option" on page 1884	Specifies the type of action to take when a SAS data set or a SAS catalog is detected as damaged.
	"ENGINE= System Option" on page 1901	Specifies the default access method for SAS libraries.
	"FILESYNC= System Option" on page 1907	Specifies when operating system buffers that contain contents of permanent SAS files are written to disk.
	"FIRSTOBS= System Option" on page 1908	Specifies the observation number or external file record that SAS processes first.
	"IBUFNO= System Option" on page 1924	Specifies an optional number of extra buffers to be allocated for navigating an index file.

Category	SAS System Options	Description
	"IBUFSIZE= System Option" on page 1925	Specifies the buffer page size for an index file.
	"_LAST_= System Option" on page 1934	Specifies the most recently created data set.
	"MERGENOBY System Option" on page 1943	Specifies the type of message that is issued when MERGE processing occurs without an associated BY statement.
	"OBS= System Option" on page 1948	Specifies the observation that is used to determine the last observation to process, or specifies the last record to process.
	"REPLACE System Option" on page 1982	Specifies whether permanently stored SAS data sets can be replaced.
	"REUSE= System Option" on page 1983	Specifies whether SAS reuses space when observations are added to a compressed SAS data set.
	"SQLCONSTDATETIME System Option" on page 2005	Specifies whether the SQL procedure replaces references to the DATE, TIME, DATETIME, and TODAY functions in a query with their equivalent constant values before the query executes.
	"SQLREDUCEPUT= System Option" on page 2006	For the SQL procedure, specifies the engine type that a query uses for which optimization is performed by replacing a PUT function in a query with a logically equivalent expression.
	"SQLREDUCEPUTOBS= System Option" on page 2008	For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the minimum number of observations that must be in a table in order for PROC SQL to consider optimizing the PUT function in a query.
	"SQLREDUCEPUTVALUES= System Option" on page 2009	For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the maximum number of SAS format values that can exist in a PUT function expression in order for PROC SQL to consider optimizing the PUT function in a query.
	"SQLREMERGE System Option" on page 2010	Specifies whether the SQL procedure can process queries that use remerging of data.
	"SQLUNDOPOLICY= System Option" on page 2011	Specifies whether the SQL procedure keeps or discards updated data if errors occur while the data is being updated.
	"UTILLOC= System Option" on page 2043	Specifies one or more file system locations in which applications can store utility files.
	"VALIDFMTNAME= System Option" on page 2048	Specifies the maximum size (32 characters or 8 characters) that user-created format and informat names can be before an error or warning is issued.
	"VALIDVARNAME= System Option" on page 2049	Specifies the rules for valid SAS variable names that can be created and processed during a SAS session.

Category	SAS System Options	Description
	"VARLENCHK= System Option" on page 2050	Specifies the type of message to write to the SAS log when the input data set is read using the SET, MERGE, UPDATE, or MODIFY statements.
Graphics: Driver settings	"DEVICE= System Option" on page 1882	Specifies the device driver to which SAS/GRAPH sends procedure output.
	"GSTYLE System Option" on page 1917	Specifies whether ODS styles can be used in the generation of graphs that are stored as GRSEG catalog entries.
	"GWINDOW System Option" on page 1918	Specifies whether SAS displays SAS/GRAPH output in the GRAPH window.
	"MAPS= System Option" on page 1942	Specifies the location of the SAS library that contains SAS/GRAPH map data sets.
Input control: Data Processing	"BYSORTED System Option" on page 1857	Specifies whether observations in one or more data sets are sorted in alphabetic or numeric order or are grouped in another logical order.
	"CAPS System Option" on page 1858	Specifies whether to convert certain types of input to uppercase.
	"CARDIMAGE System Option" on page 1859	Specifies whether SAS processes source and data lines as 80-byte cards.
	"DATESTYLE= System Option" on page 1879	Specifies the sequence of month, day, and year when ANYDTDTE, ANYDTDTM, or ANYDTTME informat data is ambiguous.
	"INVALIDDATA= System Option" on page 1931	Specifies the value that SAS assigns to a variable when invalid numeric data is encountered.
	"S= System Option" on page 1987	Specifies the length of statements on each line of a source statement and the length of data on lines that follow a DATALINES statement.
	"S2= System Option" on page 1990	Specifies the length of statements on each line of a source statement from a %INCLUDE statement, an autoexec file, or an autocall macro file.
	"S2V= System Option" on page 1993	Specifies the starting position to begin reading a file that is specified in a %INCLUDE statement, an autoexec file, or an autocall macro file with a variable length record format.
	"SEQ= System Option" on page 1996	Specifies the length of the numeric portion of the sequence field in input source lines or data lines.
	"SPOOL System Option" on page 2004	Specifies whether SAS statements are written to a utility data set in the WORK data library.
	"YEARCUTOFF= System Option" on page 2058	Specifies the first year of a 100-year span that is used by date informats and functions to read a two–digit year.
Input control: Data processing	"INTERVALDS= System Option" on page 1930	Specifies one or more interval name and value pairs, where the value is a SAS data set that contains user-supplied holidays. The interval can be used as an argument to the INTNX and INTCK functions.

Category	SAS System Options	Description
Log and procedure output control: ODS Printing	"BINDING= System Option" on page 1849	Specifies the binding edge for duplexed printed output.
	"BOTTOMMARGIN= System Option" on page 1850	Specifies the size of the margin at the bottom of a printed page.
	"COLLATE System Option" on page 1871	Specifies whether to collate multiple copies of printed output.
	"COLORPRINTING System Option" on page 1872	Specifies whether to print in color if color printing is supported.
	"COPIES= System Option" on page 1874	Specifies the number of copies to print.
	"DEFLATION= System Option" on page 1880	Specifies the level of compression for device drivers that support the Deflate compression algorithm.
	"DUPLEX System Option" on page 1894	Specifies whether duplex (two-sided) printing is enabled.
	"FONTEMBEDDING System Option" on page 1912	Specifies whether font embedding is enabled in Universal Printer and SAS/GRAPH printing.
	"FONTRENDERING= System Option" on page 1913	Specifies whether SAS/GRAPH devices that are based on the SASGDGIF, SASGDTIF, and SASGDIMG modules render fonts by using the operating system or by using the FreeType engine.
	"GSTYLE System Option" on page 1917	Specifies whether ODS styles can be used in the generation of graphs that are stored as GRSEG catalog entries.
	"JPEGQUALITY= System Option" on page 1932	Specifies the JPEG quality factor that determines the ratio of image quality to the level of compression for JPEG files produced by the SAS/GRAPH JPEG device driver.
	"LEFTMARGIN= System Option" on page 1934	Specifies the print margin for the left side of the page.
	"ORIENTATION= System Option" on page 1954	Specifies the paper orientation to use when printing to a printer.
	"PAPERDEST= System Option" on page 1959	Specifies the name of the output bin to receive printed output.
	"PAPERSIZE= System Option" on page 1960	Specifies the paper size to use for printing.
	"PAPERSOURCE= System Option" on page 1961	Specifies the name of the paper bin to use for printing.
	"PAPERTYPE= System Option" on page 1962	Specifies the type of paper to use for printing.
	"PRINTERPATH= System Option" on page 1978	Specifies the name of a registered printer to use for Universal Printing.

Category	SAS System Options	Description
	"RIGHTMARGIN= System Option" on page 1984	Specifies the print margin for the right side of the page for output directed to an ODS printer destination.
	"TEXTURELOC= System Option" on page 2036	Specifies the location of textures and images that are used by ODS styles.
	"TOPMARGIN= System Option" on page 2039	Specifies the print margin at the top of the page for output directed to an ODS printer destination.
	"UNIVERSALPRINT System Option" on page 2040	Specifies whether to enable Universal Printing services.
	"UPRINTCOMPRESSION System Option" on page 2041	Specifies whether to enable compression of file created by some Universal Printer and SAS/GRAPH devices.
Log and procedure output control: PDF	"PDFACCESS System Option" on page 1964	Specifies whether text and graphics from PDF documents can be read by screen readers for the visually impaired.
	"PDFASSEMBLY System Option" on page 1965	Specifies whether PDF documents can be assembled.
	"PDFCOMMENT System Option" on page 1966	Specifies whether PDF document comments can be modified.
	"PDFCONTENT System Option" on page 1967	Specifies whether the contents of a PDF document can be changed.
	"PDFCOPY System Option" on page 1968	Specifies whether text and graphics from a PDF document can be copied.
	"PDFFILLIN System Option" on page 1970	Specifies whether PDF forms can be filled in.
	" PDFPAGELAYOUT= System Option" on page 1971	Specifies the page layout for PDF documents.
	" PDFPAGEVIEW= System Option" on page 1972	Specifies the page viewing mode for PDF documents.
	"PDFPASSWORD= System Option" on page 1972	Specifies the password to use to open a PDF document and the password used by a PDF document owner.
	"PDFPRINT= System Option" on page 1974	Specifies the resolution to print PDF documents.
	"PDFSECURITY= System Option" on page 1975	Specifies the printing permissions for PDF documents.
Log and procedure output control: Procedure output	"BYLINE System Option" on page 1856	Specifies whether to print BY lines above each BY group.
	"CENTER System Option" on page 1862	Specifies whether to center or left align SAS procedure output.
	"FORMCHAR= System Option" on page 1914	Specifies the default output formatting characters.
	"FORMDLIM= System Option" on page 1916	Specifies a character to delimit page breaks in SAS output.

Category	SAS System Options	Description
	"LABEL System Option" on page 1933	Specifies whether SAS procedures can use labels with variables.
	"PAGENO= System Option" on page 1957	Resets the SAS output page number.
	"PRINTINIT System Option" on page 1979	Specifies whether to initialize the SAS procedure output file.
	"SKIP= System Option" on page 1997	Specifies the number of lines to skip at the top of each page of SAS output.
	"SYSPRINTFONT= System Option" on page 2032	Specifies the default font to use for printing, which can be overridden by explicitly specifying a font and an ODS style.
Log and procedure output control: Procedure output	"DATE System Option" on page 1878	Specifies whether to print the date and time that a SAS program started.
	"DETAILS System Option" on page 1881	Specifies whether to include additional information when files are listed in a SAS library.
	"DTRESET System Option" on page 1893	Specifies whether to update the date and time in the SAS log and in the procedure output file.
	"LINESIZE= System Option" on page 1936	Specifies the line size for the SAS log and for SAS procedure output.
	"MISSING= System Option" on page 1944	Specifies the character to print for missing numeric values.
	"NUMBER System Option" on page 1948	Specified whether to print the page number in the title line of each page of SAS output.
	"PAGEBREAKINITIAL System Option" on page 1956	Specifies whether to begin the SAS log and procedure output files on a new page.
	"PAGESIZE= System Option" on page 1958	Specifies the number of lines that compose a page of SAS output.
Log and procedure output control: SAS log and procedure output	"DATE System Option" on page 1878	Specifies whether to print the date and time that a SAS program started.
	"DETAILS System Option" on page 1881	Specifies whether to include additional information when files are listed in a SAS library.
	"DTRESET System Option" on page 1893	Specifies whether to update the date and time in the SAS log and in the procedure output file.
	"LINESIZE= System Option" on page 1936	Specifies the line size for the SAS log and for SAS procedure output.
	"MISSING= System Option" on page 1944	Specifies the character to print for missing numeric values.
	"NUMBER System Option" on page 1948	Specified whether to print the page number in the title line of each page of SAS output.
	"PAGEBREAKINITIAL System Option" on page 1956	Specifies whether to begin the SAS log and procedure output files on a new page.

Category	SAS System Options	Description
	"PAGESIZE= System Option" on page 1958	Specifies the number of lines that compose a page of SAS output.
Log and procedure output control: SAS log	"CPUID System Option" on page 1877	Specifies whether the CPU identification number is written to the SAS log.
	"DATE System Option" on page 1878	Specifies whether to print the date and time that a SAS program started.
	"DETAILS System Option" on page 1881	Specifies whether to include additional information when files are listed in a SAS library.
	"DMSLOGSIZE= System Option" on page 1888	Specifies the maximum number of rows that the SAS Log window can display.
	"DTRESET System Option" on page 1893	Specifies whether to update the date and time in the SAS log and in the procedure output file.
	"ECHOAUTO System Option" on page 1895	Specifies whether the statements in the autoexec file are written to the SAS log as they are executed.
	"ERRORS= System Option" on page 1905	Specifies the maximum number of observations for which SAS issues complete error messages.
	"LINESIZE= System Option" on page 1936	Specifies the line size for the SAS log and for SAS procedure output.
	"LOGPARM= System Option" on page 1937	Specifies when SAS log files are opened, closed, and, in conjunction with the LOG= system option, how they are named.
	"MISSING= System Option" on page 1944	Specifies the character to print for missing numeric values.
	"MSGLEVEL= System Option" on page 1945	Specifies the level of detail in messages that are written to the SAS log.
	"NEWS= System Option" on page 1946	Specifies an external file that contains messages to be written to the SAS log, immediately after the header.
	"NOTES System Option" on page 1947	Specifies whether notes are written to the SAS log.
	"NUMBER System Option" on page 1948	Specified whether to print the page number in the title line of each page of SAS output.
	"OVP System Option" on page 1955	Specifies whether overprinting of error messages to make them bold, is enabled.
	"PAGEBREAKINITIAL System Option" on page 1956	Specifies whether to begin the SAS log and procedure output files on a new page.
	"PAGESIZE= System Option" on page 1958	Specifies the number of lines that compose a page of SAS output.
	"PRINTMSGLIST System Option" on page 1980	Specifies whether to print all messages to the SAS log or to print only top-level messages to the SAS log.
	"SOURCE System Option" on page 2003	Specifies whether SAS writes source statements to the SAS log.
	"SOURCE2 System Option" on page 2004	Specifies whether SAS writes secondary source statements from included files to the SAS log.

Category	SAS System Options	Description
Log and procedure output control: SVG	"SVGCONTROLBUTTONS" on page 2018	Specifies whether to display the paging control buttons and an index in a multipage SVG document.
	"SVGHEIGHT= System Option" on page 2019	Specifies the height of the viewport unless the SVG output is embedded in another SVG output; specifies the value of the height attribute of the outermost <svg> element in the SVG file.
	"SVGPRESERVEASPECTRATIO= System Option" on page 2021	Specifies whether to force uniform scaling of SVG output; specifies the preserveAspectRatio attribute on the outermost <svg> element.
	"SVGTITLE= System Option" on page 2023	Specifies the title in the title bar of the SVG output; specifies the value of the <title> element in the SVG file.
	"SVGVIEWBOX= System Option" on page 2024	Specifies the coordinates, width, and height that are used to set the viewBox attribute on the outermost <svg> element, which enables SVG output to scale to the viewport.
	"SVGWIDTH= System Option" on page 2026	Specifies the width of the viewport unless the SVG output is embedded in another SVG output; specifies the value of the width attribute in the outermost <svg> element in the SVG file.
	"SVGX= System Option" on page 2028	Specifies the x-axis coordinate of one corner of the rectangular region into which an embedded <svg> element is placed; specifies the x attribute in the outermost <svg> element in an SVG file.
	"SVGY= System Option" on page 2029	Specifies the y-axis coordinate of one corner of the rectangular region into which an embedded <svg> element is placed; specifies the y attribute in the outermost <svg> element in an SVG file.
Sort: Procedure options	"SORTDUP= System Option" on page 1998	Specifies whether the SORT procedure removes duplicate variables based on all variables in a data set or the variables that remain after the DROP or KEEP data set options have been applied.
	"SORTEQUALS System Option" on page 1999	Specifies whether observations in the output data set with identical BY variable values have a particular order.
	"SORTSIZE= System Option" on page 2000	Specifies the amount of memory that is available to the SORT procedure.
	"SORTVALIDATE System Option" on page 2002	Specifies whether the SORT procedure verifies if a data set is sorted according to the variables in the BY statement when a user-specified sort order is denoted in the sort indicator.
System administration: Code generation	"CGOPTIMIZE= System Option" on page 1862	Specifies the level of optimization to perform during code compilation.
System administration: Installation	"SETINIT System Option" on page 1996	Specifies whether site license information can be altered.
System administration: Memory	"SORTSIZE= System Option" on page 2000	Specifies the amount of memory that is available to the SORT procedure.

Category	SAS System Options	Description
	"SUMSIZE= System Option" on page 2017	Specifies a limit on the amount of memory that is available for data summarization procedures when class variables are active.
System administration: Performance	"BUFNO= System Option" on page 1851	Specifies the number of buffers to be allocated for processing SAS data sets.
	"BUFSIZE= System Option" on page 1853	Specifies the permanent buffer page size for output SAS data sets.
	"CGOPTIMIZE= System Option" on page 1862	Specifies the level of optimization to perform during code compilation.
	"CMPMODEL= System Option" on page 1868	Specifies the output model type for the MODEL procedure.
	"CMPOPT= System Option" on page 1868	Specifies the type of code generation optimizations to use in the SAS language compiler.
	"COMPRESS= System Option" on page 1872	Specifies the type of compression of observations to use for output SAS data sets.
	"CPUCOUNT= System Option" on page 1875	Specifies the number of processors that the thread-enabled applications should assume will be available for concurrent processing.
	"SQLREDUCEPUT= System Option" on page 2006	For the SQL procedure, specifies the engine type that a query uses for which optimization is performed by replacing a PUT function in a query with a logically equivalent expression.
	"SQLREDUCEPUTOBS= System Option" on page 2008	For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the minimum number of observations that must be in a table in order for PROC SQL to consider optimizing the PUT function in a query.
	"SQLREDUCEPUTVALUES= System Option" on page 2009	For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the maximum number of SAS format values that can exist in a PUT function expression in order for PROC SQL to consider optimizing the PUT function in a query.
	"THREADS System Option" on page 2037	Specifies that SAS use threaded processing if it is available.
System administration: SQL	"SQLCONSTDATETIME System Option" on page 2005	Specifies whether the SQL procedure replaces references to the DATE, TIME, DATETIME, and TODAY functions in a query with their equivalent constant values before the query executes.
	"SQLREDUCEPUT= System Option" on page 2006	For the SQL procedure, specifies the engine type that a query uses for which optimization is performed by replacing a PUT function in a query with a logically equivalent expression.
	"SQLREDUCEPUTOBS= System Option" on page 2008	For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the minimum number of observations that must be in a table in order for PROC SQL to consider optimizing the PUT function in a query.

Category	SAS System Options	Description
	"SQLREDUCEPUTVALUES= System Option" on page 2009	For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the maximum number of SAS format values that can exist in a PUT function expression in order for PROC SQL to consider optimizing the PUT function in a query.
	"SQLREMERGE System Option" on page 2010	Specifies whether the SQL procedure can process queries that use remerging of data.
	"SQLUNDOPOLICY= System Option" on page 2011	Specifies whether the SQL procedure keeps or discards updated data if errors occur while the data is being updated.
System administration: Security	"PDFPASSWORD= System Option" on page 1972	Specifies the password to use to open a PDF document and the password used by a PDF document owner.
	"PDFSECURITY= System Option" on page 1975	Specifies the printing permissions for PDF documents.
	"RLANG System Option" on page 1985	Specifies whether SAS executes R language statements.

Dictionary

APPEND= System Option

Appends a value to the existing value of the specified system option.

Valid in: OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: APPEND= System Option in the documentation for your operating environment

Syntax

APPEND=(*system-option-1=argument-1 system-option-n=argument-n*)

Syntax Description

system-option
can be CMPLIB, FMTSEARCH, MAPS, SASAUTOS, or SASSCRIPT.

argument
specifies a new value that you want to append to the current value of *system-option*. *argument* can be any value that could be specified for *system-option* if *system-option* is set using the OPTIONS statement.

Details

If you specify a new value for the CMPLIB=, FMTSEARCH=, MAPS=, SASAUTOS=, or SASSCRIPT= system options, the new value replaces the value of the option. Instead of replacing the value, you can use the APPEND= system option to append a new value to the current value of the option.

Comparison

The APPEND= system option adds a new value to the end of the current value of the CMPLIB=, FMTSEARCH=, MAPS=, SASAUTOS=, or SASSCRIPT= system options. The INSERT= system option adds a new value as the first value of one of these system options.

Examples

The following table shows the results of adding a value to the end of the FMTSEARCH= option value:

Current FMTSEARCH= Value	Value of APPEND= System Option	New FMTSEARCH= Value
(WORK LIBRARY)	(fmtsearch=(abc def))	(WORK LIBRARY ABC DEF)

See Also

System Option:

"INSERT= System Option" on page 1929

APPLETLOC= System Option

Specifies the location of Java applets.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

Syntax

APPLETLOC="*base-URL*"

Syntax Description

"base-URL"
specifies the address where the SAS Java applets are located. The maximum address length is 256 characters.

Details

The APPLETLOC= system option specifies the base location (typically a URL) of Java applets. These applets are typically accessed from an intranet server or a local CD-ROM.

Examples

Some examples of the *base-URL* are

- "*file:///e:\java*"
- "*http://server.abc.com/SAS/applets*"

AUTHPROVIDERDOMAIN System Option

Associates a domain suffix with an authentication provider.

Valid in: configuration file, SAS invocation

Alias: AUTHPD

Category: Environment control: Initialization and operation

PROC OPTIONS GROUP= EXECMODES

Syntax

AUTHPROVIDERDOMAIN *provider* : *domain*

AUTHPROVIDERDOMAIN (*provider–1* : *domain–1*<, *…provider-n* : *domain-n*>)

Note: In UNIX operating environments, you must insert an escape character before each parenthesis. For example,

```
-authproviderdomain \(ADIR:MyDomain, LDAP:sas\)
```

△

Syntax Description

provider
specifies the authentication provider that is associated with a domain. The following are valid values for *provider*:

ADIR — specifies that the authentication provider be a Microsoft Active Directory server that accepts a bind containing user names and passwords for authentication.

HOSTUSER
: specifies that user names and passwords be authenticated by using the authentication processing that is provided by the host operating system.

  Operating Environment Information: Under the Windows operating environment, assigning the authentication provider using the HOSTUSER domain is the same as assigning the authentication provider using the AUTHSERVER system option. You might want to use the AUTHPROVIDERDOMAIN system option when you specify multiple authentication providers. △

LDAP
: specifies that the authentication provider use a directory server to specify the bind distinguished name (BINDDN) and a password for authentication.

domain
: specifies a site-specific domain name. Quotation marks are required if the domain name contains blanks.

Details

SAS is able to provide authentication of a user through the use of many authentication providers. The AUTHPROVIDERDOMAIN= system option associates a domain suffix with an authentication provider. This association enables the SAS server to choose the authentication provider by the domain name that is presented.

When a domain suffix is not specified or the domain suffix is unknown, authentication is performed on the user ID and password by the host operating system.

Parenthesis are required when you specify more than one set of *provider : domain* pairs.

The maximum length for the AUTHPROVIDERDOMAIN option value is 1,024 characters.

To use the Microsoft Active Directory or LDAP authentication providers, these environment variables must be set in the server or spawner startup script:

Microsoft Active Directory Server:

AD_PORT=*Microsoft Active Directory port number*

AD_HOST=*Microsoft Active Directory host name*

LDAP Server:

LDAP_PORT=*LDAP port number*

LDAP_BASE=*base distinguished name*

LDAP_HOST=*LDAP host_name*

LDAP Server for users connecting with a user ID instead of a distinguished name (DN):

LDAP_PRIV_DN=*privileged DN that is allowed to search for users*

LDAP_PRIV_PW=*LDAP_PRIV_DN password*

Note: If the LDAP server allows anonymous binds, then LDAP_PRIV_DN and LDAP_PRIV_PW are not required. △

In addition to setting these environment variables, you can set the LDAP_IDATTR environment variable to the name of the person-entry LDAP attribute that stores the user ID if the attribute does not contain the default value of **`uid`**.

Examples

The following examples show you how to specify the AUTHPROVIDERDOMAIN option:

- **`-authpd ldap:sas`** causes the SAS server to send credentials for users who log on as *anything*@sas to LDAP for authentication.
- **`-authpd adir:sas`** causes the SAS server to send credentials for users who log on as *anything*@sas to Active Directory for authentication.
- **`-authproviderdomain (hostuser:'my domain', ldap:sas)`** causes the SAS server to send credentials for users who log on as the following:
  - When a user logs on as *anything*@'my domain', authentication is provided by the operating system authentication system
  - When a user logs on as *anything*@sas, authentication is provided by LDAP

See Also

System option:

"PRIMARYPROVIDERDOMAIN= System Option" on page 1976

AUTOSAVELOC= System Option

Specifies the location of the Program Editor autosave file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Display

Restriction: The location that is specified by the AUTOSAVELOC= system option is valid only for the Program Editor. This option does not apply to the Enhanced Editor.

PROC OPTIONS GROUP= ENVDISPLAY

See: AUTOSAVELOC System Option under UNIX OpenVMS

Syntax

AUTOSAVELOC= "*location*"

Syntax Description

location
specifies the pathname of the autosave file. If *location* contains spaces or is specified in an OPTIONS statement, then enclose *location* in quotation marks.

See Also

"Saving Program Editor Files Using Autosave" in the *SAS Companion for Windows*.

"Program Editor Window" in the SAS Help and Documentation.

BINDING= System Option

Specifies the binding edge for duplexed printed output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Restriction: This option is ignored if the printer does not support duplex (two-sided) printing.

Syntax

BINDING=DEFAULTEDGE | LONGEDGE | SHORTEDGE

Syntax Description

DEFAULT | DEFAULTEDGE
specifies that duplexing is done using the default binding edge.

LONG | LONGEDGE
specifies the long edge as the binding edge for duplexed output.

SHORT | SHORTEDGE
specifies the short edge as the binding edge for duplexed output.

Details

The binding edge setting determines how the paper is oriented before output is printed on the second side.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*.

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

See Also

System Option:

"DUPLEX System Option" on page 1894

BOTTOMMARGIN= System Option

Specifies the size of the margin at the bottom of a printed page.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

BOTTOMMARGIN=*margin-size*<*margin-unit*>

Syntax Description

margin-size
specifies the size of the margin.

Restriction: The bottom margin should be small enough so that the top margin plus the bottom margin is less than the height of the paper.

Interactions: Changing the value of this option might result in changes to the value of the PAGESIZE= system option.

<margin-unit>
specifies the units for margin-size. The margin-unit can be *in* for inches or *cm* for centimeters. <*margin-unit*> is saved as part of the value of the BOTTOMMARGIN system option.

Default: inches

Details

All margins have a minimum that is dependent on the printer and the paper size. The default value of the BOTTOMMARGIN system option is **`0.00 in`**.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see ODS statements in the *SAS Output Delivery System: User's Guide*.

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

See Also

System Options:

"LEFTMARGIN= System Option" on page 1934
"RIGHTMARGIN= System Option" on page 1984
"TOPMARGIN= System Option" on page 2039

BUFNO= System Option

Specifies the number of buffers to be allocated for processing SAS data sets.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
System administration: Performance

PROC OPTIONS GROUP= SASFILES
PERFORMANCE

See: BUFNO= System Option in the documentation for your operating environment.

Syntax

BUFNO=*n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the number of buffers to be allocated in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of `8` specifies 8 bytes, and a value of `3m` specifies 3,145,728 bytes.

Tip: Use the notation that best fits the memory size of your system.

***hex*X**
specifies the number of buffers as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the value `2dx` specifies 45 buffers.

MIN
sets the minimum number of buffers to 0, which causes SAS to use the minimum optimal value for the operating environment. This is the default.

MAX
sets the number of buffers to the maximum possible number in your operating environment, up to the largest four-byte, signed integer which is 2^{31}-1, or approximately 2 billion.

Details

The number of buffers is not a permanent attribute of the data set; it is valid only for the current SAS session or job.

BUFNO= applies to SAS data sets that are opened for input, output, or update.

Using BUFNO= can improve execution time by limiting the number of input/output operations that are required for a particular SAS data set. The improvement in execution time, however, comes at the expense of increased memory consumption.

You can estimate the number of buffers you need from the data set page size and the amount of memory in your system. The data set page size can be specified by the BUFSIZE= system option or by the BUFSIZE= data set option. If the default is used, SAS uses the minimal optimal page size for the operating environment. You can find the page size for a data set in the output of the CONTENTS procedure. Once you have the data set page size and the amount of memory available, you can estimate the number of buffers you need. If the number of buffers is too large, SAS might not have enough memory to process the DATA or PROC step. You can change the page size for a data set by recreating the data set using the BUFSIZE= data set option.

Operating Environment Information: Under the Window operating environment, if the SGIO system option is set, the maximum number of bytes that can be processed in an I/O operation is 64MB. Therefore, *number-of-buffers* x *page-size* <= 64MB. △

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Comparisons

- You can override the BUFNO= system option by using the BUFNO= data set option.
- To request that SAS allocate the number of buffers based on the number of data set pages and index file pages, use the SASFILE statement.

See Also

Data Set Option:

"BUFNO= Data Set Option" on page 15

System Option:

"BUFSIZE= System Option" on page 1853

Statements:

"SASFILE Statement" on page 1755

Procedures:

The Contents Procedure

BUFSIZE= System Option

Specifies the permanent buffer page size for output SAS data sets.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
System administration: Performance

PROC OPTIONS GROUP= SASFILES
PERFORMANCE

See: BUFSIZE= System Option in the documentation for your operating environment.

Syntax

BUFSIZE=*n* | *n*K | *n*M | *n*G | *n*T |*hex*X | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the page size in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **8** specifies 8 bytes, and a value of **3m** specifies 3,145,728 bytes.

The default is 0, which causes SAS to use the minimum optimal page size for the operating environment.

***hex*X**
specifies the page size as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **2dx** sets the page size to 45 bytes.

MAX
sets the page size to the maximum possible number in your operating environment, up to the largest four-byte, signed integer, which is 2^{31}-1, or approximately 2 billion bytes.

Details

The page size is the amount of data that can be transferred from a single input/output operation to one buffer. The page size is a permanent attribute of the data set and is used when the data set is processed.

A larger page size can improve execution time by reducing the number of times SAS has to read from or write to the storage medium. However, the improvement in execution time comes at the expense of increased memory consumption.

To change the page size, use a DATA step to copy the data set and either specify a new page or use the SAS default.

Note: If you use the COPY procedure to copy a data set to another library that is allocated with a different engine, the specified page size of the data set is not retained. △

Operating Environment Information: The default value for BUFSIZE= is determined by your operating environment and is set to optimize sequential access. To improve performance for direct (random) access, you should change the value for BUFSIZE=. For the default setting and possible settings for direct access, see the BUFSIZE= system option in the SAS documentation for your operating environment. △

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Comparisons

The BUFSIZE= system option can be overridden by the BUFSIZE= data set option.

See Also

Data Set Option:

"BUFSIZE= Data Set Option" on page 16

System Option:

"BUFNO= System Option" on page 1851

BYERR System Option

Specifies whether SAS produces errors when the SORT procedure attempts to process a _NULL_ data set.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

BYERR | NOBYERR

Syntax Description

BYERR
specifies that SAS issue an error message and stop processing if the SORT procedure attempts to sort a _NULL_ data set.

NOBYERR
specifies that SAS ignore the error message and continue processing if the SORT procedure attempts to sort a _NULL_ data.

Comparisons

The VNFERR system option sets the error flag for a missing variable when a _NULL_ data set is used. The DSNFERR system option specifies how SAS responds when a SAS data set is not found.

See Also

System Options:

"DSNFERR System Option" on page 1892

"VNFERR System Option" on page 2054

"BY-Group Processing" in *SAS Language Reference: Concepts*

BYLINE System Option

Specifies whether to print BY lines above each BY group.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LISTCONTROL

Syntax

BYLINE | NOBYLINE

Syntax Description

BYLINE
specifies that BY lines are printed above each BY group.

NOBYLINE
suppresses the automatic printing of BY lines.

Details

Use NOBYLINE to suppress the automatic printing of BY lines in procedure output. You can then use #BYVAL, #BYVAR, or #BYLINE to display BYLINE information in a TITLE statement.

These SAS procedures perform their own BY line processing by displaying output for multiple BY groups on the same page:

- MEANS
- PRINT
- STANDARD
- SUMMARY
- TTEST (in SAS/STAT software).

With these procedures, NOBYLINE causes a page eject between BY groups. For PROC PRINT, the page eject between BY groups has the same effect as specifying the right most BY variable in the PAGEBY statement.

See Also

Statements:

#BYVAL, #BYVAR, and #BYLINE in the “TITLE Statement” on page 1779

“BY-Group Processing in SAS Programs” in *SAS Language Reference: Concepts*

BYSORTED System Option

Specifies whether observations in one or more data sets are sorted in alphabetic or numeric order or are grouped in another logical order.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

BYSORTED | NOBYSORTED

Syntax Description

BYSORTED
specifies that observations in a data set or data sets are sorted in alphabetic or numeric order.

Requirement: When you use the BYSORTED option, observations must be ordered or indexed according to the values of BY variables.

Interaction: If both the BYSORTED system option and the NOTSORTED statement option on a BY statement are specified, then the NOTSORTED option in the BY statement takes precedence over the BYSORTED system option.

Tip: If BYSORTED is specified, then SAS assumes that the data set is ordered by the BY variable. BYSORTED should be used if the data set is ordered by the BY variable for better performance.

NOBYSORTED
specifies that observations with the same BY value are grouped together but are not necessarily sorted in alphabetic or numeric order.

Tip: When the NOBYSORTED option is specified, you do not have to specify NOTSORTED on every BY statement to access the data sets.

Tip: NOBYSORTED is useful if you have data that falls into other logical groupings such as chronological order or linguistic order. NOBYSORTED allows BY processing to continue without failure when a data set is not actually sorted in alphabetic or numeric order.

Note: If a procedure ignores the NOTSORTED option in a BY statement, then it ignores the NOBYSORTED system option also. △

Details

The requirement for ordering or indexing observations according to the values of BY variables is suspended for BY-group processing when you use the NOBYSORTED option. By default, BY-group processing requires that your data be sorted in alphabetic or numeric order. If your data is grouped in any other way but alphabetic or numeric, then you must use the NOBYSORTED option to allow BY-processing to continue without failure. For more information about BY-group processing, see "BY-Group Processing in SAS Programs" in *SAS Language Reference: Concepts*.

See Also

Statements:

NOTSORTED option in the "BY Statement" on page 1452.

CAPS System Option

Specifies whether to convert certain types of input to uppercase.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

CAPS | NOCAPS

Syntax Description

CAPS

specifies that SAS translate lowercase characters to uppercase in these types of input:

- data following CARDS, CARDS4, DATALINES, DATALINES4, and PARMCARDS statements
- text enclosed in single or double quotation marks
- values in VALUE and INVALUE statements in the FORMAT procedure
- titles, footnotes, variable labels, and data set labels
- constant text in macro definitions
- values of macro variables
- parameter values passed to macros.

Note: Data read from external files and SAS data sets are not translated to uppercase. △

NOCAPS

specifies that lowercase characters that occur in the types of input that are listed above are not translated to uppercase.

Comparisons

The CAPS system option and the CAPS command both specify whether input is converted to uppercase. The CAPS command, which is available in windows that allow text editing, can act as a toggle. The CAPS command converts all text that is entered from the keyboard to uppercase. If either the CAPS system option or the CAPS command is in effect, all applicable input is translated to uppercase.

See Also

Command:

CAPS in SAS Help and Documentation

CARDIMAGE System Option

Specifies whether SAS processes source and data lines as 80-byte cards.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

See: CARDIMAGE System Option in the documentation for your operating environment.

Syntax

CARDIMAGE | NOCARDIMAGE

Syntax Description

CARDIMAGE
specifies that SAS source and data lines be processed as if they were punched card images—all exactly 80 bytes long and padded with blanks. That is, column 1 of a line is treated as if it immediately followed column 80 of the previous line. Therefore, *tokens* can be split across lines. (A *token* is a character or series of characters that SAS treats as a discrete word.)

Strings in quotation marks (literal tokens) that begin on one line and end on another are treated as if they contained blanks out to column 80 of the first line. Data lines longer than 80 bytes are split into two or more 80-byte lines. Data lines are not truncated regardless of their length.

NOCARDIMAGE
specifies that SAS source and data lines not be treated as if they were 80-byte card images. When NOCARDIMAGE is in effect, the end of a line is always treated as the end of the last token, except for strings in quotation marks. Strings in quotation marks can be split across lines. Other types of tokens cannot be split across lines under any circumstances. Strings in quotation marks that are split across lines are not padded with blanks.

Operating Environment Information: CARDIMAGE is generally used in the z/OS operating environment; NOCARDIMAGE is used in other operating environments. △

Examples

Consider the following DATA step:

```
data;
   x='A
B';
run;
```

If CARDIMAGE is in effect, the variable X receives a value that consists of 78 characters: the A, 76 blanks, and the B. If NOCARDIMAGE is in effect, the variable X receives a value that consists of two characters: AB, with no intervening blanks.

CATCACHE= System Option

Specifies the number of SAS catalogs to keep open in cache memory.

Valid in: configuration file, SAS invocation

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

See: CATCACHE= System Option in the documentation for your operating environment.

Syntax

CATCACHE=*n* | *hex*X | MIN | MAX |

Syntax Description

n
specifies any integer greater than or equal to 0 in terms of bytes. If $n > 0$, SAS places up to that number of open-file descriptors in cache memory instead of closing the catalogs.

***hex*X**
specifies the number of open-file descriptors that are kept in cache memory as a hexadecimal number. You must specify the value beginning with a number (0-9), followed by an X. For example, the value **`2dx`** sets the number of catalogs to keep open to 45 catalogs.

MIN
sets the number of open-file descriptors that are kept in cache memory to 0.

MAX
sets the number of open-file descriptors that are kept in cache memory to the largest, signed, 4–byte integer representable in your operating environment.

Tip: The recommended maximum setting for this option is 10.

Details

Use the CATCACHE= system option to tune an application by avoiding the overhead of repeatedly opening and closing the same SAS catalogs.

CAUTION:
When using both the CBUFNO= and CATCACHE= options, if one of the option's value is set higher than zero, the other option must be set to zero. △

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For more information, see the SAS documentation for your operating environment. △

Operating Environment Information: Some system settings might affect the default setting. For more information, see the documentation for your operating system. △

CBUFNO= System Option

Specifies the number of extra page buffers to allocate for each open SAS catalog.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

CBUFNO=*n* | *n*K | *n*M | *n*G | *n*T | *hex*X |MIN | MAX

Syntax Description

***n* | *n*K | nM | nG | *n*T**
specifies the number of extra page buffers in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes; 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **`8`** specifies 8 bytes, and a value of **`3m`** specifies 3,145,728 bytes.

MIN
sets the number of extra page buffers to 0.

MAX
sets the number of extra page buffers to 20.

***hex*X**
specifies the number of extra page buffers as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`0ax`** sets the number of extra page buffers to 10 buffers.

Details

The CBUFNO= option is similar to the BUFNO= option that is used for SAS data set processing.

Increasing the value for the CBUFNO= option might result in fewer I/O operations when your application reads very large objects from catalogs. Increasing this value also comes with the normal tradeoff between performance and memory usage. If memory is a serious constraint for your system, you should not increase the value of the CBUFNO= option. Do not increase the value of the CBUFNO= option if you have increased the value of the CATCACHE= option.

CAUTION:
When using both the CBUFNO= and CATCACHE= options, if one of the option's value is set higher than zero, the other option must be set to zero. △

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For more information, see the SAS documentation for your operating environment. △

CENTER System Option

Specifies whether to center or left align SAS procedure output.

Alias: CENTRE

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LISTCONTROL

Syntax

CENTER | NOCENTER

Syntax Description

CENTER
centers SAS procedure output.

NOCENTER
left aligns SAS procedure output.

CGOPTIMIZE= System Option

Specifies the level of optimization to perform during code compilation.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Alias: CGOPT

Category: System administration: Performance

System administration: Code generation

PROC OPTIONS GROUP= PERFORMANCE
CODEGEN

Syntax

CGOPTIMIZE=0 | 1 | 2 | 3

Syntax Description

0
specifies not to perform optimization.

1
specifies to perform stage 1 optimization. Stage 1 optimization removes redundant instructions, missing value checks, and repetitive computations for array subscriptions; detects patterns of instructions and replaces them with more efficient sequences.

2
specifies to perform stage 2 optimization. Stage 2 performs optimizations that pertain to the SAS register.

Interaction: Stage 2 optimization for a large DATA step program can result in a significant increase in compilation time and thus overall execution time.

3
specifies to perform full optimization, which is a combination of stages 1 and 2. This is the default value.

See Also

Reducing CPU Time by Modifying Program Compilation Optimization in *SAS Language Reference: Concepts*

CHARCODE System Option

Specifies whether specific keyboard combinations are substituted for special characters that are not on the keyboard.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Display

PROC OPTIONS GROUP= ENVDISPLAY

Syntax

CHARCODE | NOCHARCODE

Syntax Description

CHARCODE
allows certain character combinations to be substituted for special characters that might not be on your keyboard.

NOCHARCODE
does not allow substitutions for certain keyboard characters.

Details

If you do not have the following symbols on your keyboard, you can use these character combinations to create the symbols that you need when CHARCODE is active:

Symbol	Characters
back quote (‘)	?:
backslash (\)	?,
left brace ({)	?(
right brace (})	?)
logical not sign (¬ or ^)	?=
left square bracket ([)	?<
right square bracket (])	?>
underscore (_)	?-
vertical bar (|)	?/

Examples

This statement produces the output [TEST TITLE]:

```
title '?<TEST TITLE?>';
```

CLEANUP System Option

For an out-of-resource condition, specifies whether to perform an automatic cleanup or a user-specified cleanup.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Environment control: Error handling
PROC OPTIONS GROUP= ERRORHANDLING
See: CLEANUP System Option in the documentation for your operating environment.

Syntax

CLEANUP | NOCLEANUP

Syntax Description

CLEANUP
specifies that during the entire session, SAS attempts to perform automatic, continuous clean-up of resources that are not essential for execution. Nonessential resources include resources that are not visible to the user (for example, cache memory) and resources that are visible to the user (for example, the KEYS windows).

When CLEANUP is in effect and an out-of-resource condition occurs (except for a disk-full condition), a dialog box is not displayed, and no intervention is required by the user. When CLEANUP is in effect and a disk-full condition occurs, a dialog box displays that allows the user to decide how to proceed.

NOCLEANUP
specifies that SAS allow the user to choose how to handle an out-of-resource condition. When NOCLEANUP is in effect and SAS cannot execute because of a lack of resources, SAS automatically attempts to clean up resources that are not visible to the user (for example, cache memory). However, resources that are visible to the user (for example, windows) are not automatically cleaned up. Instead, a dialog box appears that allows the user to choose how to proceed.

Details

This table lists the dialog box choices:

Dialog Box Choice	Action
Free windows	clears all windows not essential for execution.
Clear paste buffers	deletes paste buffer contents.
Deassign inactive librefs	prompts user for librefs to delete.
Delete definitions of all SAS macros and macro variables	deletes all macro definitions and variables.
Delete SAS files	allows user to select files to delete.
Clear Log window	erases Log window contents.
Clear Output window	erases Output window contents.
Clear Program Editor window	erases Program Editor window contents.
Clear source spooling/DMS recall buffers	erases recall buffers.
More items to clean up	displays a list of other resources that can be cleaned up.

Dialog Box Choice	Action
Clean up everything	cleans up all other options that are shown on the requestor window. This selection only applies to the current clean-up request, not to the entire SAS session.
Continuous clean up	performs automatic, continuous clean-up. When continuous clean up is selected, SAS cleans up as many resources as possible in order to continue execution, and it ceases to display the requester window. Selecting continuous clean-up has the same effect as specifying CLEANUP. This selection applies to the current clean-up request and to the remainder of the SAS session.

Operating Environment Information: Some operating environments might also include these choices in the dialog box:

Dialog Box Choice	Action
Execute X command	enables the user to erase files and perform other clean-up operations.
Do nothing	halts the clean-up request and returns to the SAS session. This selection only applies to the current clean-up request, not to the entire SAS session.

If an out-of-resource condition cannot be resolved, the dialog box continues to display. In that case, see the SAS documentation for your operating environment for instructions on terminating the SAS session.

When running in modes other than a windowing environment, the operation of CLEANUP depends on your operating environment. For details, see the SAS documentation for your operating environment. △

CMPLIB= System Option

Specifies one or more SAS data sets that contain compiler subroutines to include during program compilation.

Valid in: configuration file, SAS invocation, OPTIONS statement, System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

CMPLIB=*libref.data-set* | (*libref.data-set-1 … libref.data-set-n*) | (*libref.data-set-n – libref.data-set-m*)

Syntax Description

libref.data-set
specifies the libref and the data set of the compiler subroutines that are to be included during program compilation. The *libref* and *data-set* must be valid SAS names.

libref.data-set-n – libref.data-set-m
specifies a range of compiler subroutines that are to be included during program compilation. The name of the libref and the data set must be valid SAS names that contain a numeric suffix.

Details

SAS procedures, DATA steps, and macro programs that perform non-linear statistical modeling or optimization use a SAS language compiler subsystem that compiles and executes your SAS programs. The compiler subsystem generates machine language code for the computer on which SAS is running. The SAS procedures that use the SAS language compiler are CALIS, COMPILE, GA, GENMOD, MODEL, NLIN, NLMIXED, NLP, PHREG, Risk Dimensions procedures, and SQL.

The subroutines that you want to include must already have been compiled. All the subroutines in *libref.data-set* are included.

You can specify a single *libref.data-set*, a list of *libref.data-set* names, or a range of *libref.data-set* names with numeric suffixes. When you specify more than one *libref.data-set* name, separate the names with a space and enclose the names in parentheses.

After SAS starts, you can use the APPEND or INSERT system options to add additional data sets.

Examples

Number of Libraries	OPTIONS Statement
One library	`options cmplib=sasuser.cmpl;`
Two or more libraries	`options cmplib=(sasuser.cmpl sasuser.cmplA sasuser.cmpl3);`
A range of libraries	`options cmplib=(sasuser.cmpl1 - sasuser.cmpl6);`

See Also

- System options:

  "APPEND= System Option" on page 1844

  "INSERT= System Option" on page 1929

CMPMODEL= System Option

Specifies the output model type for the MODEL procedure.

Valid in: configuration file, SAS invocation, OPTIONS statement, System Options window

Category: System administration: Performance

PROC OPTIONS GROUP= Performance

Syntax

CMPMODEL=BOTH | CATALOG | XML

Syntax Description

BOTH
specifies that the MODEL procedure create two output types for a model, one as a SAS catalog entry and the other as an XML file. This is the default.

CATALOG
specifies that the output model type is an entry in a SAS catalog.

XML
specifies that the output model type is an XML file.

See Also

The MODEL Procedure in *SAS/ETS User's Guide*

CMPOPT= System Option

Specifies the type of code generation optimizations to use in the SAS language compiler.

Valid in: configuration file, SAS invocation, OPTIONS statement, System Options window

Category: System administration: Performance

PROC OPTIONS GROUP= PERFORMANCE

Syntax

CMPOPT=*optimization-value* | (*optimization-value-1 ... optimization-value-n*) | "*optimization-value–1 ... optimization-value-n*" | ALL | NONE

NOCMPOPT

Syntax Description

optimization
specifies the type of optimization that the SAS compiler is to use. Valid values are

EXTRAMATH | NOEXTRAMATH
specifies to keep or remove mathematical operations that do not affect the outcome of a statement. When you specify EXTRAMATH, the compiler retains the extra mathematical operations. When you specify NOEXTRAMATH, the extra mathematical operations are removed.

FUNCDIFFERENCING | NOFUNCDIFFERENCING
specifies whether analytic derivatives are computed for user defined functions. When you specify NOFUNCDIFFERENCING, analytic derivatives are computed for user defined functions. When you specify FUNCDIFFERENCING, numeric differencing is used to calculate derivatives for user defined functions. The default is NOFUNCDIFFERENCING.

GUARDCHECK | NOGUARDCHECK
specifies whether to check for array boundary problems. When you specify GUARDCHECK, the compiler checks for array boundary problems. When you specify NOGUARDCHECK, the compiler does not check for array boundary problems.

Interaction: NOGUARDCHECK is set when CMPOPT is set to ALL and when CMPOPT is set to NONE.

MISSCHECK | NOMISSCHECK
specifies whether to check for missing values in the data. If the data contains a significant amount of missing data, then you can optimize the compilation by specifying MISSCHECK. If the data rarely contains missing values, then you can optimize the compilation by specifying NOMISSCHECK.

PRECISE | NOPRECISE
specifies to handle exceptions at an operation boundary or at a statement boundary. When you specify PRECISE, exceptions are handled at the operation boundary. When you specify NOPRECISE, exceptions are handled at the statement boundary.

Tip: EXTRAMATH, MISSCHECK, PRECISE, GUARDCHECK, and FUNCDIFFERENCING can be specified in any combination when you specify one or more values.

ALL
specifies that the compiler is to optimize the machine language code by using the (NOEXTRAMATH NOMISSCHECK NOPRECISE NOGUARDCHECK NOFUNCDIFFERENCING) optimization values. This is the default.

Restriction: ALL cannot be specified in combination with any other values.

NONE
specifies that the compiler is not set to optimize the machine language code by using the (EXTRAMATH MISSCHECK PRECISE NOGUARDCHECK FUNCDIFFERENCING) optimization values.

Restriction: NONE cannot be specified in combination with any other values.

NOCMPOPT
specifies to set the value of CMPOPT to ALL. The compiler is to optimize the machine language code by using the (NOEXTRAMATH NOMISSCHECK NOPRECISE NOGUARDCHECK NOFUNCDIFFERENCING) optimization values.

Restriction: NOCMPOPT cannot be specified in combination with values for the CMPOPT option.

Details

SAS procedures that perform non-linear statistical modeling or optimization use a SAS language compiler subsystem that compiles and executes your SAS programs. The compiler subsystem generates machine language code for the computer on which SAS is running. By specifying values with the CMPOPT option, the machine language code can be optimized for efficient execution. The SAS procedures that use the SAS language compiler are CALIS, COMPILE, GENMOD, MODEL, PHREG, NLIN, NLMIXED, NLP, and RISK.

To specify multiple optimization values, the values must be enclosed in either parentheses, single quotation marks, or double quotation marks. When CMPOPT is set to multiple values, the parentheses or quotation marks are retained as part of the value. They are not retained as part of the value when CMPOPT is set to a single value.

If a value is entered more than once, then the last setting is used. For example, if you specify CMPOPT=(PRECISE NOEXTRAMATH NOPRECISE), then the values that are set are NOEXTRAMATH and NOPRECISE. All leading, trailing, and embedded blanks are removed.

When you specify EXTRAMATH or NOEXTRAMATH, some of the mathematical operations that are either included or excluded in the machine language code are

$x * 1$	$x * -1$
$x \div 1$	$x \div -1$
$x + 0$	$x - 0$
$x - x$	$x \div x$
$- -x$	any operation on two literal constants

Examples

OPTIONS Statement	Result
`options cmpopt=(extramath);`	extramath
`options cmpopt="extramath missscheck precise";`	"precise extramath extramath"
`options nocmpopt;`	(noextramath nomisscheck noprecise noguardcheck nofuncdifferencing)

COLLATE System Option

Specifies whether to collate multiple copies of printed output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

COLLATE | NOCOLLATE

Syntax Description

COLLATE
specifies to collate multiple copies of printed output.

NOCOLLATE
specifies not to collate multiple copies of printed output. This is the default.

Details

When you send a print job to the printer and you want multiple copies of multiple pages, the COLLATE option controls how the pages are ordered:

COLLATE causes the pages to print consecutively: 123, 123, 123...

NOCOLLATE causes the same-numbered pages to print together: 111, 222, 333...

Note: You can also control collation with the SAS windowing environment Page Setup window, invoked with the DMPAGESETUP command. △

Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment.

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*.

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

See Also

System Option:

"COPIES= System Option" on page 1874

COLORPRINTING System Option

Specifies whether to print in color if color printing is supported.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

COLORPRINTING | NOCOLORPRINTING

Syntax Description

COLORPRINTING
specifies to attempt to print in color.

NOCOLORPRINTING
specifies not to print in color.

Details

Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment.

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*.

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

COMPRESS= System Option

Specifies the type of compression of observations to use for output SAS data sets.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
System administration: Performance

PROC OPTIONS GROUP= SASFILES
PERFORMANCE

Restriction: The TAPE engine does not support the COMPRESS= system option.

Syntax

COMPRESS=NO | YES | CHAR | BINARY

Syntax Description

NO
specifies that the observations in a newly created SAS data set are uncompressed (fixed-length records).

Alias: N | OFF

YES | CHAR
specifies that the observations in a newly created SAS data set are compressed (variable-length records) by SAS using RLE (Run Length Encoding). RLE compresses observations by reducing repeated consecutive characters (including blanks) to two-byte or three-byte representations.

Alias: Y, ON

Tip: Use this compression algorithm for character data.

Note: COMPRESS=CHAR is accepted by Version 7 and later versions. △

BINARY
specifies that the observations in a newly created SAS data set are compressed (variable-length records) by SAS using RDC (Ross Data Compression). RDC combines run-length encoding and sliding-window compression to compress the file.

Tip: This method is highly effective for compressing medium to large (several hundred bytes or larger) blocks of binary data (numeric variables). Because the compression function operates on a single record at a time, the record length needs to be several hundred bytes or larger for effective compression.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Details

Compressing a file is a process that reduces the number of bytes required to represent each observation. Advantages of compressing a file include reduced storage requirements for the file and fewer I/O operations necessary to read or write to the data during processing. However, more CPU resources are required to read a compressed file (because of the overhead of uncompressing each observation), and there are situations when the resulting file size might increase rather than decrease.

Use the COMPRESS= system option to compress all output data sets that are created during a SAS session. Use the option only when you are creating SAS data files (member type DATA). You cannot compress SAS views, because they contain no data.

Once a file is compressed, the setting is a permanent attribute of the file, which means that to change the setting, you must re-create the file. That is, to uncompress a file, specify COMPRESS=NO for a DATA step that copies the compressed file.

Note: For the COPY procedure, the default value CLONE uses the compression attribute from the input data set for the output data set. If the engine for the input data set does not support the compression attribute, then PROC COPY uses the current value of the COMPRESS= system option. For more information about CLONE and NOCLONE, see COPY statement in the DATASETS procedure in the *Base SAS Procedures Guide*. This interaction does not apply when using SAS/SHARE or SAS/CONNECT. △

Comparisons

The COMPRESS= system option can be overridden by the COMPRESS= option in the LIBNAME statement and the COMPRESS= data set option.

The data set option POINTOBS=YES, which is the default, determines that a compressed data set can be processed with random access (by observation number) rather than sequential access. With random access, you can specify an observation number in the FSEDIT procedure and the POINT= option in the SET and MODIFY statements.

When you create a compressed file, you can also specify REUSE=YES (as a data set option or system option) in order to track and reuse space. With REUSE=YES, new observations are inserted in space freed when other observations are updated or deleted. When the default REUSE=NO is in effect, new observations are appended to the existing file.

POINTOBS=YES and REUSE=YES are mutually exclusive. That is, they cannot be used together. REUSE=YES takes precedence over POINTOBS=YES. That is, if you set REUSE=YES, SAS automatically sets POINTOBS=NO.

The TAPE engine does not support the COMPRESS= system option, but the engine does support the COMPRESS= data set option.

The XPORT engine does not support compression.

See Also

Data Set Options:

"COMPRESS= Data Set Option" on page 19

"POINTOBS= Data Set Option" on page 48

"REUSE= Data Set Option" on page 56

Statements:

"LIBNAME Statement" on page 1656

System Option:

"REUSE= System Option" on page 1983

"Compressing Data Files" in *SAS Language Reference: Concepts*

COPIES= System Option

Specifies the number of copies to print.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

COPIES=*n*

Syntax Description

n
specifies the number of copies.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*. For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

See Also

System Option:
"COLLATE System Option" on page 1871

CPUCOUNT= System Option

Specifies the number of processors that the thread-enabled applications should assume will be available for concurrent processing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: System administration: Performance

PROC OPTIONS GROUP= PERFORMANCE

Default: Under Windows, OpenVMS, and z/OS, the default is ACTUAL. Under UNIX, the default is either ACTUAL or 4 for systems that have more than four processors.

Interaction: If the THREADS system option is set to NOTHREADS, the CPUCOUNT= option has no effect.

Syntax

CPUCOUNT= 1 - 1024 | ACTUAL

Syntax Description

1-1024
is the number of CPUs that SAS will assume are available for use by thread-enabled applications.

Tip: The value is typically set to the actual number of CPUs available to the current process by your configuration.

Tip: Setting CPUCOUNT= to a number greater than the actual number of available CPUs might result in reduced overall performance of SAS.

ACTUAL
returns the number of physical processors that are associated with the operating system where SAS is executing. If the operating system is executing in a partition, the value of the CPUCOUNT system is the number of physical processors that are associated with the operating system in that partition.

Tip: This number can be less than the physical number of CPUs if the SAS process has been restricted by system administration tools.

Tip: Setting CPUCOUNT= to ACTUAL at any time causes the option to be reset to the number of physical processors that are associated with the operating system at that time. If the operating system is executing in a partition, the value of the CPUCOUNT system is the number of physical processors that are associated with the operating system in that partition.

Tip: If your system supports Simultaneous Multi-Threading (SMT), hyperthreading, or Chip Multi-Threading (CMT), the value of the CPUCOUNT= option represents the number of such threads on the system.

Details

Certain procedures have been modified to take advantage of multiple CPUs by threading the procedure processing. The Base SAS engine also uses threading to create an index. The CPUCOUNT= option provides the information that is needed to make decisions about the allocation of threads.

Changing the value of CPUCOUNT= affects the degree of parallelism each thread-enabled process attempts to achieve. Setting CPUCOUNT to a number greater than the actual number of available CPUs might result in reduced overall performance of SAS.

Comparisons

When the related system option THREADS is in effect, threading will be active where available. The value of the CPUCOUNT= option affects the performance of THREADS by suggesting how many system CPUs are available for use by thread-enabled SAS procedures.

See Also

System Options:

"THREADS System Option" on page 2037

"UTILLOC= System Option" on page 2043

"Support for Parallel Processing" in *SAS Language Reference: Concepts*.

CPUID System Option

Specifies whether the CPU identification number is written to the SAS log.

Valid in: configuration file, SAS invocation
Category: Log and procedure output control: SAS log
PROC OPTIONS GROUP= LOGCONTROL

Syntax

CPUID | NOCPUID

Syntax Description

CPUID
specifies that the CPU identification number is printed at the top of the SAS log after the licensing information.

NOCPUID
specifies that the CPU identification number is not written to the SAS log.

See Also

The SAS Log in *SAS Language Reference: Concepts*

DATASTMTCHK= System Option

Specifies which SAS statement keywords are prohibited from being specified as a one-level DATA step name to protect against overwriting an input data set.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Files: SAS Files
PROC OPTIONS GROUP= SASFILES

Syntax

DATASTMTCHK=COREKEYWORDS | ALLKEYWORDS | NONE

Syntax Description

COREKEYWORDS
prohibits certain words as one-level SAS data set names in the DATA statement. They can appear as two-level names. The following keywords cannot appear as one-level SAS data set names:

MERGE
RETAIN
SET
UPDATE.

For example, SET is not acceptable in the DATA statement, but SAVE.SET and WORK.SET are acceptable. COREKEYWORDS is the default.

ALLKEYWORDS
prohibits any keyword that can begin a statement in the DATA step (for example, ABORT, ARRAY, INFILE) as a one-level data set name in the DATA statement.

NONE
provides no protection against overwriting SAS data sets.

Details

If you omit a semicolon in the DATA statement, you can overwrite an input data set if the next statement is SET, MERGE, or UPDATE. Different, but significant, problems arise when the next statement is RETAIN. DATASTMTCHK= enables you to protect yourself against overwriting the input data set.

DATE System Option

Specifies whether to print the date and time that a SAS program started.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

Syntax

DATE | NODATE

Syntax Description

DATE
specifies that the date and the time that the SAS program started are printed at the top of each page of the SAS log and any output that is created by SAS.

Note: In an interactive SAS session, the date and time are noted only in the output window. △

NODATE
specifies that the date and the time are not printed.

See Also

The SAS Log in *SAS Language Reference: Concepts*

DATESTYLE= System Option

Specifies the sequence of month, day, and year when ANYDTDTE, ANYDTDTM, or ANYDTTME informat data is ambiguous.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Environment control: Language control
Input control: Data Processing
PROC OPTIONS GROUP= INPUTCONTROL
LANGUAGECONTROL

Syntax

DATESTYLE= MDY | MYD | YMD | YDM | DMY | DYM | LOCALE

Syntax Description

MDY
specifies that SAS set the order as month, day, year.

MYD
specifies that SAS set the order as month, year, day.

YMD
specifies that SAS set the order as year, month, day.

YDM
specifies that SAS set the order as year, day, month.

DMY
specifies that SAS set the order as day, month, year.

DYM
specifies that SAS set the order as day, year, month.

LOCALE
specifies that SAS set the order based on the value that corresponds to the LOCALE= system option value and is one of the following: MDY | MYD | YMD | YDM | DMY | DYM.

Details

System option DATESTYLE= identifies the order of month, day, and year. The default value is LOCALE. The default LOCALE system option value is English, therefore, the default DATESTYLE order is MDY.

Operating Environment Information: See "Locale Values" in *SAS National Language Support (NLS): Reference Guide* to get the default settings for each locale option value. △

See Also

System Option:

"LOCALE System Option: UNIX, Windows, OpenVMS, and z/OS" in *SAS National Language Support (NLS): Reference Guide*

Informats:

"ANYDTDTE*w*. Informat" on page 1299

"ANYDTDTM*w*. Informat" on page 1301

"ANYDTTME*w*. Informat" on page 1304

DEFLATION= System Option

Specifies the level of compression for device drivers that support the Deflate compression algorithm.

Alias: DEFLATE

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Requirement: The UPRINTCOMPRESSION system option must be set in order to compress files.

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

DEFLATION=*n* | **MIN** | **MAX**

Syntax Description

n

specifies the level of compression. The larger the number, the greater the compression. For example, *n*=0 is the minimum compression level (completely uncompressed), and *n*=9 is the maximum compression level.

Default: 6

Range: 0–9

MIN

specifies the minimum compression level of 0.

MAX

specifies the maximum compression level of 9.

Details

The DEFLATION= system option controls the level of compression for device drivers that support Deflate compression. The PRINTERPATH= system option must be set to

one of the following SAS device drivers that support Deflate compression: the PDF device driver or the SVG Universal Printer drivers.

The ODS PRINTER statement option, COMPRESS=, takes precedence over the DEFLATION system option.

See Also

System options:

"PRINTERPATH= System Option" on page 1978

"UPRINTCOMPRESSION System Option" on page 2041

Statements:

"ODS PRINTER Statement" in the *SAS Output Delivery System: User's Guide*

DETAILS System Option

Specifies whether to include additional information when files are listed in a SAS library.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

Syntax

DETAILS | NODETAILS

Syntax Description

DETAILS
includes additional information when some SAS procedures and windows display a listing of files in a SAS library.

NODETAILS
does not include additional information.

Details

The DETAILS specification sets the default display for these components of SAS:

- the CONTENTS procedure
- the DATASETS procedure.

The type and amount of additional information that displays depends on which procedure or window you use.

See Also

The SAS Log in *SAS Language Reference: Concepts*

DEVICE= System Option

Specifies the device driver to which SAS/GRAPH sends procedure output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Alias: DEV=

Category: Graphics: Driver settings

PROC OPTIONS GROUP= GRAPHICS

See: DEVICE= System Option in the documentation for your operating environment.

Syntax

DEVICE=*device-driver-specification*

Syntax Description

device-driver-specification
specifies the name of a device driver.

Details

If you omit the device-driver name, you are prompted to enter a driver name when you execute a procedure that produces graphics.

Operating Environment Information: The syntax that is shown applies to the OPTIONS statement. However, when you specify DEVICE= either on the command-line or in a configuration file, the syntax is specific to your operating environment and might include additional or alternate punctuation. △

See Also

Device Drivers in *SAS/GRAPH: Reference, Second Edition*

DKRICOND= System Option

Specifies the level of error detection to report when a variable is missing from an input data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

DKRICOND=ERROR | WARN | WARNING | NOWARN | NOWARNING

Syntax Description

ERROR
sets the error flag and writes an error message to the SAS log when a variable is missing from an input data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

WARN|WARNING
writes a warning message to the SAS log when a variable is missing from an input data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

NOWARN|NOWARNING
does not write a warning message to the SAS log when a variable is missing from an input data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

Examples

In the following statements, if the variable X is not in data set B and DKRICOND=ERROR, SAS sets the error flag to 1 and displays error messages:

```
data a;
   set b(drop=x);
run;
```

See Also

System Option:

"DKROCOND= System Option" on page 1883

DKROCOND= System Option

Specifies the level of error detection to report when a variable is missing for an output data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

DKROCOND=ERROR | WARN | WARNING | NOWARN | NOWARNING

Syntax Description

ERROR
sets the error flag and writes an error message to the SAS log when a variable is missing for an output data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

WARN | WARNING
writes a warning message to the SAS log when a variable is missing for an output data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

NOWARN | NOWARNING
does not write a warning message to the SAS log when a variable is missing for an output data set during the processing of a DROP=, KEEP=, or RENAME= data set option.

Examples

In the following statements, if the variable X is not in data set A and DKROCOND=ERROR, SAS sets the error flag to 1 and displays error messages:

```
data a;
   drop x;
run;
```

See Also

System Option:
"DKRICOND= System Option" on page 1883

DLDMGACTION= System Option

Specifies the type of action to take when a SAS data set or a SAS catalog is detected as damaged.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

DLDMGACTION=FAIL | ABORT | REPAIR | NOINDEX | PROMPT

Syntax Description

FAIL
stops the step and issues an error message to the log immediately. This is the default for batch mode.

ABORT
terminates the step and issues an error message to the log, and ends the SAS session.

REPAIR
For data files, automatically repairs and rebuilds indexes and integrity constraints, unless the data file is truncated. You use the REPAIR statement to restore the truncated data file. It issues a warning message to the log. This is the default for interactive mode. For catalogs, automatically deletes catalog entries for which an error occurs during the repair process.

NOINDEX
For data files, automatically repairs the data file without the indexes and integrity constraints, deletes the index file, updates the data file to reflect the disabled indexes and integrity constraints, and limits the data file to be opened only in INPUT mode. A warning is written to the SAS log instructing you to execute the PROC DATASETS REBUILD statement to correct or delete the disabled indexes and integrity constraints. For more information, see the "REBUILD Statement" in the "DATASETS Procedure" in *Base SAS Procedures Guide* and "Recovering Disabled Indexes and integrity Constraints" in *SAS Language Reference: Concepts*.

Restriction: NOINDEX does not apply to damaged catalogs or libraries, only data files.

PROMPT
For data sets, displays a dialog box where you can specify either FAIL, ABORT, REPAIR, or NOINDEX. For a damaged catalog or library, PROMPT displays a dialog box where you can specify either FAIL, ABORT, or REPAIR.

DMR System Option

Specifies whether to enable SAS to invoke a server session for use with a SAS/CONNECT client.

Valid in: configuration file, SAS invocation

Category: Environment control: Initialization and operation

PROC OPTIONS GROUP= EXECMODES

Syntax

DMR | NODMR

Syntax Description

DMR
enables you to invoke a remote SAS session in order to connect with a SAS/CONNECT client.

NODMR
disables you from invoking a remote SAS session.

Details

You normally invoke the remote SAS session from a local session by including DMR with the SAS command in a script that contains a TYPE statement. (A *script* is a text file that contains statements to establish or terminate the SAS/CONNECT link between the local and the remote SAS sessions.)

The following SAS execution mode invocation option has precedence over this option:

- OBJECTSERVER

DMR overrides all other SAS execution mode invocation options. See "Order of Precedence" on page 1829 for more information about invocation option precedence.

See Also

DMR information in *SAS/CONNECT User's Guide*

DMS System Option

Specifies whether to invoke the SAS windowing environment and display the Log, Editor, and Output windows.

Valid in: configuration file, SAS invocation
Category: Environment control: Initialization and operation
PROC OPTIONS GROUP= EXECMODES

Syntax

DMS | NODMS

Syntax Description

DMS
invokes the SAS windowing environment and displays the Log, an Editor window, and Output windows.

NODMS
invokes an interactive line mode SAS session.

Details

When you invoke SAS and you are using a configuration file or the command line to control your system option settings, it is possible to create a situation where some system option settings conflict with other system option settings. The following invocation system options, in order, have precedence over the DMS invocation system option:

1. OBJECTSERVER.
2. DMR
3. SYSIN

If you specify DMR while using another invocation option of equal precedence to invoke SAS, SAS uses the last option that is specified. See "Order of Precedence" on page 1829 for more information about invocation option precedence.

See Also

System Options:

"DMR System Option" on page 1885

"DMSEXP System Option" on page 1887

"EXPLORER System Option" on page 1906

DMSEXP System Option

Specifies whether to invoke the SAS windowing environment and display the Explorer, Editor, Log, Output, and Results windows.

Valid in: configuration file, SAS invocation

Category: Environment control: Initialization and operation

PROC OPTIONS GROUP= EXECMODES

Syntax

DMSEXP | NODMSEXP

Syntax Description

DMSEXP
invokes SAS with the Explorer, Editor, Log, Output, and Results windows active.

NODMSEXP
invokes SAS with the Editor, Log, and Output windows active.

Details

In order to set DMSEXP or NODMSEXP, the DMS option must be set. The following SAS execution mode invocation options, in order, have precedence over this option:

1. OBJECTSERVER.

2 DMR

3 SYSIN

If you specify DMSEXP with another execution mode invocation option of equal precedence, SAS uses only the last option listed. See "Order of Precedence" on page 1829 for more information about invocation option precedence.

See Also

System Options:

"DMS System Option" on page 1886

"DMR System Option" on page 1885

"EXPLORER System Option" on page 1906

DMSLOGSIZE= System Option

Specifies the maximum number of rows that the SAS Log window can display.

Valid in: configuration file, SAS invocation

Category: Environment control: Display

Log and procedure output control: SAS log

Restriction: This option is valid only in the SAS windowing environment.

PROC OPTIONS GROUP= ENVDISPLAY

LOGCONTROL

Syntax

DMSLOGSIZE= *n* | *n*K | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K**

specifies the maximum number of rows that can be displayed in the SAS windowing environment Log window in multiples of 1 (n) or 1,024 (nK). For example, a value of 800 specifies 800 rows, and a value of 3K specifies 3,072 rows. Valid values range from 500 to 999999. The default is 99999.

***hex*X**

specifies the maximum number of rows that can be displayed in the SAS windowing environment Log window as a hexadecimal value. You must specify the value beginning with a number (0-9), followed by an X. For example, **`2ffx`** specifies 767 rows and **`0A00x`** specifies 2,560 rows.

MIN

specifies to set the maximum number of rows that can be displayed in the SAS windowing environment Log window to 500.

MAX
specifies to set the maximum number of rows that can be displayed in the SAS windowing environment Log window to 999999.

Details

When the maximum number of rows have been displayed in the Log window, SAS prompts you to either file, print, save, or clear the Log window.

See Also

System Option:
"DMSOUTSIZE= System Option" on page 1889
"The SAS Log" in *SAS Language Reference: Concepts*

DMSOUTSIZE= System Option

Specifies the maximum number of rows that the SAS Output window can display.

Valid in: configuration file, SAS invocation
Category: Environment control: Display
Restriction: This option is valid only in the SAS windowing environment.
PROC OPTIONS GROUP= ENVDISPLAY

Syntax

DMSOUTSIZE= *n* | *n*K | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K**
specifies the maximum number of rows that can be displayed in the SAS windowing environment Output window in multiples of 1 (n) or 1,024 (nK). For example, a value of 800 specifies 800 rows, and a value of 3K specifies 3,072 rows. Valid values range from 500 to 999999. The default is 99999.

***hex*X**
specifies the maximum number of rows that can be displayed in the SAS windowing environment Output window as a hexadecimal value. You must specify the value beginning with a number (0-9), followed by an X. For example, **`2ffx`** specifies 767 rows and **`0A00x`** specifies 2,560 rows.

MIN
specifies to set the maximum number of rows that can be displayed in the SAS windowing environment Output window to 500.

MAX
specifies to set the maximum number of rows that can be displayed in the SAS windowing environment Output window to 999999.

Details

When the maximum number of rows have been displayed in the Output window, SAS prompts you to either file, print, save, or clear the Output window.

See Also

System Option:

"DMSLOGSIZE= System Option" on page 1888

DMSPGMLINESIZE= System Option

Specifies the maximum number of characters in a Program Editor line.

Valid in: configuration file, SAS invocation

Category: Environment control: Display

PROC OPTIONS GROUP= ENVDISPLAY

Syntax

DMSPGMLINESIZE= *n*

Syntax Description

n

specifies the maximum number of characters in a Program Editor line.

Default: 136

Range: 136–960

DMSSYNCHK System Option

In the SAS windowing environment, specifies whether to enable syntax check mode for DATA step and PROC step processing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

DMSSYNCHK | NODMSSYNCHK

Syntax Description

DMSSYNCHK
enables syntax check mode for statements that are submitted within the SAS windowing environment.

NODMSSYNCHK
does not enable syntax check mode for statements that are submitted within the SAS windowing environment.

Details

If a syntax or semantic error occurs in a DATA step after the DMSSYNCHK option is set, then SAS enters syntax check mode, which remains in effect from the point where SAS encountered the error to the end of the code that was submitted. After SAS enters syntax mode, all subsequent DATA step statements and PROC step statements are validated.

While in syntax check mode, only limited processing is performed. For a detailed explanation of syntax check mode, see "Syntax Check Mode" in the "Error Processing in SAS" section of *SAS Language Reference: Concepts*.

CAUTION:
Place the OPTIONS statement that enables DMSSYNCHK before the step for which you want it to take effect. If you place the OPTIONS statement inside a step, then DMSSYNCHK will not take effect until the beginning of the next step. △

If NODMSSYNCHK is in effect, SAS processes the remaining steps even if an error occurs in the previous step.

Comparisons

You use the DMSSYNCHK system option to validate syntax in an interactive session by using the SAS windowing environment. You use the SYNTAXCHECK system option to validate syntax in a non-interactive or batch SAS session. You can use the ERRORCHECK= option to specify the syntax check mode for the LIBNAME statement, the FILENAME statement, the %INCLUDE statement, and the LOCK statement in SAS/SHARE.

See Also

System options:

"ERRORCHECK= System Option" on page 1904
"SYNTAXCHECK System Option" on page 2031

"Error Processing" in *SAS Language Reference: Concepts*

DSNFERR System Option

When a SAS data set cannot be found, specifies whether SAS issues an error message.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

DSNFERR | NODSNFERR

Syntax Description

DSNFERR
specifies that SAS issue an error message and stop processing if a reference is made to a SAS data set that does not exist.

NODSNFERR
specifies that SAS ignore the error message and continue processing if a reference is made to a SAS data set that does not exist. The data set reference is treated as if _NULL_ had been specified.

Comparisons

- DSNFERR is similar to the BYERR system option, which issues an error message and stops processing if the SORT procedure attempts to sort a _NULL_ data set.
- DSNFERR is similar to the VNFERR system option, which sets the error flag for a missing variable when a _NULL_ data set is used.

See Also

System Options:

"BYERR System Option" on page 1855
"VNFERR System Option" on page 2054

DTRESET System Option

Specifies whether to update the date and time in the SAS log and in the procedure output file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

Syntax

DTRESET | NODTRESET

Syntax Description

DTRESET
specifies that SAS update the date and time in the titles of the SAS log and the procedure output file.

NODTRESET
specifies that SAS not update the date and time in the titles of the SAS log and the procedure output file.

Details

The DTRESET system option updates the date and time in the titles of the SAS log and the procedure output file. This update occurs when the page is being written. The smallest time increment that is reflected is minutes.

The DTRESET option is especially helpful in obtaining a more accurate date and time stamp when you run long SAS jobs.

When you use NODTRESET, SAS displays the date and time that the job originally started.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

DUPLEX System Option

Specifies whether duplex (two-sided) printing is enabled.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Restriction: This option is ignored if the printer does not support duplex (two-sided) printing.

Syntax

DUPLEX| NODUPLEX

Syntax Description

DUPLEX
specifies that duplex (two-sided) printing is enabled.

Interaction: When DUPLEX is selected, the setting of the BINDING= option determines how the paper is oriented before output is printed on the second side.

NODUPLEX
specifies that duplex (two-sided) printing is not enabled. This is the default.

Details

Note that duplex (two-sided) printing can be used only on printers that support duplex output.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings for some SAS system options might vary both by operating environment and by site. Option values might also vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

See Also

System Option:

"BINDING= System Option" on page 1849

For information about declaring an ODS printer destination, see the ODS PRINTER Statement in *SAS Output Delivery System: User's Guide*.

For information about SAS Universal Printing, see Printing with SAS in *SAS Language Reference: Concepts*.

ECHOAUTO System Option

Specifies whether the statements in the autoexec file are written to the SAS log as they are executed.

Valid in: configuration file, SAS invocation

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

Syntax

ECHOAUTO | NOECHOAUTO

Syntax Description

ECHOAUTO
specifies that the SAS statements in the autoexec file are written to the SAS log as they are executed.

Requirement: To print autoexec file statements in the SAS log, the SOURCE system option must be set.

NOECHOAUTO
specifies that SAS statements in the autoexec file are not written in the SAS log, even though they are executed.

Details

Regardless of the setting of this option, messages that result from errors in the autoexec files are printed in the SAS log.

See Also

System Option:

"SOURCE System Option" on page 2003

The SAS Log in *SAS Language Reference: Concepts*

EMAILAUTHPROTOCOL= System Option

Specifies the authentication protocol for SMTP E-mail.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Communications: Email

PROC OPTIONS GROUP= EMAIL

Syntax

EMAILAUTHPROTOCOL= NONE | LOGIN

Syntax Description

LOGIN
specifies that the LOGIN authentication protocol is used. For more information about the order of authentication, see "Sending E-Mail through SMTP" in *SAS Language Reference: Concepts*.

Note: When you specify LOGIN, you might also need to specify EMAILID and EMAILPW. If you omit EMAILID, SAS will look up your user ID and use it. If you omit EMAILPW, no password is used. △

NONE
specifies that no authentication protocol is used.

Comparisons

For the SMTP access method, use this option in conjunction with the EMAILID=, EMAILPW=, EMAILPORT, and EMAILHOST system options. EMAILID= provides the user name, EMAILPW= provides the password, EMAILPORT specifies the port to which the SMTP server is attached, EMAILHOST specifies the SMTP server that supports e-mail access for your site, and EMAILAUTHPROTOCOL= provides the protocol.

See Also

System Options:

"EMAILHOST= System Option" on page 1897
"EMAILID= System Option" on page 1898
"EMAILPORT System Option" on page 1899
"EMAILPW= System Option" on page 1900

EMAILFROM System Option

When sending e-mail by using SMTP, specifies whether the e-mail option FROM is required in either the FILE or FILENAME statement.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Communications: Email

PROC OPTIONS GROUP= EMAIL

Syntax

EMAILFROM | NOEMAILFROM

Syntax Description

EMAILFROM
specifies that the FROM e-mail option is required when sending e-mail by using either the FILE or FILENAME statements.

NOEMAILFROM
specifies that the FROM e-mail option is not required when sending e-mail by using either the FILE or FILENAME statements.

See Also

Statements:

"FILE Statement" on page 1503
"FILENAME Statement, EMAIL (SMTP) Access Method" on page 1532

EMAILHOST= System Option

Specifies one or more SMTP servers that support e-mail access.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Communications: Email
PROC OPTIONS GROUP= EMAIL

Syntax

EMAILHOST= *server*

EMAILHOST=(*'server-1' 'server-2'* <*…'server-n'*>)

Syntax Description

server
specifies one or more Simple Mail Transfer Protocol (SMTP) server domain names for your site.

Note: The system administrator for your site will provide this information. △

Range: The maximum number of characters that can be specified for SMTP servers is 1,024

Requirement: When more than one server name is specified, the list must be enclosed in parentheses and each server name must be enclosed in single or double quotation marks..

Details

When more than one SMTP server is specified, SAS attempts to connect to e-mail servers in the order that they are specified. E-mail is delivered to the first server that

SAS connects to. If SAS is not able to connect to any of the specified servers, the attempt to deliver e-mail fails and SAS returns an error.

Operating Environment Information: To enable the SMTP interface that SAS provides, you must also specify the EMAILSYS=SMTP system option. For information about EMAILSYS, see the documentation for your operating environment. △

Comparisons

For the SMTP access method, use this option in conjunction with the EMAILID=, EMAILAUTHPROTOCOL=, EMAILPORT, and EMAILPW system options. EMAILID= provides the user name, EMAILPW= provides the password, EMAILPORT specifies the port to which the SMTP server is attached, EMAILHOST specifies SMTP servers that supports e-mail access for your site, and EMAILAUTHPROTOCOL= provides the protocol.

See Also

System Option:

"EMAILAUTHPROTOCOL= System Option" on page 1895
"EMAILID= System Option" on page 1898
"EMAILPORT System Option" on page 1899
"EMAILPW= System Option" on page 1900

EMAILID= System Option

Identifies an e-mail sender by specifying either a logon ID, an e-mail profile, or an e-mail address.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Communications: Email

PROC OPTIONS GROUP= EMAIL

Syntax

EMAILID =*logonid* | *profile* | *emailaddress*

Syntax Description

logonid
specifies the logon ID for the user running SAS.

Maximum: The maximum number of characters is 32,000.

profile
see documentation for your e-mail system to determine the profile name.

email-address
specifies the fully qualified e-mail address of the user running SAS.

Requirement: The e-mail address is valid only when SMTP is enabled.

Requirement: If the value of *email-address* contains a space, you must enclose it in double quotation marks.

Details

The EMAILID= system option specifies the logon ID, profile, or e-mail address to use with your e-mail system.

Comparisons

For the SMTP access method, use this option in conjunction with the EMAILAUTHPROTOCOL=, EMAILPW=, EMAILPORT, and EMAILHOST system options. EMAILID= provides the user name, EMAILPW= provides the password, EMAILPORT specifies the port to which the SMTP server is attached, EMAILHOST specifies the SMTP server that supports e-mail access for your site, and EMAILAUTHPROTOCOL= provides the protocol.

See Also

System Options:

"EMAILAUTHPROTOCOL= System Option" on page 1895
"EMAILHOST= System Option" on page 1897
"EMAILPORT System Option" on page 1899
"EMAILPW= System Option" on page 1900

EMAILPORT System Option

Specifies the port that the SMTP server is attached to.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Communications: Email

PROC OPTIONS GROUP= EMAIL

Syntax

EMAILPORT <*port-number*>

Syntax Description

port-number
specifies the port number that is used by the SMTP server that you specified on the EMAILHOST option.

Note: The system administrator for your site will provide this information. △

Details

Operating Environment Information: If you use the SMTP protocol that SAS provides, you must also specify the EMAILSYS SMTP system option. For information about EMAILSYS, see the documentation for your operating environment. △

Comparisons

For the SMTP access method, use this option in conjunction with the EMAILID=, EMAILAUTHPROTOCOL= , EMAILPW= , and EMAILHOST system options. EMAILID= provides the user name, EMAILPW= provides the password, EMAILPORT specifies the port to which the SMTP server is attached, EMAILHOST specifies the SMTP server that supports e–mail access for your site, and EMAILAUTHPROTOCOL= provides the protocol.

See Also

System Option:

"EMAILAUTHPROTOCOL= System Option" on page 1895
"EMAILHOST= System Option" on page 1897
"EMAILID= System Option" on page 1898
"EMAILPW= System Option" on page 1900

EMAILPW= System Option

Specifies an e-mail logon password.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Communications: Email

PROC OPTIONS GROUP= EMAIL

Syntax

EMAILPW= "*password*"

Syntax Description

PASSWORD

specifies the logon password for your logon name.

Restriction: If "*password*" contains a space, you must enclose the value in double quotation marks.

Details

You can use encoded e-mail passwords. When a password is encoded with PROC PWENCODE, the output string includes a tag that identifies the string as having been

encoded. An example of a tag is {sas001}. The tag indicates the encoding method. Encoding a password enables you to avoid e-mail access authentication with a password in plaintext. Passwords that start with "{sas" trigger an attempt to be decoded. If the decoding succeeds, then that decoded password is used. If the decoding fails, then the password is used as is. For more information, see PROC PWENCODE in the *Base SAS Procedures Guide*.

Operating Environment Information: In the Windows operating system, SAS will prompt you for an e–mail ID and a password if the EMAILSYS system option is set to MAPI or VIM, or if you do not specify the EMAILID and EMAILPW system options at invocation, or if you are not otherwise logged on to your e-mail system. If the EMAILSYS system option is set to SMTP, SAS will not prompt you for an e–mail ID and a password. △

Comparisons

For the SMTP access method, use this option in conjunction with the EMAILID=, EMAILAUTHPROTOCOL=, EMAILPORT, and EMAILHOST system options. EMAILID= provides the user name, EMAILPW= provides the password, EMAILPORT specifies the port to which the SMTP server is attached, EMAILHOST specifies the SMTP server that supports e-mail access for your site, and EMAILAUTHPROTOCOL= provides the protocol.

See Also

System Options:

"EMAILAUTHPROTOCOL= System Option" on page 1895

"EMAILHOST= System Option" on page 1897

"EMAILID= System Option" on page 1898

"EMAILPORT System Option" on page 1899

ENGINE= System Option

Specifies the default access method for SAS libraries.

Valid in: configuration file, SAS invocation

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

See: ENGINE= System Option in the documentation for your operating environment.

Syntax

ENGINE=*engine-name*

Syntax Description

engine-name
specifies an engine name.

Details

The ENGINE= system option specifies which default engine name is associated with a SAS library. The default engine is used when a SAS library points to an empty directory or a new file. The default engine is also used on directory-based systems, which can store more than one SAS file type within a directory. For example, some operating environments can store SAS files from multiple versions in the same directory.

Operating Environment Information: Valid engine names depend on your operating environment. For details, see the SAS documentation for your operating environment. △

See Also

"SAS I/O Engines" in *SAS Language Reference: Concepts*

ERRORABEND System Option

Specifies whether SAS responds to errors by terminating.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Alias: ERRABEND | NOERRABEND

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

ERRORABEND | NOERRORABEND

Syntax Description

ERRORABEND
specifies that SAS terminate for most errors (including syntax errors and file not found errors) that would normally cause it to issue an error message, set OBS=0, and go into syntax-check mode (if syntax checking is enabled). SAS also terminates if an error occurs in any global statement other than the LIBNAME and FILENAME statements.

Tip: Use the ERRORABEND system option with SAS production programs, which presumably should not encounter any errors. If errors are encountered and ERRORABEND is in effect, SAS brings the errors to your attention immediately

by terminating. ERRORABEND does not affect how SAS handles notes such as invalid data messages.

NOERRORABEND
specifies that SAS handle errors normally, that is, issue an error message, set OBS=0, and go into syntax-check mode (if syntax checking is enabled).

See Also

System options:

"ERRORBYABEND System Option" on page 1903

"ERRORCHECK= System Option" on page 1904

"Global Statements" on page 1434

ERRORBYABEND System Option

Specifies whether SAS ends a program when an error occurs in BY-group processing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

ERRORBYABEND | NOERRORBYABEND

Syntax Description

ERRORBYABEND
specifies that SAS ends a program for BY-group error conditions that would normally cause it to issue an error message.

NOERRORBYABEND
specifies that SAS handle BY-group errors normally, that is, by issuing an error message and continuing processing.

Details

If SAS encounters one or more BY-group errors while ERRORBYABEND is in effect, SAS brings the errors to your attention immediately by ending your program. ERRORBYABEND does not affect how SAS handles notes that are written to the SAS log.

Note: Use the ERRORBYABEND system option with SAS production programs that should be error free. △

See Also

System Option:

"ERRORABEND System Option" on page 1902

ERRORCHECK= System Option

Specifies whether SAS enters syntax-check mode when errors are found in the LIBNAME, FILENAME, %INCLUDE, and LOCK statements.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

ERRORCHECK=NORMAL | STRICT

Syntax Description

NORMAL
specifies not to place the SAS program into syntax-check mode when an error occurs in a LIBNAME or FILENAME statement, or in a LOCK statement in SAS/SHARE software. In addition, the program or session does not terminate when a %INCLUDE statement fails due to a non-existent file.

STRICT
specifies to place the SAS program into syntax-check mode when an error occurs in a LIBNAME or FILENAME statement, or in a LOCK statement in SAS/SHARE software. If the ERRORABEND system option is set and an error occurs in either a LIBNAME or FILENAME statement, SAS terminates. In addition, SAS terminates when a %INCLUDE statement fails due to a non-existent file.

See Also

System option:

"ERRORABEND System Option" on page 1902

ERRORS= System Option

Specifies the maximum number of observations for which SAS issues complete error messages.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling
Log and procedure output control: SAS log

PROC OPTIONS GROUP= ERRORHANDLING
LOGCONTROL

Syntax

ERRORS=*n* | *n*K | *n*M | *n*G | *n*T | MIN | MAX | *hex*X

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the number of observations for which SAS issues error messages in terms of 1 (n); 1,024 (nK); 1,048,576 (nM); 1,073,741,824 (nG); or 1,099,511,627,776 (nT). For example, a value of **8** specifies eight observations, and a value of **3M** specifies 3,145,728 observations.

MIN
sets the number of observations for which SAS issues error messages to 0.

MAX
sets the maximum number of observations for which SAS issues error messages to the largest signed, 4–byte integer representable in your operating environment.

***hex*X**
specifies the maximum number of observations for which SAS issues error messages as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **2dx** sets the maximum number of observations for which SAS issues error messages to 45 observations.

Details

If data errors are detected in more than *n* observations, processing continues, but SAS does not issue error messages for the additional errors.

Note: If you set ERRORS=0 and an error occurs, or if the maximum number of errors has been reached, a warning message displays in the log which states that the limit set by the ERRORS option has been reached. △

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

EXPLORER System Option

Specifies whether to invoke the SAS windowing environment and display only the Explorer window.

Valid in: configuration file, SAS invocation
Category: Environment control: Initialization and operation
PROC OPTIONS GROUP= EXECMODES

Syntax

EXPLORER | NOEXPLORER

Syntax Description

EXPLORER
specifies that the SAS session be invoked with only the Explorer window.

NOEXPLORER
specifies that the SAS session be invoked without the Explorer window.

Details

The following SAS execution mode invocation options, in order, have precedence over this option:

1. OBJECTSERVER.
2. DMR
3. SYSIN

If you specify EXPLORER with another execution mode invocation option of equal precedence, SAS uses only the last option listed. See “Order of Precedence” on page 1829 for more information about invocation option precedence.

See Also

System Options:

“DMS System Option” on page 1886

“DMSEXP System Option” on page 1887

FILESYNC= System Option

Specifies when operating system buffers that contain contents of permanent SAS files are written to disk.

Valid in: configuration file, SAS invocation

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

See: FILESYNC= System Option in the documentation for your operating environment.

Syntax

FILESYNC= SAS | CLOSE | HOST | SAVE

Syntax Description

SAS
specifies that SAS requests the operating system to force buffered data to be written to disk when it is best for the integrity of the SAS file.

CLOSE
specifies that SAS requests the operating system to force buffered data to be written to disk when the SAS file is closed.

HOST
specifies that the operating system schedules when the buffered data for a SAS file is written to disk. This is the default.

SAVE
specifies that the buffers are written to disk when the SAS file is saved.

Details

By using the FILESYNC= system option, SAS can tell the operating system when to force data that is temporarily stored in operating system buffers to be written to disk. Only SAS files in a permanent SAS library are affected; files in a temporary library are not affected.

If you specify a value other than the default value of HOST, the following occurs:

- the length of time it takes to run a SAS job increases
- the small chance of loosing data in the event of a system failure is further reduced

Consult with your system administrator before you change the value of the FILESYNC= system option to a value other than the default value.

Operating Environment Information: Under z/OS, the FILESYNC= system option affects SAS files only in UNIX file system (UFS) libraries. For more information, see "FILESYNC= System Option" in *SAS Companion for z/OS* △

FIRSTOBS= System Option

Specifies the observation number or external file record that SAS processes first.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

FIRSTOBS= *n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the number of the first observation or external file record to process, with *n* being an integer. Using one of the letter notations results in multiplying the integer by a specific value. That is, specifying K (kilo) multiplies the integer by 1,024; M (mega) multiplies by 1,048,576 ; G (giga) multiplies by 1,073,741,824; or T (tera) multiplies by 1,099,511,627,776. For example, a value of **`8`** specifies the eighth observations or records, and a value of **`3m`** specifies observation or record 3,145,728.

***hex*X**
specifies the number of the first observation or the external file record to process as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`2dx`** specifies the 45th observation.

MIN
sets the number of the first observation or external file record to process to 1. This is the default.

MAX
sets the number of the first observation to process to the maximum number of observations in the data sets or records in the external file, up to the largest eight-byte, signed integer, which is 2^{63}-1, or approximately 9.2 quintillion observations.

Details

The FIRSTOBS= system option is valid for all steps for the duration of your current SAS session or until you change the setting. To affect any single SAS data set, use the FIRSTOBS= data set option.

You can apply FIRSTOBS= processing to WHERE processing. For details, see "Processing a Segment of Data That Is Conditionally Selected" in *SAS Language Reference: Concepts*.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the documentation for your operating environment. △

Comparisons

- You can override the FIRSTOBS= system option by using the FIRSTOBS= data set option and by using the FIRSTOBS= option as a part of the INFILE statement.
- While the FIRSTOBS= system option specifies a starting point for processing, the OBS= system option specifies an ending point. The two options are often used together to define a range of observations or records to be processed.

Examples

If you specify FIRSTOBS=50, SAS processes the 50th observation of the data set first.

This option applies to every input data set that is used in a program or a SAS process. In this example, SAS begins reading at the 11th observation in the data sets OLD, A, and B:

```
options firstobs=11;

data a;
   set old; /* 100 observations */
run;

data b;
   set a;
run;

data c;
   set b;
run;
```

Data set OLD has 100 observations, data set A has 90, B has 80, and C has 70. To avoid decreasing the number of observations in successive data sets, use the FIRSTOBS= data set option in the SET statement. You can also reset FIRSTOBS=1 between a DATA step and a PROC step.

See Also

Data Set Option:

"FIRSTOBS= Data Set Option" on page 25

Statement:

"INFILE Statement" on page 1591

System Option:

"OBS= System Option" on page 1948

FMTERR System Option

When a variable format cannot be found, specifies whether SAS generates an error or continues processing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

FMTERR | NOFMTERR

Syntax Description

FMTERR
specifies that when SAS cannot find a specified variable format, it generates an error message and does not allow default substitution to occur.

NOFMTERR
replaces missing formats with the *w*. or $*w*. default format, issues a note, and continues processing.

See Also

System Option:

"FMTSEARCH= System Option" on page 1910

FMTSEARCH= System Option

Specifies the order in which format catalogs are searched.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: FMTSEARCH= System Option under OpenVMS

Syntax

FMTSEARCH=(*catalog-specification-1... catalog-specification-n*)

Syntax Description

catalog-specification
searches format catalogs in the order listed, until the desired member is found. The value of *catalog-specification* can be either *libref* or *libref.catalog*. If only the libref is given, SAS assumes that FORMATS is the catalog name.

Details

The WORK.FORMATS catalog is always searched first, and the LIBRARY.FORMATS catalog is searched next, unless one of them appears in the FMTSEARCH= list.

If a catalog appears in the FMTSEARCH= list, the catalog is searched in the order in which it appears in the list. If a catalog in the list does not exist, that particular item is ignored and searching continues.

Operating Environment Information: Under the Windows, UNIX, and z/OS operating environments, you can use the APPEND or INSERT system options to add additional *catalog-specification*. For details, see the documentation for the APPEND and INSERT system options. △

Examples

If you specify FMTSEARCH=(ABC DEF.XYZ GHI), SAS searches for requested formats or informats in this order:

1. WORK.FORMATS
2. LIBRARY.FORMATS
3. ABC.FORMATS
4. DEF.XYZ
5. GHI.FORMATS.

If you specify FMTSEARCH=(ABC WORK LIBRARY) SAS searches in this order:

1. ABC.FORMATS
2. WORK.FORMATS
3. LIBRARY.FORMATS.

Because WORK appears in the FMTSEARCH list, WORK.FORMATS is not automatically searched first.

See Also

System Option:

"APPEND= System Option" on page 1844
"INSERT= System Option" on page 1929
"FMTERR System Option" on page 1910

FONTEMBEDDING System Option

Specifies whether font embedding is enabled in Universal Printer and SAS/GRAPH printing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

FONTEMBEDDING | NOFONTEMBEDDING

Syntax Description

FONTEMBEDDING
specifies to enable font embedding. This is the default.

NOFONTEMBEDDING
specifies to disable font embedding.

Details

When FONTEMBEDDING is set, fonts can be embedded, or included, in the output files that are created by the Universal Printer and SAS/GRAPH. Output files with embedded fonts do not rely on fonts being installed on the computer that is used to view or print the output file. Embedding fonts increases the file size.

When NOFONTEMBEDDING is set, the output files rely on the fonts being installed on the computer that is used to view or print the font.

When you print or create PostScript files, if the specified font is recognized by SAS but is not available on the printer, SAS substitutes the most similar, standard font in the output. For example, the Helvetica font would replace any occurrence of Albany AMT. This guarantees that the printer is capable of printing the text.

To determine which fonts will be substituted for a given printer, use the Print Setup window, the Registry Editor, or the REGEDIT procedure to display the Printer Setup properties. Under **Fonts**, any individual fonts that are listed will be recognized by the printer. All other fonts, including those that are available via a link in the SAS Registry, will be substituted in the document when the document is created.

FONTRENDERING= System Option

Specifies whether SAS/GRAPH devices that are based on the SASGDGIF, SASGDTIF, and SASGDIMG modules render fonts by using the operating system or by using the FreeType engine.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

FONTRENDERING=HOST_PIXELS | FREETYPE_POINTS

Syntax Description

HOST_PIXELS
specifies that fonts are rendered by the operating system and that font size is requested in pixels.

Operating Environment Information: On z/OS, HOST_PIXELS is not supported. If HOST_PIXELS is specified, SAS uses FREETYPE_POINTS as the value for this option. △

FREETYPE_POINTS
specifies that fonts are rendered by the FreeType engine and that font size is requested in points. This is the default.

Details

Use the FONTRENDERING= system option to specify how SAS/GRAPH devices that are based on the SASGDGIF, SASGDTIF, and SASGDIMG modules render fonts. When the operating system renders fonts, the font size is requested in pixels. When the FreeType engine renders fonts, the font size is requested in points.

Use the GDEVICE procedure to determine which module a SAS/GRAPH device uses:

```
proc gdevice c=sashelp.devices browse nofs;
   list devicename;
quit;
```

For example,

```
proc gdevice c=sashelp.devices browse nofs;
  list gif;
quit;
```

The following is partial output from the GDEVICE procedure output:

```
                                          GDEVICE procedure
                                  Listing from SASHELP.DEVICES - Entry GIF

Orig Driver: GIF                Module:   SASGDGIF  Model:     6031
Description: GIF File Format                             Type: EXPORT
*** Institute-supplied ***
Lrows:  43  Xmax:   8.333 IN    Hsize:    0.000 IN  Xpixels:        800
Lcols:  88  Ymax:   6.250 IN    Vsize:    0.000 IN  Ypixels:        600
Prows:   0                      Horigin:  0.000 IN
Pcols:   0                      Vorigin:  0.000 IN
Aspect:   0.000                 Rotate:
Driver query: Y                 Queued messages: N
```

The **Module** entry names the module used by the device.

See Also

"SAS/GRAPH Fonts" in *SAS/GRAPH: Reference, Second Edition*

FONTSLOC= System Option

Specifies the location of the fonts that are supplied by SAS; names the default font file location for registering fonts that use the FONTREG procedure.

Valid in: configuration file, SAS invocation

Category: Environment control: Display

PROC OPTIONS GROUP= ENVDISPLAY

See: FONTSLOC= System Option in the documentation for your operating environment

Syntax

FONTSLOC= "*location*"

Syntax Description

"location"
specifies a fileref or the location of the SAS fonts that are used during the SAS session.

Note: If "*location*" is a fileref, you *do not* need to enclose the value in quotation marks. △

FORMCHAR= System Option

Specifies the default output formatting characters.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LISTCONTROL

See: FORMCHAR= System Option in the documentation for your operating environment.

Syntax

FORMCHAR= *'formatting-characters'*

Syntax Description

'formatting-characters'
specifies any string or list of strings of characters up to 64 bytes long. If fewer than 64 bytes are specified, the string is padded with blanks on the right.

Tip: For consistent results when you move your document to different computers, issue the following OPTIONS statement before using ODS destinations other than the Listing destination:

```
options formchar="|----|+|---+=|-/\<>*";
```

Details

Formatting characters are used to construct tabular output outlines and dividers for various procedures, such as the FREQ, REPORT, and TABULATE procedures. If you omit formatting characters as an option in the procedure, the default specifications given in the FORMCHAR= system option are used. Note that you can also specify a hexadecimal character constant as a formatting character. When you use a hexadecimal constant with this option, SAS interprets the value of the hexadecimal constant as appropriate for your operating system.

Note: To ensure that row and column separators and boxed tabular reports are printed legibly when using the standard forms characters, you must use these resources:

- either the SAS Monospace or the SAS Monospace Bold font
- a printer that supports TrueType fonts

△

See Also

For further information about how Base SAS procedures use formatting characters, see the *Base SAS Procedures Guide*. For procedures in other products that use formatting characters, see the documentation for that product.

"Printing with SAS" in *SAS Language Reference: Concepts*

FORMDLIM= System Option

Specifies a character to delimit page breaks in SAS output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LISTCONTROL

Syntax

FORMDLIM=*'delimiting-character'*

Syntax Description

'delimiting-character'
specifies in quotation marks a character written to delimit pages. Normally, the delimit character is null, as in this statement:

```
options formdlim='';
```

Details

When the delimit character is null, a new physical page starts whenever a new page occurs. However, you can conserve paper by allowing multiple pages of output to appear on the same page. For example, this statement writes a line of dashes (- -) where normally a page break would occur:

```
options formdlim='-';
```

When a new page is to begin, SAS skips a single line, writes a line consisting of the dashes that are repeated across the page, and skips another single line. There is no skip to the top of a new physical page. Resetting FORMDLIM= to null causes physical pages to be written normally again.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

FORMS= System Option

If forms are used for printing, specifies the default form to use.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Display

PROC OPTIONS GROUP= ENVDISPLAY

Syntax

FORMS=*form-name*

Syntax Description

form-name
specifies the name of the form.

Tip: To create a customized form, use the FSFORM command in a windowing environment.

Details

The default form contains settings that control various aspects of interactive windowing output, including printer selection, text body, and margins. The FORMS= system option also customizes output from the PRINT command (when FORM= is omitted) or output from interactive windowing procedures.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

GSTYLE System Option

Specifies whether ODS styles can be used in the generation of graphs that are stored as GRSEG catalog entries.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Graphics: Driver settings
Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= GRAPHICS
ODSPRINT

Syntax

GSTYLE | NOGSTYLE

Syntax Description

GSTYLE
specifies that ODS styles can be used in the generation of graphs that are stored as GRSEG catalog entries. If no style is specified, the default style for the given output destination is used. This is the default.

NOGSTYLE
specifies to not use ODS styles in the generation of graphs that are stored as GRSEG catalog entries.

Tip: Use NOGSTYLE for compatibility of graphs generated before SAS 9.2.

Details

The GSTYLE system option affects only graphic output that is generated using GRSEGs. The GSTYLE option does not affect the use of ODS styles in graphs that are generated by the following means:

- Java device driver
- ActiveX device driver
- SAS/GRAPH statistical graphic procedures
- SAS/GRAPH template language
- ODS GRAPHICS ON statement

GWINDOW System Option

Specifies whether SAS displays SAS/GRAPH output in the GRAPH window.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Graphics: Driver settings

PROC OPTIONS GROUP= GRAPHICS

Syntax

GWINDOW | NOGWINDOW

Syntax Description

GWINDOW
displays SAS/GRAPH software output in the GRAPH window, if your site licenses SAS/GRAPH software and if your personal computer has graphics capability.

NOGWINDOW
displays graphics outside of the windowing environment.

HELPBROWSER= System Option

Specifies the browser to use for SAS Help and ODS output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Help

PROC OPTIONS GROUP= HELP

Syntax

HELPBROWSER=REMOTE | SAS

Syntax Description

REMOTE
specifies to use the remote browser for the Help. The location of the remote browser is determined by the HELPHOST and the HELPPORT system options. This is the default value for the OpenVMS, UNIX, z/OS, and Windows 64-bit operating environments.

SAS
specifies to use the SAS browser for the Help. This is the default for the Windows 32-bit operating environment.

See Also

System options:

"HELPHOST System Option" on page 1921

"HELPPORT= System Option" on page 1922

Viewing Output and Help in the SAS Remote Browser in *SAS Companion for OpenVMS on HP Integrity Servers*

Viewing Output and Help in the SAS Remote Browser in *SAS Companion for UNIX Environments*

Viewing Output and Help in the SAS Remote Browser in *SAS Companion for Windows*

Using the SAS Remote Browser in *SAS Companion for z/OS*

HELPENCMD System Option

Specifies whether SAS uses the English version or the translated version of the keyword list for the command–line Help.

Valid in: configuration file, SAS invocation

Category: Environment control: Files

PROC OPTIONS GROUP= HELP

Syntax

HELPENCMD | NOHELPENCMD

Syntax Description

HELPENCMD
specifies that SAS use the English version of the keyword list for the command-line help, although the index will still be displayed with translated keywords. This is the default.

NOHELPENCMD
specifies that SAS use the translated version of the keyword list for the command-line help, if a translated version exists.

Details

Set NOHELPENCMD if you want the command-line help to locate keywords by using the localized terms. By default, all terms on the command line will be read as English.

See Also

System Options:

HELPINDEX System Option in *SAS Companion for Windows* , *SAS Companion for UNIX Environments*, and *SAS Companion for OpenVMS on HP Integrity Servers*

HELPLOC System Option in *SAS Companion for Windows*, *SAS Companion for UNIX Environments*, *SAS Companion for OpenVMS on HP Integrity Servers*, and *SAS Companion for z/OS*

HELPTOC System Option in *SAS Companion for Windows*, *SAS Companion for UNIX Environments*, and *SAS Companion for OpenVMS on HP Integrity Servers*

HELPHOST System Option

Specifies the name of the computer where the remote browser is to send Help and ODS output.

Default: NULL

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Help

PROC OPTIONS GROUP= HELP

See: HELPHOST= System Option under OpenVMS UNIX Windowsz/OS

Syntax

HELPHOST="*host*"

"*host*"
specifies the name of the computer where the remote help is to be displayed. Quotation marks or parentheses are required. The maximum number of characters is 2,048.

Details

Operating Environment Information: If you do not specify the HELPHOST option, the location where SAS displays the Help depends on your operating environment. See the HELPHOST system option in the documentation for your operating environment. △

See Also

"HELPBROWSER= System Option" on page 1919

"HELPPORT= System Option" on page 1922

Viewing Output and Help in the SAS Remote Browser in the *SAS Companion for OpenVMS on HP Integrity Servers*

Viewing Output and Help in the SAS Remote Browser in the *SAS Companion for UNIX Environments*

Viewing Output and Help in the SAS Remote Browser in the *SAS Companion for Windows*

Using the SAS Remote Browser in the *SAS Companion for z/OS*

HELPPORT= System Option

Specifies the port number for the remote browser client.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Environment control: Help
PROC OPTIONS GROUP= HELP

Syntax

HELPPORT=*port-number*

port-number
specifies the port number for the SAS Remote Browser Server.
Range: 0–65535
Default: 0

Details

When HELPPORT is set to 0, SAS uses the default port number for the remote browser server.

See Also

"HELPBROWSER= System Option" on page 1919
"HELPHOST System Option" on page 1921
Viewing Output and Help in the SAS Remote Browser in the *SAS Companion for OpenVMS on HP Integrity Servers*
Viewing Output and Help in the SAS Remote Browser in the *SAS Companion for UNIX Environments*
Viewing Output and Help in the SAS Remote Browser in the *SAS Companion for Windows*
Using the SAS Remote Browser in the *SAS Companion for z/OS*

HTTPSERVERPORTMAX= System Option

Specifies the highest port number that can be used by the SAS HTTP server for remote browsing.

Valid in: configuration file, SAS invocation
Category: Communications: Networking and encryption
PROC OPTIONS GROUP= Communications

Syntax

HTTPSERVERPORTMAX=*max-port-number*

Syntax Description

max-port-number
specifies the highest port number that can be used by the SAS HTTP server for remote browsing.

Range: 0–65535

Default: 0

Details

Use the HTTPSERVERPORTMAX= and HTTPSERVERPORTMIN= system options to specify a range of port values that the remote browser HTTP server can use to dynamically assign a port number when a firewall is configured between SAS and the HTTP server.

See Also

System options:

"HTTPSERVERPORTMIN= System Option" on page 1923

HTTPSERVERPORTMIN= System Option

Specifies the lowest port number that can be used by the SAS HTTP server for remote browsing.

Valid in: configuration file, SAS invocation

Category: Communications: Networking and encryption

PROC OPTIONS GROUP= Communications

Syntax

HTTPSERVERPORTMIN=*min-port-number*

Syntax Description

min-port-number
specifies the lowest port number that can be used by the SAS HTTP server for remote browsing.

Range: 0–65535

Default: 0

Details

Use the HTTPSERVERPORTMIN and HTTPSERVERPORTMAX system options to specify a range of port values that the remote browser HTTP server can use to

dynamically assign a port number when a firewall is configured between SAS and the HTTP server.

See Also

System option:

"HTTPSERVERPORTMAX= System Option" on page 1922

IBUFNO= System Option

Specifies an optional number of extra buffers to be allocated for navigating an index file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Default: 0

Syntax

IBUFNO=*n* | *n*K | *n*M | *n*G | *n*T |*hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the number of extra index buffers to be allocated in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **8** specifies eight buffers, and a value of **3k** specifies 3,072 buffers.

Restriction: Maximum value is 10,000.

***hex*X**
specifies the number of extra index buffers as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **2dx** specifies 45 buffers.

MIN
sets the number of extra index buffers to 0. This is the default.

MAX
sets the maximum number of extra index buffers to 10,000.

Details

An index is an optional SAS file that you can create for a SAS data file in order to provide direct access to specific observations. The index file consists of entries that are organized into hierarchical levels, such as a tree structure, and connected by pointers. When an index is used to process a request, such as for WHERE processing, SAS does a

binary search on the index file and positions the index to the first entry that contains a qualified value. SAS uses the value's identifier to directly access the observation that contains the value. SAS requires memory for buffers when an index is actually used. The buffers are not required unless SAS uses the index, but they must be allocated in preparation for the index that is being used.

SAS automatically allocates a minimal number of buffers in order to navigate the index file. Typically, you do not need to specify extra buffers. However, using IBUFNO= to specify extra buffers could improve execution time by limiting the number of input/output operations that are required for a particular index file. However, the improvement in execution time comes at the expense of increased memory consumption.

Note: Whereas too few buffers allocated to the index file decrease performance, over allocation of index buffers creates performance problems as well. Experimentation is the best way to determine the optimal number of index buffers. For example, experiment with **`ibufno=3`**, then **`ibufno=4`**, and so on, until you find the least number of buffers that produces satisfactory performance results. △

See Also

"Understanding SAS Indexes" in *SAS Language Reference: Concepts*.

System Option:

"IBUFSIZE= System Option" on page 1925

IBUFSIZE= System Option

Specifies the buffer page size for an index file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Restriction: Specify a page size before the index file is created. After it is created, you cannot change the page size.

Syntax

IBUFSIZE=*n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**

specifies the page size to process in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **`8`** specifies 8 bytes, and a value of **`3k`** specifies 3,072 bytes.

The default is 0, which causes SAS to use the minimum optimal page size for the operating environment.

***hex*X**
specifies the page size as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`2dx`** sets the page size to 45 bytes.

MAX
sets the page size for an index file to the maximum possible number. For IBUFSIZE=, the value is 32,767 bytes.

Details

An index is an optional SAS file that you can create for a SAS data file in order to provide direct access to specific observations. The index file consists of entries that are organized into hierarchical levels, such as a tree structure, and connected by pointers. When an index is used to process a request, such as for WHERE processing, SAS does a search on the index file in order to rapidly locate the requested records.

Typically, you do not need to specify an index page size. However, the following situations could require a different page size:

- The page size affects the number of levels in the index. The more pages there are, the more levels in the index. The more levels, the longer the index search takes. Increasing the page size allows more index values to be stored on each page, thus reducing the number of pages (and the number of levels). The number of pages required for the index varies with the page size, the length of the index value, and the values themselves. The main resource that is saved when reducing levels in the index is I/O. If your application is experiencing a lot of I/O in the index file, increasing the page size might help. However, you must re-create the index file after increasing the page size.
- The index file structure requires a minimum of three index values to be stored on a page. If the length of an index value is very large, you might get an error message that the index could not be created because the page size is too small to hold three index values. Increasing the page size should eliminate the error.

Note: Experimentation is the best way to determine the optimal index page size. △

See Also

"Understanding SAS Indexes" in *SAS Language Reference: Concepts*.
"IBUFNO= System Option" on page 1924

INITCMD System Option

Specifies an application invocation command and optional SAS windowing environment or text editor commands that SAS executes before processing AUTOEXEC file during SAS invocation.

Valid in: configuration file, SAS invocation
Category: Environment control: Initialization and operation
PROC OPTIONS GROUP= EXECMODES

Syntax

INITCMD "*command-1 <windowing-command-n>*"

Syntax Description

command-1
specifies any SAS command that invokes an application window. Some valid values are:

AF

ANALYST

ASSIST

DESIGN

EIS

FORECAST

GRAPH

HELP

IMAGE

LAB

MINER

PHCLINICAL

PHKINETICS

PROJMAN

QUERY

RUNEIS

SQC

XADX.

Interaction: If you specify FORECAST for *command-1*, you cannot use *windowing-command-n*.

windowing-command-n
specifies a valid windowing command or text editor command. Separate multiple commands with semicolons. These commands are processed in sequence. If you use a windowing command that impacts flow, such as the BYE command, it might delay or prohibit processing.

Restriction: Do not use the *windowing-command-n* argument when you enter a command for an application that submits SAS statements or commands during initialization of the application, that is, during autoexec file initialization.

Details

The INITCMD system option suppresses the Log, Output, Program Editor, and Explorer windows when SAS starts so that application window is the first screen that you see. The suppressed windows do not appear, but you can activate them. You can use the ALTLOG option to direct log output for viewing. If windows are initiated by an autoexec file or the INITSTMT option, the window that is displayed by the INITCMD option is displayed last. When you exit an application that is invoked with the INITCMD option, your SAS session ends.

You can use the INITCMD option in a windowing environment only. Otherwise, the option is ignored and a warning message is issued. If *command-1* is not a valid command, the option is ignored and a warning message is issued.

The following SAS execution mode invocation options, in order, have precedence over this option:

1. OBJECTSERVER.
2. DMR
3. SYSIN

If you specify INITCMD with another execution mode invocation option of equal precedence, SAS uses only the last option listed. See "Order of Precedence" on page 1829 for more information about invocation option precedence.

Examples

```
INITCMD "AFA c=mylib.myapp.primary.frame dsname=a.b"
INITCMD "ASSIST; FSVIEW SASUSER.CLASS"
```

INITSTMT= System Option

Specifies a SAS statement to execute after any statements in the autoexec file and before any statements from the SYSIN= file.

Valid in: configuration file, SAS invocation

Alias: IS=

Category: Environment control: Initialization and operation

PROC OPTIONS GROUP= EXECMODES

See: INITSTMT= System Option under Windows OpenVMS

Syntax

INITSTMT=*'statement'*

Syntax Description

'statement'

specifies any SAS statement or statements.

Requirements: *statement* must be able to run on a step boundary.

Operating Environment Information: On the command line or in a configuration file, the syntax is specific to your operating environment. The SYSIN= system option might not be supported by your operating environment. For details, see the SAS documentation for your operating environment. △

Comparisons

INITSTMT= specifies the SAS statements to be executed at SAS initialization, and the TERMSTMT= system option specifies the SAS statements to be executed at SAS termination.

Examples

Here is an example of using this option on UNIX:

```
sas -initstmt '%put you have used the initstmt; data x; x=1;
   run;'
```

See Also

System Option:

"TERMSTMT= System Option" on page 2035

INSERT= System Option

Inserts the specified value as the first value of the specified system option.

Valid in: OPTIONS statement, SAS System Option Window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: INSERT= System Option in the documentation for your operating environment.

Syntax

INSERT=(*system-option-1=argument-1 system-option-n=argument-n*)

Syntax Description

system-option
can be CMPLIB, FMTSEARCH, MAPS, SASAUTOS, or SASSCRIPT.

argument
specifies a new value that you want as the first value of *system-option*.
argument can be any value that could be specified for *system-option* if *system-option* is set using the OPTIONS statement.

Details

If you specify a new value for the CMPLIB=, FMTSEARCH=, MAPS=, SASAUTOS=, or SASSCRIPT= system options, the new value replaces the value of the option. Instead of replacing the value, you can use the INSERT= system option to add an additional value to the option as the first value of the option.

Comparison

The INSERT= system option adds a new value to the beginning of the current value of the CMPLIB=, FMTSEARCH=, MAPS=, SASAUTOS=, or SASSCRIPT= system options. The APPEND= system option adds a new value to the end of one of these system options.

Examples

The following table shows the results of adding a value to the beginning of the FMTSEARCH= option value:

Current FMTSEARCH= Value	Value of INSERT= System Option	New FMTSEARCH= Value
(WORK LIBRARY)	(fmtsearch=(abc def))	(ABC DEF WORK LIBRARY)

See Also

System option:

"APPEND= System Option" on page 1844

INTERVALDS= System Option

Specifies one or more interval name and value pairs, where the value is a SAS data set that contains user-supplied holidays. The interval can be used as an argument to the INTNX and INTCK functions.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data processing

PROC OPTIONS GROUP= INPUTCONTROL

Requirement: The set of interval-value pairs must be enclosed in parentheses.

Syntax

INTERVALDS=(*interval-1=libref.dataset-name-1 <interval-n=libref.dataset-name-n>*)

Syntax Description

interval

specifies the name of an interval. The value of *interval* is the data set that is named in *libref.dataset-name*.

Requirement: When you specify multiple intervals, the interval name must not be the same as another interval.

libref.dataset-name

specifies the libref and the data set name of the file that contains user-supplied holidays.

Details

The INTCK and INTNX functions specify *interval* as the interval name in the function argument list to reference a data set that names user-supplied intervals.

The same *libref.dataset-name* can be assigned to different intervals. An error occurs when more than one *interval* of the same name is defined for the INTERVALDS system option.

Examples

This example assigns a single data set to an interval on the SAS command line or in a configuration file.

```
-intervalds (mycompany=mycompany.holidays)
```

The next example assigns multiple intervals using the OPTIONS statement. The intervals subsid1 and subsid2 are assigned the same libref and data set name.

```
options intervalds=(mycompany=mycompany.holidays subsid1=subsid.holidays
        subsid2=subsid.holidays);
```

See Also

Functions:

"INTCK Function" on page 833

"INTNX Function" on page 848

About Date and Time Intervals in *SAS Language Reference: Concepts*

INVALIDDATA= System Option

Specifies the value that SAS assigns to a variable when invalid numeric data is encountered.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

INVALIDDATA='*character*'

Syntax Description

'character'

specifies the value to be assigned, which can be a letter (A through Z, a through z), a period (.), or an underscore (_). The default value is a period.

Details

The INVALIDDATA= system option specifies the value that SAS is to assign to a variable when invalid numeric data is read with an INPUT statement or the INPUT function.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

JPEGQUALITY= System Option

Specifies the JPEG quality factor that determines the ratio of image quality to the level of compression for JPEG files produced by the SAS/GRAPH JPEG device driver.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Requirement: The DEVICE graphic option must be set to the SAS/GRAPH JPEG device driver.

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

JPEGQUALITY= *n* | **MIN** | **MAX**

Syntax Description

n
specifies an integer that indicates the JPEG quality factor. The quality of the image increases with larger numbers and decreases with smaller numbers. JPEG files are compressed less for higher-quality images. Therefore, the JPEG file size is greater for higher-quality images. For example, *n*=100 is completely uncompressed and the image quality is highest. When *n*=0, the image is produced at the maximum compression level with the lowest quality.

Range: 0–100

Default: 75

MIN
specifies to set the JPEG quality factor to 0, which has the lowest image quality and the greatest file compression.

MAX
specifies to set the JPEG quality factor to 100, which has the highest image quality with no file compression.

Details

The optimal quality value varies for each image. The default value of 75 is a good starting value that you can use to optimize the quality of an image within a compressed file. You can increase or decrease the value until you are satisfied with the image quality. Values between 50 and 95 produce the best quality images.

When the value is 24 or less, some viewers might not be able to display the JPEG file. When you create such a file, SAS writes the following caution to the SAS log:

```
Caution: quantization tables are too coarse for baseline JPEG.
```

See Also

Graph options:

DEVICE

LABEL System Option

Specifies whether SAS procedures can use labels with variables.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LISTCONTROL

Syntax

LABEL | NOLABEL

Syntax Description

LABEL
specifies that SAS procedures can use labels with variables. The LABEL system option must be in effect before the LABEL option of any procedure can be used.

NOLABEL
specifies that SAS procedures cannot use labels with variables. If NOLABEL is specified, the LABEL option of a procedure is ignored.

Details

A *label* is a string of up to 256 characters that can be written by certain procedures in place of the variable's name.

See Also

Data Set Option:

"LABEL= Data Set Option" on page 37

Statements:

"ODS PROCLABEL Statement" in *SAS Output Delivery System: User's Guide*.

LAST= System Option

Specifies the most recently created data set.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

LAST=*SAS-data-set*

Syntax Description

SAS-data-set
specifies a SAS data set name.

Restriction: No data set options are allowed.

Restriction: Use *libref.membername* or *membername* syntax, not a string that is enclosed in quotation marks, to specify a SAS data set name.

Note: You can use quotation marks in the *libref.membername* or *membername* syntax if the libref or member name is associated with a SAS/ACCESS engine that supports member names with syntax that requires quoting or name literal (n-literal) specification. For more information, see the SAS/ACCESS documentation. △

Details

By default, SAS automatically keeps track of the most recently created SAS data set. Use the _LAST_= system option to override the default.

LAST= is not allowed with data set options.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

LEFTMARGIN= System Option

Specifies the print margin for the left side of the page.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

LEFTMARGIN=*margin-size*<*margin-unit*>

Syntax Description

margin-size
specifies the size of the left print margin.

Restriction: The left margin should be small enough so that the left margin plus the right margin is less than the width of the paper.

Interactions: Changing the value of this option might result in changes to the value of the LINESIZE= system option.

<margin-unit>
specifies the units for margin-size. The margin-unit can be *in* for inches or *cm* for centimeters. <*margin-unit*> is saved as part of the value of the BOTTOMMARGIN system option whether or not it is specified.

Default: inches

Details

All margins have a minimum that is dependent on the printer and the paper size. The default value of the LEFTMARGIN system option is **`0.00 in`**.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*.

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

See Also

System Options:

"BOTTOMMARGIN= System Option" on page 1850
"RIGHTMARGIN= System Option" on page 1984
"TOPMARGIN= System Option" on page 2039

LINESIZE= System Option

Specifies the line size for the SAS log and for SAS procedure output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Alias: LS=

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

See: LINESIZE= System Option in the documentation for your operating environment.

Syntax

LINESIZE=*n* | MIN | MAX | *hex*X

Syntax Description

n
specifies the number of characters in a line.

MIN
sets the number of characters in a line to 64.

MAX
sets the number of characters in a line to 256.

***hex*X**
specifies the number of characters in a line as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`OFAx`** sets the line size of the SAS procedure output to 250.

Details

The LINESIZE= system option specifies the line size (printer line width) in characters for the SAS log and the SAS output that are used by the DATA step and procedures. The LINESIZE= system option affects the following output:

- the Output window for the ODS LISTING destination
- output produced for an ODS markup destination by a DATA step where the FILE statement destination is PRINT (the FILE PRINT ODS statement is not affected by the LINESIZE= system option)
- procedures that produce only characters that cannot be scaled, such as the PLOT procedure, the CALENDAR procedure, the TIMEPLOT procedure, the FORMS procedure, and the CHART procedure

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

LOGPARM= System Option

Specifies when SAS log files are opened, closed, and, in conjunction with the LOG= system option, how they are named.

Valid in: configuration file, SAS invocation

Restriction: LOGPARM= is valid only in line mode and in batch mode

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

See: LOGPARM= System Option in the documentation for your operating environment.

Syntax

LOGPARM=

"<**OPEN**= APPEND | REPLACE | REPLACEOLD>
<**ROLLOVER**= AUTO | NONE | SESSION | *n* | *n*K | *n*M | *n*G>
<**WRITE**= BUFFERED | IMMEDIATE>"

Syntax Description

OPEN=APPEND | REPLACE | REPLACEOLD
when a log file already exists, specifies how the contents of the existing file are treated.

APPEND
appends the log when opening an existing file. If the file does not already exist, a new file is created.

REPLACE
overwrites the current contents when opening an existing file. If the file does not already exist, a new file is created.

REPLACEOLD
replaces files that are more than one day old. If the file does not already exist, a new file is created.

Operating Environment Information: For z/OS, see the SAS documentation for your operating environment for limitations on the use of OPEN=REPLACEOLD. △

Default: REPLACE

ROLLOVER=AUTO|NONE|SESSION | *n* | *n*G | *n*M | *n*G
specifies when or if the SAS log "rolls over". That is, when the current log is closed and a new one is opened.

AUTO
causes an automatic "rollover" of the log when the directives in the value of the LOG= option change, that is, the current log is closed and a new log file is opened.

Interaction: The name of the new log file is determined by the value of the LOG= system option. If LOG= does not contain a directive, however, the name would never change, so the log would never roll over, even when ROLLOVER=AUTO.

NONE
specifies that rollover does not occur, even when a change occurs in the name that is specified with the LOG= option.

Interaction: If the LOG= value contains any directives, they do not resolve. For example, if Log="#b.log" is specified, the directive "#" does not resolve, and the name of the log file remains "#b.log".

SESSION
at the beginning of each SAS session, opens the log file, resolves directives that are specified in the LOG= system option, and uses its resolved value to name the new log file. During the course of the session, no rollover is performed.

n | *n*K | *n*M | *n*G
causes the log to rollover when the log reaches a specific size, stated in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); or 1,073,741,824 (gigabytes). When the log reaches the specified size, it is closed and renamed by appending "old" to the log filename, and if it exists, the lock file for a server log. For example, a filename of 2008Dec01.log would be renamed 2008Dec01old.log. A new log file is opened using the name specified in the LOG= option.

CAUTION:
Old log files can be overwritten. SAS maintains only one old log file with the same name as the open log file. If rollover occurs more than once, the old log file is overwritten. △

Restriction: The minimum log file size is 10K.

See also: "Log Filenames" in *SAS Language Reference: Concepts*

Default: NONE

Interaction: Rollover is triggered by a change in the value of the LOG= option.

Restriction: Rollover will not occur more often than once a minute.

See Also: LOG= system option under Windows, UNIX, z/OS

WRITE=BUFFERED | IMMEDIATE
specifies when content is written to the SAS log.

BUFFERED
writes content to the SAS log only when a buffer is full in order to increase efficiency.

IMMEDIATE
writes to the SAS log each time that statements are submitted that produce content for the SAS log. SAS does no buffering of log messages.

Default: BUFFERED

Tip: Under Windows, the buffered log contents are written periodically, using an interval that is specified by SAS.

Details

The LOGPARM= system option controls the opening and closing of SAS log files when SAS is operating in batch mode or in line mode. This option also controls the naming of new log files, in conjunction with the LOG= system option and the use of directives in the value of LOG=.

Using directives in the value of the LOG= system option enables you to control when logs are open and closed and how they are named, based on actual time events, such as time, month, and day of week.

Operating Environment Information: Under the Windows and UNIX operating environments, you can begin directives with either the % symbol or the # symbol, and use both symbols in the same directive. For example, -log=mylog%b#C.log.

Under z/OS, begin directives only with the # symbol. For example, -log=mylog#b#c.log.

Under OpenVMS, begin directives only with the % symbol. For example, -log=mylog%b%c.log. △

The following table contains a list of directives that are valid in LOG= values:

Table 7.4 Directives for Controlling the Name of SAS Log Files

Directive	Description	Range
%a or #a	Locale's abbreviated day of week	Sun–Sat
%A or #A	Locale's full day of week	Sunday–Saturday
%b or #b	Local's abbreviated month	Jan–Dec
%B or #B	Locale's full month	January–December
%C or #C	Century number	00–99
%d or #d	Day of the month	01–31
%H or #H	Hour	00–23
%j or #j	Julian day	001–366
%l or #l *	User name	alphanumeric string that is the name of the user that started SAS
%M or #M	Minutes	00–59
%m or #m	Month number	01–12
%n or #n	Current system node name (without domain name)	none
%p or #p *	Process ID	alphanumeric string that is the SAS session process ID
%s or #s	Seconds	00–59
%u or #u	Day of week	1= Monday–7=Sunday
%v or #v *	Unique identifier	alphanumeric string that creates a log filename that does not currently exist
%w or #w	Day of week	0=Sunday–6=Saturday

Directive	Description	Range
%W or #W	Week number (Monday as first day; all days in new year preceding first Monday are in week 00)	00–53
%y or #y	Year without century	00–99
%Y or #Y	Full year	1970–9999
%%	Percent escape writes a single percent sign in the log filename.	%
##	Pound escape writes a single pound sign in the log filename.	#

* Because %v, %l, and %p are not a time-based format, the log filename will never change after it has been generated. Therefore, the log will never roll over. In these situations,specifying ROLLOVER=AUTO is equivalent to specifying ROLLOVER=SESSION.

Operating Environment Information: See the SAS companion for z/OS for limitations on the length of the log filename under z/OS. △

Note: Directives that you specify in the LOG= system option are not the same as the conversion characters that you specify to format logging facility logs. Directives specify a format for a log name. Conversion characters specify a format for log messages. Directives and conversion characters that use the same characters might function differently. △

Note: If you start SAS in batch mode or server mode and the LOGCONFIGLOC= option is specified, logging is done by the SAS logging facility. The traditional SAS log option LOGPARM= is ignored. The traditional SAS log option LOG= is honored only when the %S{App.Log} conversion character is specified in the logging configuration file. For more information, see SAS Logging Facility in *SAS Logging: Configuration and Programming Reference*. △

Examples

Operating Environment Information: The LOGPARM= system option is executed when SAS is invoked. When you invoke SAS at your site, the form of the syntax is specific to your operating environment. See the SAS documentation for your operating environment for details. △

- *Rolling over the log at a certain time and using directives to name the log according to the time:* If this command is submitted at 9:43 AM, this example creates a log file called test0943.log, and the log rolls over each time the log filename changes. In this example, at 9:44 AM, the test0943.log file will be closed, and the test0944.log file will be opened.

```
sas -log "test%H%M.log" -logparm "rollover=auto"
```

- *Preventing log rollover but using directives to name the log:* For a SAS session that begins at 9:34 AM, this example creates a log file named test0934.log, and prevents the log file from rolling over:

```
sas -log "test%H%M.log" -logparm "rollover=session"
```

- *Preventing log rollover and preventing the resolution of directives:* This example creates a log file named test%H%M.log, ignores the directives, and prevents the log file from rolling over during the session:

```
sas -log "test%H%M.log" -logparm "rollover=none"
```

- *Creating log files with unique identifiers:* This example uses a unique identifier to create a log file with a unique name:

```
sas -log "test%v.log" -logparm "rollover=session"
```

  SAS replaces the directive %v with *process_ID*v*n*, where *process_ID* is a numeric process identifier that is determined by the operating system and *n* is an integer number, starting with 1. The letter v that is between *process_ID* and *n* is always a lowercase letter.

  For this example, *process_ID* is 3755. If the file does not already exist, SAS creates a log file with the name test3755v1.log. If test3755v1.log does exist, SAS attempts to create a log file by incrementing *n* by 1, and this process continues until SAS can generate a log file. For example, if the file test3755v1.log exists, SAS attempts to create the file test3755v2.log.

- *Naming a log file by the user that started SAS:* This example creates a log filename that contains the user name that started the SAS session:

```
sas -log "%l.log" -logparm "rollover=session";
```

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

LRECL= System Option

Specifies the default logical record length to use for reading and writing external files.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: External files

PROC OPTIONS GROUP= EXTFILES

Syntax

LRECL=*n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

n
specifies the logical record length in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **`32`** specifies 32 bytes, and a value of **`32k`** specifies 32,767 bytes.

Default: 256

Range: 1–32767

***hex*X**
specifies the logical record length as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`2dx`** sets the logical record length to 45 characters.

MIN
specifies a logical record length of 1.

MAX
specifies a logical record length of 32,767.

Details

The logical record length for reading or writing external files is first determined by the LRECL= option in the access method statement, function, or command that is used to read or write an individual file, or the DDName value in the z/OS operating environment. If the logical record length is not specified by any of these means, SAS uses the value that is specified by the LRECL= system option.

Use a value for the LRECL= system option that is not an arbitrary large value. Large values for this option can result in excessive use of memory, which can degrade performance.

Operating Environment Information: Under z/OS, the LRECL= system option is recognized only for reading and writing HFS files. △

MAPS= System Option

Specifies the location of the SAS library that contains SAS/GRAPH map data sets.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Graphics: Driver settings

PROC OPTIONS GROUP= GRAPHICS

See: MAPS= System Option in the documentation for your operating environment

Syntax

MAPS=*location-of-maps*

Syntax Description

location-of-maps
specifies either a physical path, an environment variable, or a libref to locate the SAS/GRAPH map data sets.

Restriction: If you specify a libref, you must specify the MAPS option in the configuration file.

Operating Environment Information: The syntax shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For more information, see the SAS documentation for your operating environment. △

Operating Environment Information: Under the Windows, UNIX, and z/OS operating environments, you can use the APPEND or INSERT system options to add additional *location-of-maps*. For more information, see the documentation for the APPEND and INSERT system options. △

See Also

System Options:

"APPEND= System Option" on page 1844

"INSERT= System Option" on page 1929

MERGENOBY System Option

Specifies the type of message that is issued when MERGE processing occurs without an associated BY statement.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

MERGENOBY= NOWARN |WARN | ERROR

Syntax Description

NOWARN
specifies that no warning message is issued. This is the default.

WARN
specifies that a warning message is issued.

ERROR
specifies that an error message is issued.

MISSING= System Option

Specifies the character to print for missing numeric values.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

Syntax

MISSING=<'>*character*<'>

Syntax Description

<'>character<'>
specifies the value to be printed. The value can be any character. Single or double quotation marks are optional. The period is the default.

Operating Environment Information: The syntax that is shown above applies to the OPTIONS statement. However, when you specify the MISSING= system option on the command line or in a configuration file, the syntax is specific to your operating environment and might include additional or alternate punctuation. For details, see the SAS documentation for your operating environment. △

Details

The MISSING= system option does not apply to special missing values such as .A and .Z.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

MSGLEVEL= System Option

Specifies the level of detail in messages that are written to the SAS log.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

Syntax

MSGLEVEL= N | I

Syntax Description

N
specifies to print notes, warnings, CEDA message, and error messages only. N is the default.

I
specifies to print additional notes pertaining to index usage, merge processing, and sort utilities, along with standard notes, warnings, CEDA message, and error messages.

Details

Some of the conditions under which the MSGLEVEL= system option applies are as follows:

- If MSGLEVEL=I, SAS writes informative messages to the SAS log about index processing. In general, when a WHERE expression is executed for a data set with indexes, the following information appears in the SAS log:
  - if an index is used, a message displays that specifies the name of the index
  - if an index is not used but one exists that could optimize at least one condition in the WHERE expression, messages provide suggestions that describe what you can do to influence SAS to use the index. For example, a message could suggest sorting the data set into index order or to specify more buffers.
  - a message displays the IDXWHERE= or IDXNAME= data set option value if the setting can affect index processing.
- If MSGLEVEL=I, SAS writes a warning to the SAS log when a MERGE statement would cause variables to be overwritten.
- If MSGLEVEL=I, SAS writes a message that indicates which sorting product was used.
- For informative messages about queries by an application to a SAS/SHARE server, MSGLEVEL=I must be set for the SAS session where the SAS/SHARE server is running. The messages are written to the SAS log for the SAS session that runs the SAS/SHARE server.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

MULTENVAPPL System Option

Specifies whether the fonts available in a SAS application font selector window lists only the SAS fonts that are available in all operating environments.

Valid in: configuration file, SAS invocation, OPTIONS statement

Category: Environment control: Initialization and operation

PROC OPTIONS GROUP= EXECMODES

Syntax

MULTENVAPPL | NOMULTENVAPPL

Syntax Description

MULTENVAPPL
specifies that an application font selector window list only the SAS fonts.

NOMULTENVAPPL
specifies that an application font selector window list only the operating environment fonts.

Details

The MULTENVAPPL system option enables applications that support a font selection window, such as SAS/AF, SAS/FSP, SAS/EIS, or SAS/GIS, to choose a SAS font that is supported in all operating environments. Choosing a SAS font ensures portability of applications across all operating environments.

When NOMULTENVAPPL is in effect, the application font selector window has available only the fonts that are specific to your operating environment. SAS might need to resize operating environment fonts, which could result in text that is difficult to read. If the application is ported to another environment and the font is not available, a font is selected by the operating environment.

NEWS= System Option

Specifies an external file that contains messages to be written to the SAS log, immediately after the header.

Valid in: configuration file, SAS invocation

Category: Environment control: Files
Log and procedure output control: SAS log

PROC OPTIONS GROUP= ENVFILES
LOGCONTROL

See: NEWS= System Option in the documentation for your operating environment.

Syntax

NEWS=*external-file*

Syntax Description

external-file
specifies an external file.

Operating Environment Information: A valid file specification and its syntax are specific to your operating environment. Although the syntax is generally consistent with the command line syntax of your operating environment, it might include additional or alternate punctuation. For details, see the SAS documentation for your operating environment. △

Details

The NEWS file can contain information for uses, including news items about SAS.

The contents of the NEWS file are written to the SAS log immediately after the SAS header.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

NOTES System Option

Specifies whether notes are written to the SAS log.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

Syntax

NOTES | NONOTES

Syntax Description

NOTES
specifies that SAS write notes to the SAS log.

NONOTES
specifies that SAS does not write notes to the SAS log. NONOTES does not suppress error and warning messages.

Details

You must specify NOTES for SAS programs that you send to SAS for problem determination and resolution.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

NUMBER System Option

Specified whether to print the page number in the title line of each page of SAS output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

Syntax

NUMBER | NONUMBER

Syntax Description

NUMBER
specifies that SAS print the page number on the first title line of each page of SAS output.

NONUMBER
specifies that SAS not print the page number on the first title line of each page of SAS output.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

OBS= System Option

Specifies the observation that is used to determine the last observation to process, or specifies the last record to process.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

See: OBS= System Option in the documentation for your operating environment

Syntax

OBS= *n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies a number to indicate when to stop processing, with *n* being an integer. Using one of the letter notations results in multiplying the integer by a specific value. That is, specifying K (kilo) multiplies the integer by 1,024; M (mega) multiplies by 1,048,576; G (giga) multiplies by 1,073,741,824; or T (tera) multiplies by 1,099,511,627,776. For example, a value of **`20`** specifies 20 observations or records, while a value of **`3m`** specifies 3,145,728 observations or records.

***hex*X**
specifies a number to indicate when to stop processing as a hexadecimal value. You must specify the value beginning with a number (0–9), followed by an X. For example, the hexadecimal value F8 must be specified as **`0F8x`** in order to specify the decimal equivalent of 248. The value **`2dx`** specifies the decimal equivalent of 45.

MIN
sets the number to 0 to indicate when to stop processing.

Interaction: If OBS=0 and the NOREPLACE option is in effect, then SAS can still take certain actions because it actually executes each DATA and PROC step in the program, using no observations. For example, SAS executes procedures, such as CONTENTS and DATASETS, that process libraries or SAS data sets. External files are also opened and closed. Therefore, even if you specify OBS=0, when your program writes to an external file with a PUT statement, an end-of-file mark is written, and any existing data in the file is deleted.

MAX
sets the number to indicate when to stop processing to the maximum number of observations or records, up to the largest eight-byte, signed integer, which is 2^{63}-1, or approximately 9.2 quintillion. This is the default.

Details

OBS= tells SAS when to stop processing observations or records. To determine when to stop processing, SAS uses the value for OBS= in a formula that includes the value for OBS= and the value for FIRSTOBS=. The formula is

`(obs - firstobs) + 1 = results`

For example, if OBS=10 and FIRSTOBS=1 (which is the default for FIRSTOBS=), the result is 10 observations or records, that is **`(10 - 1) + 1 = 10`**. If OBS=10 and FIRSTOBS=2, the result is nine observations or records, that is, **`(10 - 2) + 1 = 9`**.

OBS= is valid for all steps during your current SAS session or until you change the setting.

You can also use OBS= to control analysis of SAS data sets in PROC steps.

If SAS is processing a raw data file, OBS= specifies the last line of data to read. SAS counts a line of input data as one observation, even if the raw data for several SAS data set observations is on a single line.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Comparisons

- An OBS= specification from either a data set option or an INFILE statement option takes precedence over the OBS= system option.
- While the OBS= system option specifies an ending point for processing, the FIRSTOBS= system option specifies a starting point. The two options are often used together to define a range of observations to be processed.

Examples

Example 1: Using OBS= to Specify When to Stop Processing Observations This example illustrates the result of using OBS= to tell SAS when to stop processing observations. This example creates a SAS data set, executes the OPTIONS statement by specifying FIRSTOBS=2 and OBS=12, and executes the PRINT procedure. The result is 11 observations, that is, **`(12 - 2) + 1 = 11`**. The result of OBS= in this situation appears to be the observation number that SAS processes last, because the output starts with observation 2, and ends with observation 12, but this result is only a coincidence.

```
data Ages;
   input Name $ Age;
   datalines;
Miguel 53
Brad 27
Willie 69
Marc 50
Sylvia 40
Arun 25
Gary 40
Becky 51
Alma 39
Tom 62
Kris 66
Paul 60
Randy 43
Barbara 52
Virginia 72
;
options firstobs=2 obs=12;
proc print data=Ages;
run;
```

Output 7.4 PROC PRINT Output Using OBS= and FIRSTOBS=

```
                    The SAS System                          1

                       Obs    Name       Age

                         2    Brad        27
                         3    Willie      69
                         4    Marc        50
                         5    Sylvia      40
                         6    Arun        25
                         7    Gary        40
                         8    Becky       51
                         9    Alma        39
                        10    Tom         62
                        11    Kris        66
                        12    Paul        60
```

Example 2: Using OBS= with WHERE Processing This example illustrates the result of using OBS= along with WHERE processing. The example uses the data set that was created in Example 1, which contains 15 observations, and the example assumes a new SAS session with the defaults FIRSTOBS=1 and OBS=MAX.

First, here is the PRINT procedure with a WHERE statement. The subset of the data results in 12 observations:

```
proc print data=Ages;
   where Age LT 65;
run;
```

Output 7.5 PROC PRINT Output Using a WHERE Statement

```
                    The SAS System                          1

                       Obs    Name       Age

                         1    Miguel      53
                         2    Brad        27
                         4    Marc        50
                         5    Sylvia      40
                         6    Arun        25
                         7    Gary        40
                         8    Becky       51
                         9    Alma        39
                        10    Tom         62
                        12    Paul        60
                        13    Randy       43
                        14    Barbara     52
```

Executing the OPTIONS statement with OBS=10 and the PRINT procedure with the WHERE statement results in 10 observations, that is, **(10 – 1) + 1 = 10**. Note that with WHERE processing, SAS first subsets the data and then SAS applies OBS= to the subset.

```
options obs=10;
proc print data=Ages;
   where Age LT 65;
run;
```

Output 7.6 PROC PRINT Output Using a WHERE Statement and OBS=

```
                    The SAS System                          2

                         Obs    Name      Age

                           1    Miguel     53
                           2    Brad       27
                           4    Marc       50
                           5    Sylvia     40
                           6    Arun       25
                           7    Gary       40
                           8    Becky      51
                           9    Alma       39
                          10    Tom        62
                          12    Paul       60
```

The result of OBS= appears to be how many observations to process, because the output consists of 10 observations, ending with the observation number 12. However, the result is only a coincidence. If you apply FIRSTOBS=2 and OBS=10 to the subset, the result is nine observations, that is, **`(10 - 2) + 1 = 9`**. OBS= in this situation is neither the observation number to end with nor how many observations to process; the value is used in the formula to determine when to stop processing.

```
options firstobs=2 obs=10;
proc print data=Ages;
   where Age LT 65;
run;
```

Output 7.7 PROC PRINT Output Using WHERE Statement, OBS=, and FIRSTOBS=

```
                    The SAS System                          3

                         Obs    Name      Age

                           2    Brad       27
                           4    Marc       50
                           5    Sylvia     40
                           6    Arun       25
                           7    Gary       40
                           8    Becky      51
                           9    Alma       39
                          10    Tom        62
                          12    Paul       60
```

Example 3: Using OBS= When Observations Are Deleted This example illustrates the result of using OBS= for a data set that has deleted observations. The example uses the data set that was created in Example 1, with observation 6 deleted. The example also assumes a new SAS session with the defaults FIRSTOBS=1 and OBS=MAX.

First, here is PROC PRINT output of the modified file:

```
&proc print data=Ages;
run;
```

Output 7.8 PROC PRINT Output Showing Observation 6 Deleted

```
                    The SAS System                          1

                       Obs    Name        Age

                         1    Miguel       53
                         2    Brad         27
                         3    Willie       69
                         4    Marc         50
                         5    Sylvia       40
                         7    Gary         40
                         8    Becky        51
                         9    Alma         39
                        10    Tom          62
                        11    Kris         66
                        12    Paul         60
                        13    Randy        43
                        14    Barbara      52
                        15    Virginia     72
```

Executing the OPTIONS statement with OBS=12, then the PRINT procedure, results in 12 observations, that is, **`(12 - 1) + 1 = 12`**:

```
options obs=12;
proc print data=Ages;
run;
```

Output 7.9 PROC PRINT Output Using OBS=

```
                    The SAS System                          2

                       Obs    Name       Age

                         1    Miguel      53
                         2    Brad        27
                         3    Willie      69
                         4    Marc        50
                         5    Sylvia      40
                         7    Gary        40
                         8    Becky       51
                         9    Alma        39
                        10    Tom         62
                        11    Kris        66
                        12    Paul        60
                        13    Randy       43
```

The result of OBS= appears to be how many observations to process, because the output consists of 12 observations, ending with the observation number 13. However, if you apply FIRSTOBS=2 and OBS=12, the result is 11 observations, that is **`(12 - 2) + 1 = 11`**. OBS= in this situation is neither the observation number to end with nor how many observations to process; the value is used in the formula to determine when to stop processing.

```
options firstobs=2 obs=12;
proc print data=Ages;
run;
```

Output 7.10 PROC PRINT Output Using OBS= and FIRSTOBS=

```
                    The SAS System                          3

                        Obs    Name      Age

                          2    Brad       27
                          3    Willie     69
                          4    Marc       50
                          5    Sylvia     40
                          7    Gary       40
                          8    Becky      51
                          9    Alma       39
                         10    Tom        62
                         11    Kris       66
                         12    Paul       60
                         13    Randy      43
```

See Also

Data Set Options:

"FIRSTOBS= Data Set Option" on page 25
"OBS= Data Set Option" on page 39
"REPLACE= Data Set Option" on page 55

System Option:

"FIRSTOBS= System Option" on page 1908

ORIENTATION= System Option

Specifies the paper orientation to use when printing to a printer.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

ORIENTATION=PORTRAIT | LANDSCAPE | REVERSEPORTRAIT | REVERSELANDSCAPE

Syntax Description

PORTRAIT
specifies the paper orientation as portrait. This is the default.

LANDSCAPE
specifies the paper orientation as landscape.

REVERSEPORTRAIT
specifies the paper orientation as reverse portrait to enable printing on paper with prepunched holes. The reverse side of the page is printed upside down.

REVERSELANDSCAPE
specifies the paper orientation as reverse landscape to enable printing on paper with prepunched holes. The reverse side of the page is printed upside down.

Details

Changing the value of this option might result in changes to the values of the portable LINESIZE= and PAGESIZE= system options.

Operating Environment Information: Most SAS system options are initialized with default settings when you invoke SAS. However, the default settings for some SAS system options vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

OVP System Option

Specifies whether overprinting of error messages to make them bold, is enabled.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

Syntax

OVP | NOOVP

Syntax Description

OVP
specifies that overprinting of error messages is enabled.

NOOVP
specifies that overprinting of error messages is disabled. This is the default.

Details

When OVP is specified, error messages are emphasized when SAS overprints the error message two additional times with overprint characters.

When output is displayed to a monitor, OVP is overridden and is changed to NOOVP.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

PAGEBREAKINITIAL System Option

Specifies whether to begin the SAS log and procedure output files on a new page.

Valid in: configuration file, SAS invocation

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

See: PAGEBREAKINITIAL System Option in the documentation for your operating environment.

Syntax

PAGEBREAKINITIAL | NOPAGEBREAKINITIAL

Syntax Description

PAGEBREAKINITIAL
specifies to begin the SAS log and procedure output files on a new page.

NOPAGEBREAKINITIAL
specifies not to begin the SAS log and procedure output files on a new page. NOPAGEBREAKINITIAL is the default.

Details

The PAGEBREAKINITIAL option inserts a page break at the start of the SAS log and procedure output files.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

PAGENO= System Option

Resets the SAS output page number.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LISTCONTROL

See: PAGENO= System Option in the documentation for your operating environment.

Syntax

PAGENO=*n* | *n*K | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K**
specifies the page number in multiples of 1 (*n*); 1,024 (*n*K). For example, a value of **`8`** sets the page number to 8 and a value of **`3k`** sets the page number to 3,072.

***hex*X**
specifies the page number as a hexadecimal number. You must specify the value beginning with a number (0-9), followed by an X. For example, the value **`2dx`** sets the page number to 45.

MIN
sets the page number to the minimum number, 1.

MAX
specifies the maximum page number as the largest signed, four–byte integer that is representable in your operating environment.

Details

The PAGENO= system option specifies a beginning page number for the next page of output that SAS produces. Use PAGENO= to reset page numbering during a SAS session.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

PAGESIZE= System Option

Specifies the number of lines that compose a page of SAS output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Alias: PS=

Category: Log and procedure output control: SAS log and procedure output
Log and procedure output control: SAS log
Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LOG_LISTCONTROL
LISTCONTROL
LOGCONTROL

See: PAGESIZE= System Option in the documentation for your operating environment.

Syntax

PAGESIZE=*n* | *n*K | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K**
specifies the number of lines that compose a page in terms of lines (*n*)or units of 1,024 lines (*n*K).

hex
specifies the number of lines that compose a page as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`2dx`** sets the number of lines that compose a page to 45 lines.

MIN
sets the number of lines that compose a page to the minimum setting, 15.

MAX
sets the number of lines that compose a page to the maximum setting, 32,767.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, valid values and range vary with your operating environment. For details, see the SAS documentation for your operating environment. △

Details

The PAGESIZE= system option affects the following output:

- the Output window for the ODS LISTING destination
- the ODS markup destinations when the PRINT option is used in the FILE statement in a DATA step (the FILE PRINT ODS statement is not affected by the PAGESIZE= system option)
- procedures that produce characters that cannot be scaled, such as the PLOT procedure, the CALENDAR procedure, the TIMEPLOT procedure, the FORMS procedure, and the CHART procedure

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

PAPERDEST= System Option

Specifies the name of the output bin to receive printed output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Restriction: This option is ignored if the printer does not have multiple output bins.

Syntax

PAPERDEST=*printer-bin-name*

Syntax Description

printer-bin-name
specifies the bin to receive printed output.

Restriction: Maximum length is 200 characters.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*.

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

See Also

System Options:

"PAPERSIZE= System Option" on page 1960
"PAPERSOURCE= System Option" on page 1961
"PAPERTYPE= System Option" on page 1962

PAPERSIZE= System Option

Specifies the paper size to use for printing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Language control
Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= LANGUAGECONTROL
ODSPRINT

Syntax

PAPERSIZE=*paper_size_name* | *("width_value"* <,> *"height_value")* | *('width_value'*<,> *'height_value')* | *(width_value height_value)*

Syntax Description

paper_size_name
specifies a predefined paper size. The default is either LETTER or A4, depending on the locale.

Default: Letter

Valid Values: Refer to the Registry Editor, or use PROC REGISTRY to obtain a listing of supported paper sizes. Additional values can be added.

Requirement: When the name of a predefined paper size contains spaces, enclose the name in single or double quotation marks.

Restriction: The maximum length is 200 characters.

("width_value", "height_value")
specifies paper width and height as positive floating-point values.

Default: inches

Range: *in* or *cm* for width_value, height_value

Details

If you specify a predefined paper size or a custom size that is not supported by your printer, the printer default paper size is used. The printer default paper size is locale dependent and can be changed using the Page Setup dialog box.

Fields that specify values for paper sizes can either be separated by blanks or commas.

Note: Changing the value of this option can result in changes to the values of the portable LINESIZE= and PAGESIZE= system options. △

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options can vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see ODS statements in the *SAS Output Delivery System: User's Guide*.

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

Examples

The first OPTIONS statement sets a paper size value that is a paper size name from the SAS Registry. The second OPTIONS statement sets a specific width and height for a paper size.

```
options papersize="480x640 Pixels";
options papersize=("4.5" "7");
```

In the first example, quotation marks are required because there is a space in the name.

In the second example, quotation marks are not required. When no measurement units are specified, SAS writes the following warning to the SAS log:

```
WARNING: Units were not specified on the PAPERSIZE option. Inches will be used.
WARNING: Units were not specified on the PAPERSIZE option. Inches will be used.
```

You can avoid the warning message by adding the unit type, **in** or **cm**, to the value with no space separating the value and the unit type:

```
options papersize=(4.5in 7in);
```

See Also

System Options:

"PAPERDEST= System Option" on page 1959
"PAPERSOURCE= System Option" on page 1961
"PAPERTYPE= System Option" on page 1962

PAPERSOURCE= System Option

Specifies the name of the paper bin to use for printing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Restriction: This option is ignored if the printer does not have multiple input bins.

Syntax

PAPERSOURCE=*printer-bin-name*

Syntax Description

printer-bin-name
specifies the bin that sends paper to the printer.

Operating Environment Information: For instructions on how to specify a printer bin, see the SAS documentation for your operating environment. △

See Also

System Options:

"PAPERDEST= System Option" on page 1959

"PAPERSIZE= System Option" on page 1960

"PAPERTYPE= System Option" on page 1962

For information about declaring an ODS printer destination, see the ODS PRINTER statement in *SAS Output Delivery System: User's Guide*.

For information about SAS Universal Printing, see Printing with SAS in *SAS Language Reference: Concepts*.

PAPERTYPE= System Option

Specifies the type of paper to use for printing.

Valid in: configuration file, SAS invocation, OPTIONS statement SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

PAPERTYPE=*paper-type-string*

Syntax Description

paper-type-string
specifies the type of paper. Maximum length is 200.

Range: Values vary by printer, site, and operating environment.

Default: Values vary by site and operating environment.

Operating Environment Information: For instructions on how to specify the type of paper, see the SAS documentation for your operating environment. There is a very large number of possible values for this option. △

See Also

System Options:

"PAPERDEST= System Option" on page 1959

"PAPERSIZE= System Option" on page 1960

"PAPERSOURCE= System Option" on page 1961

For information about declaring an ODS printer destination, see the ODS PRINTER statement in *SAS Output Delivery System: User's Guide*

For information about SAS Universal Printing, see Printing with SAS in *SAS Language Reference: Concepts*.

PARM= System Option

Specifies a parameter string that is passed to an external program.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

Syntax

PARM=<'>*string*<'>

Syntax Description

<'>string<'>
specifies a character string that contains a parameter.

Examples

This statement passes the parameter X=2 to an external program:

```
options parm='x=2';
```

Operating Environment Information: Other methods of passing parameters to external programs depend on your operating environment and on whether you are running in interactive line mode or batch mode. For details, see the SAS documentation for your operating environment. △

PARMCARDS= System Option

Specifies the file reference to open when SAS encounters the PARMCARDS statement in a procedure.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: PARMCARDS= System Option in the documentation for your operating environment.

Syntax

PARMCARDS=*file-ref*

Syntax Description

file-ref
specifies the file reference to open.

Details

The PARMCARDS= system option specifies the file reference of a file that SAS opens when it encounters a PARMCARDS (or PARMCARDS4) statement in a procedure.

SAS writes all data lines after the PARMCARDS (or PARMCARDS4) statement to the file until it encounters a delimiter line of either one or four semicolons. The file is then closed and made available to the procedure to read. There is no parsing or macro expansion of the data lines.

Operating Environment Information: The syntax shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

PDFACCESS System Option

Specifies whether text and graphics from PDF documents can be read by screen readers for the visually impaired.

Requirement: Adobe Acrobat Reader or Professional 5.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFACCESS | NOPDFACCESS

Syntax Description

PDFACCESS
specifies that text and graphics from an ODS PDF document can be read by screen readers for the visually impaired.

NOPDFACCESS
specifies that text and graphics from an ODS PDF document cannot be read by screen readers for the visually impaired.

Details

When the PDFSECURITY system option is set to HIGH, SAS sets the PDFACCESS option. If the PDFSECURITY option is set to LOW or NONE, this option is not functional. When the PDFSECURITY option is set to NONE, screen readers can read PDF text and graphics.

The following document properties are set for this option:

Value of PDFACCESS	Value of PDFSECURITY	Document Properties
NOPDFACCESS	HIGH	**`Content Extraction for Accessibility`** is set to **`Not Allowed`**.

See Also

System Options:

"PDFSECURITY= System Option" on page 1975

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFASSEMBLY System Option

Specifies whether PDF documents can be assembled.

Requirement: Adobe Acrobat Reader or Professional 5.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFASSEMBLY | NOPDFASSEMBLY

Syntax Description

PDFASSEMBLY
specifies that PDF documents can be assembled.

NOPDFASSEMBLY
specifies that PDF documents cannot be assembled. This is the default.

Details

When a PDF document is assembled, pages can be rotated, inserted, and deleted, and bookmarks and thumbnail images can be added.

When the PDFSECURITY system option is set to HIGH, SAS sets the PDFASSEMBLY option. If the PDFSECURITY option is set to LOW or NONE, this option is not functional. When the PDFSECURITY option is set to NONE, PDF documents can be assembled.

See Also

System Options:

"PDFSECURITY= System Option" on page 1975

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFCOMMENT System Option

Specifies whether PDF document comments can be modified.

Requirement: Adobe Acrobat Reader or Professional 5.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFCOMMENT | NOPDFCOMMENT

Syntax Description

PDFCOMMENT
specifies that PDF document comments can be modified.

NOPDFCOMMENT
specifies that PDF document comments cannot be modified. This is the default.

Details

When the PDFSECURITY system option is set to either LOW or HIGH, SAS sets the PDFCOMMENT option. If the PDFSECURITY option is set to NONE, the PDFCOMMENT option is not functional and PDF document comments can be modified.

The following document properties are set for this option:

Value of PDFCOMMENT	Value of PDFSECURITY	Document Properties
NOPDFCOMMENT	LOW	**Commenting** is set to **Not Allowed** **Filling in of fields** is set to **Not Allowed**

When PDFSECURITY=LOW, the settings for the PDFCOMMENT and PDFFILLIN options are dependent on each other. A change in either of these options changes the other option to the similar setting. For example, if PDFSECURITY=LOW, and PDFCOMMENT and PDFFILLIN are set, and if the PDFCOMMENT setting is modified to NOPDFCOMMENT, then SAS sets NOPDFFILLIN. When PDFSECURITY=HIGH, PDFCOMMENT and PDFFILLIN can be set independently.

See Also

System Options:

"PDFFILLIN System Option" on page 1970

"PDFSECURITY= System Option" on page 1975

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFCONTENT System Option

Specifies whether the contents of a PDF document can be changed.

Requirement: Adobe Acrobat Reader or Professional 3.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFCONTENT | NOPDFCONTENT

Syntax Description

PDFCONTENT
specifies that the contents of a PDF document can be changed.

NOPDFCONTENT
specifies that the contents of a PDF document cannot be changed.

Details

When the PDFSECURITY option is set to either LOW or HIGH, SAS sets the PDFCONTENT option. If the PDFSECURITY option is set to NONE, this option is not functional and the PDF document can be changed.

The following document properties are set for this option:

Value of PDFCONTENT	Value of PDFSECURITY	Document Properties
PDFCONTENT	HIGH	**Page Extraction** and **Commenting** are set to **Not Allowed**.
NOPDFCONTENT	Not applicable	**Changing the Document** and **Document Assembly** are both set to **Not Allowed**.

See Also

System Options:

"PDFSECURITY= System Option" on page 1975

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFCOPY System Option

Specifies whether text and graphics from a PDF document can be copied.

Requirement: Adobe Acrobat Reader or Professional 3.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFCOPY | NOPDFCOPY

Syntax Description

PDFCOPY
specifies that text and graphics from a PDF document can be copied. This is the default.

NOPDFCOPY
specifies that text and graphics from a PDF document cannot be copied.

Details

When the PDFSECURITY system option is set to either LOW or HIGH, SAS sets the PDFCOPY option. If the PDFSECURITY option is set to NONE, this option is not functional and PDF documents can be copied.

The following document properties are set for this option:

Value of PDFCOPY	Value of PDFSECURITY	Document Properties
NOPDFCOPY	LOW	**Printing**, **Content Copying**, and **Content Copying for Accessibility** are set to **Allowed**. All other properties are set to **Not Allowed**.
NOPDFCOPY	HIGH	**Changing the Document**, **Document Assembly**, **Content Copying**, **Page Extraction**, and **Commenting** are set to **Not Allowed**.

See Also

System Options:

"PDFSECURITY= System Option" on page 1975

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFFILLIN System Option

Specifies whether PDF forms can be filled in.

Requirement: Adobe Acrobat Reader or Professional 5.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFFILLIN | NOPDFFILLIN

Syntax Description

PDFFILLIN
specifies that PDF forms can be filled in.

NOPDFFILLIN
specifies that PDF forms cannot be filled in.

Details

When the PDFSECURITY option is set to HIGH, SAS sets the PDFFILLIN option. If the PDFSECURITY option is set to LOW or NONE, this option is not functional. When the PDFSECURITY option is set to NONE, PDF forms can be filled in.

The following document properties are set for this option:

Value of PDFFILLIN	Value of PDFSECURITY	Document Properties
NOPDFFILLIN	LOW	**Changing the Document**, **Document Assembly**, **Page Extraction**, **Commenting**, **Filling of form fields**, **Signing**, and **Creation of Template Pages** are set to **Not Allowed**.

When PDFSECURITY=LOW, the settings for the PDFCOMMENT and PDFFILLIN options are dependent on each other. A change in either of these options changes the other option to the similar setting. For example, if PDFSECURITY=LOW, and PDFCOMMENT and PDFFILLIN are set, and if the PDFCOMMENT setting is modified to NOPDFCOMMENT, then SAS sets NOPDFFILLIN. When PDFSECURITY=HIGH, PDFCOMMENT and PDFFILLIN can be set independently.

See Also

System Options:

"PDFCOMMENT System Option" on page 1966

"PDFSECURITY= System Option" on page 1975

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFPAGELAYOUT= System Option

Specifies the page layout for PDF documents.

Requirement: Adobe Acrobat Reader or Professional 5.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFPAGELAYOUT= DEFAULT | SINGLEPAGE | CONTINUOUS | FACING | CONTINUOUSFACING

Syntax Description

DEFAULT
specifies to use the current page layout for Acrobat Reader. This is the default.

SINGLEPAGE
specifies to display one page at a time in the viewing area.

CONTINUOUS
specifies to display all document pages in the viewing area in a single column.

FACING
specifies to display only two pages in the viewing area, with the even pages on the left and the odd pages on the right.

Requirement: Acrobat Reader 5.0 or later version is required.

CONTINUOUSFACING
specifies to display all pages in the viewing area, two pages side by side. The even pages display on the left, and the odd pages display on the right.

See Also

System option:

" PDFPAGEVIEW= System Option" on page 1972

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFPAGEVIEW= System Option

Specifies the page viewing mode for PDF documents.

Requirement: Adobe Acrobat Reader or Professional 5.0 and later versions

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFPAGEVIEW= DEFAULT | ACTUAL | FITPAGE | FITWIDTH | FULLSCREEN

Syntax Description

DEFAULT
specifies to use the current page view setting for Acrobat Reader. This is the default.

ACTUAL
specifies to set the page view setting to 100%.

FITPAGE
specifies to view a page using the full extent of the viewing window, maintaining the height and width aspect ratio.

FITWIDTH
specifies to view a page using the full width of the viewing window. The height of the document is not scaled to fit the page.

FULLSCREEN
specifies to view a page using the full screen. This option disables the table of contents, bookmarks, and all other document access aids, such as accessing a specific page.

See Also

System option:

" PDFPAGEVIEW= System Option" on page 1972

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFPASSWORD= System Option

Specifies the password to use to open a PDF document and the password used by a PDF document owner.

Requirement: Adobe Acrobat Reader or Professional 3.0 and later versions

Alias: PDFPW

Valid in: configuration file, SAS invocation, OPTIONS statement

Category: Log and procedure output control: PDF
System administration: Security

PROC OPTIONS GROUP= PDF
SECURITY

Security

Syntax

PDFPASSWORD=(OPEN=<">*password*<"> | OPEN="" < <,>
OWNER=<">*password*<"> | OWNER="">)

PDFPASSWORD=(OWNER=<">*password*<"> | OWNER="" <<,>
OPEN=<">*password*<"> | OPEN="">)

PDFPASSWORD=(OPEN=<">*password*<"> | OPEN="")

PDFPASSWORD=(OWNER=<">*password*<"> | OWNER="")

Syntax Description

OPEN="*password*"
specifies the password to open a PDF document. The quotation marks are optional.

password
specifies a set of characters, up to 32 characters, that are used to validate that a user has permission to open a PDF document.

Restriction: The OPEN password must be different from the OWNER password.

OPEN=""
specifies to reset the password to open a PDF document to null. When the password is set to null, no password is necessary to open a PDF document. This is the default.

Restriction: Null values are not valid.

OWNER="*password*"
specifies the password for the PDF document owner. The quotation marks are optional.

password
specifies a set of characters, up to 32 characters, that are used to validate the owner of a PDF document.

Restriction: The OWNER password must be different from the OPEN password.

Restriction: Null values are not valid.

OWNER=""
specifies to reset the password used by a PDF document owner to null. When the password is set to null, the owner does not need a password for the PDF document. This is the default.

Restriction: Null values are not valid.

Details

You can set the PDFPASSWORD option at any time, but it is ignored until the PDFSECURITY system option is set to either LOW or HIGH. When the PDFSECURITY option is set to NONE, passwords for a PDF document are not needed.

See Also

System option:

" PDFPAGEVIEW= System Option" on page 1972

"PDFSECURITY= System Option" on page 1975

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFPRINT= System Option

Specifies the resolution to print PDF documents.

Requirement: Adobe Acrobat Reader or Professional 3.0 and later versions, depending on PDFPRINT setting

Valid in: configuration files, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: PDF

PROC OPTIONS GROUP= PDF

Syntax

PDFPRINT= HRES | LRES | NONE

Syntax Description

HRES
specifies to print PDF documents at the highest resolution available on the printer. This is the default for Acrobat Reader or Professional 5.0 and later versions.

Requirement: Acrobat Reader or Professional 5.0 and later versions.

LRES
specifies to print PDF documents at a lower resolution for draft-quality documents.

Requirement: Acrobat Reader or Professional 3.0 and later versions.

Restriction: PDFPRINT=LRES can be set only when the PDFSECURITY option is set to HIGH.

NONE
specifies the PDF documents have no print resolution.

Requirement: Any version of Acrobat Reader or Professional.

Restriction: PDFPRINT=NONE can be set only when the PDFSECURITY option is set to HIGH or LOW.

Details

When the PDFSECURITY option is set to NONE, PDF documents can be printed.

The following table shows the option settings for allowing high and low resolution printing:

Value of PDFPRINT	Value of PDFSECURITY	Printing Resolution Allowed
LRES	LOW	High resolution printing
LRES	HIGH	Low resolution (150 dpi) printing

See Also

System option:

" PDFPAGEVIEW= System Option" on page 1972

Securing ODS Generated PDF Files in *SAS Language Reference: Concepts*

PDFSECURITY= System Option

Specifies the printing permissions for PDF documents.

Requirements: Adobe Acrobat Reader or Professional 3.0 and later versions, unless otherwise noted.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Restriction: The PDFSECURITY option is valid for UNIX, Windows, and z/OS operating systems, but only in countries where importing encryption software is legal.

Category: Log and procedure output control: PDF
System administration: Security

PROC OPTIONS GROUP= PDF
SECURITY

Syntax

PDFSECURITY= HIGH | LOW | NONE

Syntax Description

HIGH
specifies that SAS encrypts PDF documents using a 128-bit encryption algorithm.

Requirements: When PDFSECURITY=HIGH, you must use Acrobat 5.0 or later version.

Interaction: At least one password must be set using the PDFPASSWORD= system option when PDFSECURITY=HIGH or LOW.

LOW
specifies that SAS encrypts PDF documents using a 40-bit encryption algorithm.

Interaction: At least one password must be set using the PDFPASSWORD= system option when PDFSECURITY=HIGH or LOW.

NONE
specifies that no encryption is performed on PDF documents. This is the default.

Details

The following table shows the PDF options that SAS sets when the PDFSECURITY option is set to HIGH, LOW, or NONE. When the PDFSECURITY option is set to NONE, there are no restrictions on PDF documents, and the PDF options are not functional.

Table 7.5 How SAS Sets PDF Options Values for the PDFSECURITY Settings

	PDFSECURITY Settings		
Option	**HIGH**	**LOW**	**NONE**
PDFACCESS	PDFACCESS	Not functional	Not functional
PDFASSEMBLY	PDFASSEMBLY	Not functional	Not functional
PDFCOMMENT	PDFCOMMENT	PDFCOMMENT	Not functional
PDFCONTENT	PDFCONTENT	PDFCONTENT	Not functional
PDFCOPY	PDFCOPY	PDFCOPY	Not functional
PDFFILLIN	PDFFILLIN	Not functional	Not functional
PDFPRINT	PRFPRINT=HRES	PDFPRINT=HRES	Not functional

See Also

System option:

"PDFACCESS System Option" on page 1964
"PDFASSEMBLY System Option" on page 1965
"PDFCOMMENT System Option" on page 1966
"PDFCONTENT System Option" on page 1967
"PDFCOPY System Option" on page 1968
"PDFFILLIN System Option" on page 1970
"PDFPASSWORD= System Option" on page 1972
"PDFPRINT= System Option" on page 1974

"Securing ODS Generated PDF Files" in *SAS Output Delivery System: User's Guide*

PRIMARYPROVIDERDOMAIN= System Option

Specifies the domain name of the primary authentication provider.

Valid in: configuration file, SAS invocation

Alias PRIMPD=
Category: Environment control: Initialization and operation
PROC OPTIONS GROUP= EXECMODES

Syntax

PRIMARYPROVIDERDOMAIN=*domain-name*

Syntax Description

domain-name
specifies the name of the domain that authenticates user names.

Requirement: If the domain name contains one or more spaces, the domain name must be enclosed in quotation marks.

Details

By default, users who log on to the SAS Metadata Server are authenticated by the operating system that hosts the SAS Metadata Server. You can specify an alternate authentication provider by using the AUTHPROVIDERDOMAIN= system option. User IDs that are verified by an alternate authentication provider must be in the format *user-ID*@*domain-name* (for example, user1@sas.com).

By specifying an authentication provider and a domain name that use the AUTHPROVIDERDOMAIN= and PRIMARYPROVIDERDOMAIN= system options, respectively, you enable users to log on to the SAS Metadata Server by using their usual user ID without using a domain-name suffix on the user ID. For example, by specifying the following system options, users who log on as *user-ID* or *user-ID*@mycompany.com can be verified by the authentication provider that is specified by the AUTHPROVIDERDOMAIN= system option:

```
-authproviderdomain ldap:mycompany
-primaryproviderdomain mycompany.com
```

If you specify the PRIMARYPROVIDERDOMAIN system option without specifying the AUTHPROVIDERDOMAIN system option, authentication is performed by the host provider.

Comparison

You use the AUTHPROVIDERDOMAIN system option to register and name your Active Directory provider or other LDAP provider. You use the PRIMARYPROVIDERDOMAIN system option to designate the primary authentication provider.

Examples

The following examples show the system options that you might use in a configuration file to define a primary authentication provider domain-name:

Active Directory

```
/* Environment variables that describe your Active Directory server  */
-set AD_HOST myhost
```

```
/* Define authentication provider  */
-authpd ADIR:mycomapny.com
-primpd mycompany.com
```

LDAP

```
/* Environment variables that describe your LDAP server  */
-set LDAP_HOST myhost
-set LDAP_BASE "ou=emp, o=us"

/* Define authentication provider */
-authpd LDAP:mycompany.com
-primpd mycompany.com
```

See Also

System option:

"AUTHPROVIDERDOMAIN System Option" on page 1846

AUTHSERVER System Option in the *SAS Companion for Windows*

"Direct LDAP Authentication" in the *SAS Intelligence Platform: Security Administration Guide*

PRINTERPATH= System Option

Specifies the name of a registered printer to use for Universal Printing.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Restriction: The PRINTERPATH= system option is ignored when the DEVICE= system option is set to the ActiveX or Java devices.

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

PRINTERPATH=(*'printer-name'* <fileref>)

Syntax Description

'printer-name'

must be one of the printers defined in the Registry Editor under **Core ▶ Printing ▶ Printers**

Requirement: When the *printer name* contains blanks, you must enclose it in quotation marks.

fileref
is an optional fileref. If a fileref is specified, it must be defined with a FILENAME statement or an external allocation. If a fileref is not specified, the default output destination can specify a printer in the Printer Setup dialog box, which you open by selecting **File ▶ Printer Setup**. Parentheses are required only when a *fileref* is specified.

Details

If the PRINTERPATH= option is not a null string, then Universal Printing will be used. If the PRINTERPATH= option does not specify a valid Universal Printing printer, then the default Universal Printer is used.

Comparisons

A related system option SYSPRINT specifies which operating system printer will be used for printing. PRINTERPATH= specifies which Universal Printing printer will be used for printing.

The operating system printer specified by the SYSPRINT option is used when PRINTERPATH="" (two double quotation marks with no space between them sets a null string).

Examples

The following example specifies an output destination that is different from the default:

```
options PRINTERPATH=(corelab out);
filename out 'your_file';
```

Operating Environment Information: In some operating environments, setting the PRINTERPATH= option might not change the setting of the PMENU print button, which might continue to use operating environment printing. See the SAS documentation for your operating environment for more information.

For additional information about declaring an ODS printer destination, see ODS statements in *SAS Output Delivery System: User's Guide*. △

For additional information about the SAS universal print facility, see "Printing with SAS" in *SAS Language Reference: Concepts*.

PRINTINIT System Option

Specifies whether to initialize the SAS procedure output file.

Valid in: configuration file, SAS invocation
Category: Log and procedure output control: Procedure output
PROC OPTIONS GROUP= LISTCONTROL
See: PRINTINIT System Option in the documentation for your operating environment.

Syntax

PRINTINIT | NOPRINTINIT

Syntax Description

PRINTINIT
specifies to initialize the SAS procedure output file and resets the file attributes.

Tip: Specifying PRINTINIT causes the SAS procedure output file to be cleared even when output is not generated.

NOPRINTINIT
specifies to preserve the existing procedure output file if no new output is generated. This is the default.

Tip: Specifying NOPRINTINIT causes the SAS procedure output file to be overwritten only when new output is generated.

Details

Operating Environment Information: The behavior of the PRINTINIT system option depends on your operating environment. For additional information, see the SAS documentation for your operating environment. △

PRINTMSGLIST System Option

Specifies whether to print all messages to the SAS log or to print only top-level messages to the SAS log.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

Syntax

PRINTMSGLIST | NOPRINTMSGLIST

Syntax Description

PRINTMSGLIST
specifies to print the entire list of messages to the SAS log. PRINTMSGLIST is the default.

NOPRINTMSGLIST
specifies to print only the top-level message to the SAS log.

Details

For Version 7 and later versions, the return code subsystem allows for lists of return codes. All of the messages in a list are related, in general, to a single error condition, but give different levels of information. This option enables you to see the entire list of messages or just the top-level message.

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

QUOTELENMAX System Option

If a quoted string exceeds the maximum length allowed, specifies whether SAS writes a warning message to the SAS log.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

QUOTELENMAX | NOQUOTELENMAX

Syntax Description

QUOTELENMAX
specifies that SAS write a warning message to the SAS log about the maximum length for strings in quotation marks.

NOQUOTELENMAX
specifies that SAS does not write a warning message to the SAS log about the maximum length for strings in quotation marks.

Details

If a string in quotation marks is too long, SAS writes the following warning to the SAS log:

```
WARNING 32-169: The quoted string currently being processed has become
                more than 262 characters long.  You may have unbalanced
                quotation marks.
```

If you are running a program that has long strings in quotation marks, and you do not want to see this warning, use the NOQUOTELENMAX system option to turn off the warning.

REPLACE System Option

Specifies whether permanently stored SAS data sets can be replaced.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

REPLACE | NOREPLACE

Syntax Description

REPLACE
specifies that a permanently stored SAS data set can be replaced with another SAS data set of the same name.

NOREPLACE
specifies that a permanently stored SAS data set cannot be replaced with another SAS data set of the same name, which prevents the accidental replacement of existing SAS data sets.

Details

This option has no effect on data sets in the WORK library, even if you use the WORKTERM= system option to store the WORK library files permanently.

Comparisons

The REPLACE= data set option overrides the REPLACE system option.

See Also

System Option:
"WORKTERM System Option" on page 2057

Data Set Option:
"REPLACE= Data Set Option" on page 55

REUSE= System Option

Specifies whether SAS reuses space when observations are added to a compressed SAS data set.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

REUSE=YES | NO

Syntax Description

YES
specifies to track free space and reuses it whenever observations are added to an existing compressed data set.

NO
specifies not to track free space. This is the default.

Details

If space is reused, observations that are added to the SAS data set are inserted wherever enough free space exists, instead of at the end of the SAS data set.

Specifying REUSE=NO results in less efficient usage of space if you delete or update many observations in a SAS data set. However, the APPEND procedure, the FSEDIT procedure, and other procedures that add observations to the SAS data set continue to add observations to the end of the data set, as they do for uncompressed SAS data sets.

You cannot change the REUSE= attribute of a compressed SAS data set after it is created. Space is tracked and reused in the compressed SAS data set according to the REUSE= value that was specified when the SAS data set was created, not when you add and delete observations. Even with REUSE=YES, the APPEND procedure will add observations at the end.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Comparisons

The REUSE= data set option overrides the REUSE= system option.

PERFORMANCE NOTE: When using COMPRESS=YES and REUSE=YES system options settings, observations cannot be addressed by observation number.

Note that REUSE=YES takes precedence over the POINTOBS=YES data set option setting.

See Also

System Option:

"COMPRESS= System Option" on page 1872

Data Set Options:

"COMPRESS= Data Set Option" on page 19
"REUSE= Data Set Option" on page 56

RIGHTMARGIN= System Option

Specifies the print margin for the right side of the page for output directed to an ODS printer destination.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

RIGHTMARGIN=*margin-size<margin-unit>*

Syntax Description

margin-size
specifies the size of the margin.

Restriction: The right margin should be small enough so that the left margin plus the right margin is less than the width of the paper.

Interactions: Changing the value of this option might result in changes to the value of the LINESIZE= system option.

<margin-unit>
specifies the units for margin-size. The margin-unit can be *in* for inches or *cm* for centimeters. *<margin-unit>* is saved as part of the value of the RIGHTMARGIN system option.

Default: inches

Details

All margins have a minimum that is dependent on the printer and the paper size. The default value of the RIGHTMARGIN system option is **`0.00 in`**.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see the ODS statements in *SAS Output Delivery System: User's Guide*

See Also

System Options:

"BOTTOMMARGIN= System Option" on page 1850
"LEFTMARGIN= System Option" on page 1934
"TOPMARGIN= System Option" on page 2039

RLANG System Option

Specifies whether SAS executes R language statements.

Valid in: configuration file, SAS invocation
Category: System administration: Security
PROC OPTIONS GROUP= SECURITY

Syntax

RLANG | NORLANG

Syntax Description

RLANG
specifies that SAS executes R language statements in operating environments that support the R language.

NORLANG
specifies that SAS is not to execute R language statements. This is the default value.

Details

If RLANG is specified and the R language is not supported in the operating environment, SAS writes a message to the SAS log. The message indicates that the R language is not supported and asks you to call SAS Technical Support. SAS Technical Support would like to track the operating environments where users would like SAS to execute R language statements, but the R language is not supported.

RSASUSER System Option

Specifies whether to open the SASUSER library for read access or read-write access.

Valid in: configuration file, SAS invocation

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: RSASUSER System Option in the documentation for your operating environment.

Syntax

RSASUSER | NORSASUSER

Syntax Description

RSASUSER
opens the SASUSER library in read-only mode.

NORSASUSER
opens the SASUSER library in read-write mode.

Details

The RSASUSER system option is useful for sites that use a single SASUSER library for all users and want to prevent users from modifying it. However, it is not useful when users use SAS/ASSIST software, because SAS/ASSIST requires writing to the SASUSER library.

Operating Environment Information: For network considerations about using the RSASUSER system option, see the SAS documentation for your operating environment. △

S= System Option

Specifies the length of statements on each line of a source statement and the length of data on lines that follow a DATALINES statement.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

S=*n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the length of statements and data in terms of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **`8`** specifies 8 bytes, and a value of **`3m`** specifies 3,145,728 bytes.

***hex*X**
specifies the length of statements and data as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`2dx`** sets the length of statements and data to 45.

MIN
sets the length of statements and data to 0.

MAX
sets the length of statements and data to 2,147,483,647.

Details

Input can be from either fixed-length or variable-length records. Both fixed-length and variable-length records can be sequenced or unsequenced. The location of the sequence numbers is determined by whether the file record format is fixed-length or variable-length.

SAS uses the value of S to determine whether to look for sequence numbers in the input, and to determine how to read the input:

Record Type	Value of S	SAS Looks for Sequence Numbers	How SAS Reads The Input
Fixed-length	S>0 or S=MAX	No	The value of S is used as the length of the source or data to be scanned and ignores everything beyond that length on each line.
Fixed-length	S=0 or S=MIN	Yes, at the end of the line of input.	SAS inspects the last *n* columns (where n is the value of the SEQ= system option) of the first sequence field. If those columns contain numbers, they are assumed to be sequence numbers and SAS ignores the last eight columns of each line. If the *n* columns contain non-digit characters, SAS reads the last eight columns as data columns.

Record Type	Value of S	SAS Looks for Sequence Numbers	How SAS Reads The Input
Variable-length	S>0 or S=MAX	No	The value of S is used as the starting column of the source or data to be scanned and ignores everything before that length on each line.
Variable-length	S=0 or S=MIN	Yes, at the beginning of each line of input.	SAS inspects the last *n* columns (where n is the value of the SEQ= system option) of the first sequence field. If those columns contain numbers, they are assumed to be sequence numbers and SAS ignores the first eight columns of each line. If the *n* columns contain non-digit characters, SAS reads the first eight columns as data columns.

Comparisons

The S= system option operates exactly like the S2= system option except that S2= controls input only from a %INCLUDE statement, an autoexec file, or an autocall macro file.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

See Also

System Options:

"S2= System Option" on page 1990
"S2V= System Option" on page 1993
"SEQ= System Option" on page 1996

S2= System Option

Specifies the length of statements on each line of a source statement from a %INCLUDE statement, an autoexec file, or an autocall macro file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

S2=S | *n* | *n*K | *n*M | *n*G | *n*T | MIN | MAX | *hex*X

Syntax Description

S
uses the current value of the S= system option to compute the record length of text that comes from a %INCLUDE statement, an autoexec file, or an autocall macro file.

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the length of the statements in a file that is specified in a %INCLUDE statement, an autoexec file, or an autocall macro file, in terms of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **`8`** specifies 8 bytes, and a value of **`3m`** specifies 3,145,728 bytes.

***hex*X**
specifies the length of statements as a hexadecimal number. You must specify the value beginning with a number (0 - 9), followed by an X. For example, the value **`2dx`** sets the length of statements to 45.

MIN
sets the length of statements and data to 0.

MAX
sets the length of statements and data to 2,147,483,647.

Details

Input can be from either fixed-length or variable-length records. Both fixed-length and variable-length records can be sequenced or unsequenced. The location of the sequence numbers is determined by whether the file record format is fixed-length or variable-length.

SAS uses the value of S2 to determine whether to look for sequence numbers in the input, and to determine how to read the input:

Record Type	Value of S2	SAS Looks for Sequence Numbers	How SAS Reads The Input
Fixed-length	S2>0 or S2=MAX	No	The value of S2 is used as the length of the source or data to be scanned and ignores everything beyond that length on each line.
Fixed-length	S2=0 or S2=MIN	Yes, at the end of the line of input.	SAS inspects the last *n* columns (where n is the value of the SEQ= system option) of the first sequence field. If those columns contain numbers, they are assumed to be sequence numbers and SAS ignores the last eight columns of each line. If the *n* columns contain non-digit characters, SAS reads the last eight columns as data columns.

Record Type	Value of S2	SAS Looks for Sequence Numbers	How SAS Reads The Input
Variable-length	S2>0 or S2=MAX	No	The value of S2 is used as the starting column of the source or data to be scanned and ignores everything before that length on each line.
Variable-length	S2=0 or S2=MIN	Yes, at the beginning of each line of input.	SAS inspects the last *n* columns (where n is the value of the SEQ= system option) of the first sequence field. If those columns contain numbers, they are assumed to be sequence numbers and SAS ignores the first eight columns of each line. If the *n* columns contain non-digit characters, SAS reads the first eight columns as data columns.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Comparisons

The S2= system option operates exactly like the S= system option except that the S2= option controls input from a %INCLUDE statement, an autoexec file, or an autocall macro file.

The S2= system option reads both fixed-length and variable-length record formats from a file specified in a %INCLUDE statement, an autoexec file, or an autocall macro file. The S2V= system option reads only a variable-length record format from a file specified in a%INCLUDE statement, an autoexec file, or an autocall macro file.

See Also

System Options:

"S= System Option" on page 1987
"S2V= System Option" on page 1993

"SEQ= System Option" on page 1996

S2V= System Option

Specifies the starting position to begin reading a file that is specified in a %INCLUDE statement, an autoexec file, or an autocall macro file with a variable length record format.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

S2V=S2 | S | *n* | *n*K | *n*M | *n*G | *n*T | MIN | MAX | *hex*X

Syntax Description

S2
specifies to use the current value of the S2= system option to compute the starting position of the variable-sized record to read from a %INCLUDE statement, an autoexec file, or an autocall macro file. This is the default.

S
specifies to use the current value of the S= system option to compute the starting position of the variable-sized record to read from a %INCLUDE statement, an autoexec file, or an autocall macro file.

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the starting position of the variable-length record to read that comes from a %INCLUDE statement, an autoexec file, or an autocall macro file, in terms of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **8** specifies 8 bytes, and a value of **3m** specifies 3,145,728 bytes.

MIN
sets the starting position of the variable-length record to read that comes from a %INCLUDE statement, an autoexec file, or an autocall macro, to **0**.

MAX
sets the starting position of the variable-length record to read that comes from a %INCLUDE statement, an autoexec file, or an autocall macro, to 2,147,483,647.

***hex*X**
specifies the starting position of the variable-length record to read that comes from a %INCLUDE statement, an autoexec file, or an autocall macro, as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X.

Details

Both the S2V= system option and the S2= system option specify the starting position for reading variable-sized record input from a %INCLUDE statement, an autoexec file, or an autocall macro file. When values for both options are specified, the value of the S2V= system option takes precedence over the value specified for the S2= system option.

Operating Environment Information: The syntax shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environments. For details, see the SAS documentation for your operating environment. △

Comparisons

The S2= system option specifies the starting position for reading both fixed-length and variable-length record formats for input from a %INCLUDE statement, an autoexec file, or an autocall macro file. The S2V= system option specifies the starting position for reading only variable-length record formats for input from a %INCLUDE statement, an autoexec file, or an autocall macro file.

See Also

System Options:

"S= System Option" on page 1987
"S2= System Option" on page 1990
"SEQ= System Option" on page 1996

SASHELP= System Option

Specifies the location of the SASHELP library.

Valid in: configuration file, SAS invocation
Category: Environment control: Files
PROC OPTIONS GROUP= ENVFILES
See: SASHELP= System Option in the documentation for your operating environment.

Syntax

SASHELP=*library-specification*

Syntax Description

library-specification
identifies an external library.

Details

The SASHELP= system option is set during the installation process and normally is not changed after installation.

Operating Environment Information: A valid external library specification is specific to your operating environment. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Operating Environment Information: Under the Windows, UNIX, and z/OS operating environments, you can use the APPEND or INSERT system options to add additional *library-specifications*. For more information, see the documentation for the APPEND and INSERT system options. △

See Also

System Options:

"APPEND= System Option" on page 1844

"INSERT= System Option" on page 1929

SASUSER= System Option

Specifies the SAS library to use as the SASUSER library.

Valid in: configuration file, SAS invocation

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: SASUSER= System Option in the documentation for your operating environment.

Syntax

SASUSER=*library-specification*

Syntax Description

library-specification
specifies the libref or the physical name that contains a user's Profile catalog.

Details

The library and catalog are created automatically by SAS; you do not have to create them explicitly.

Operating Environment Information: A valid library specification and its syntax are specific to your operating environment. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

SEQ= System Option

Specifies the length of the numeric portion of the sequence field in input source lines or data lines.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Input control: Data Processing
PROC OPTIONS GROUP= INPUTCONTROL

Syntax

SEQ=*n* | MIN | MAX | *hex*X

Syntax Description

n
specifies the length in terms of bytes.

MIN
sets the minimum length to 1.

MAX
sets the maximum length to 8.

Tip: When SEQ=8, all eight characters in the sequence field are assumed to be numeric.

***hex*X**
specifies the length as a hexadecimal. You must specify the value beginning with a number (0–9), followed by an X.

Details

Unless the S= or S2= system option specifies otherwise, SAS assumes an eight-character sequence field. However, some editors place some alphabetic information (for example, the filename) in the first several characters. The SEQ= value specifies the number of digits that are right-justified in the eight-character field. For example, if you specify SEQ=5 for the sequence field AAA00010, SAS looks at only the last five characters of the eight-character sequence field and, if the characters are numeric, treats the entire eight-character field as a sequence field.

See Also

System Options:
"S= System Option" on page 1987
"S2= System Option" on page 1990

SETINIT System Option

Specifies whether site license information can be altered.

Valid in: configuration file, SAS invocation
Category: System administration: Installation
PROC OPTIONS GROUP= INSTALL

Syntax

SETINIT | NOSETINIT

Syntax Description

SETINIT
in a non-windowing environment, specifies that you can change license information by running the SETINIT procedure.

NOSETINIT
specifies not to allow you to alter site license information after installation.

Details

SETINIT is set in the installation process and is not normally changed after installation. The SETINIT option is valid only in a non-windowing SAS session.

SKIP= System Option

Specifies the number of lines to skip at the top of each page of SAS output.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Log and procedure output control: Procedure output
PROC OPTIONS GROUP= LISTCONTROL

Syntax

SKIP=*n* | *hex*X | MIN | MAX

Syntax Description

n
specifies the range of lines to skip from 0 to 20.

MIN
sets the number of lines to skip to 0, so no lines are skipped.

MAX
sets the number of lines to skip to 20.

hex
specifies the number of lines to skip as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **`0ax`** specifies to skip 10 lines.

Details

The location of the first line is relative to the position established by carriage control or by the forms control buffer on the printer. Most sites define this position so that the first line of a new page begins three or four lines down the form. If this spacing is sufficient, specify SKIP=0 so that additional lines are not skipped.

The SKIP= value does not affect the maximum number of lines printed on each page, which is controlled by the PAGESIZE= system option.

SOLUTIONS System Option

Specifies whether the SOLUTIONS menu is included in SAS windows.

Valid in: configuration file, SAS invocation

Category: Environment control: Display

PROC OPTIONS GROUP= ENVDISPLAY

Syntax

SOLUTIONS | NOSOLUTIONS

Syntax Description

SOLUTIONS
specifies that the SOLUTIONS menu is included in SAS windows.

NOSOLUTIONS
specifies that the SOLUTIONS menu is not included in SAS windows.

SORTDUP= System Option

Specifies whether the SORT procedure removes duplicate variables based on all variables in a data set or the variables that remain after the DROP or KEEP data set options have been applied.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Sort: Procedure options

PROC OPTIONS GROUP= SORT

Syntax

SORTDUP=PHYSICAL | LOGICAL

Syntax Description

PHYSICAL
removes duplicates based on all the variables that are present in the data set. This is the default.

LOGICAL
removes duplicates based on only the variables remaining after the DROP= and KEEP= data set options are processed.

Details

The SORTDUP= option specifies what variables to sort to remove duplicate observations when the SORT procedure NODUPRECS option is specified.

When SORTDUP= is set to LOGICAL and NODUPRECS is specified in the SORT procedure, duplicate observations are removed based on the variables that remain after a DROP or KEEP operation on the input data set. Setting SORTDUP=LOGICAL increases the number of duplicate observations that are removed because it eliminates variables before observations are compared. Setting SORTDUP=LOGICAL might improve performance.

When SORTDUP= is set to PHYSICAL and NODUPRECS is specified in the SORT procedure, duplicate observations are removed based on all of the variables in the input data set.

See Also

The SORT Procedure in *Base SAS Procedures Guide*

SORTEQUALS System Option

Specifies whether observations in the output data set with identical BY variable values have a particular order.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System OPTIONS window

Category: Sort: Procedure options

PROC OPTIONS GROUP= SORT

Syntax

SORTEQUALS | NOSORTEQUALS

SORTEQUALS
specifies that observations with identical BY variable values are to retain the same relative positions in the output data set as in the input data set.

NOSORTEQUALS
specifies that no resources be used to control the order of observations with identical BY variable values in the output data set.

Interaction: To achieve the best sorting performance when using the THREADS= system option, specify THREADS=YES and NOSORTEQUALS.

Tip: To save resources, use NOSORTEQUALS when you do not need to maintain a specific order of observations with identical BY variable values.

Comparisons

The SORTEQUALS and NOSORTEQUALS system options set the sorting behavior of PROC SORT for your SAS session. The EQUAL or NOEQUAL option in the PROC SORT statement overrides the setting of the system option for an individual PROC step and specifies the sorting behavior for that PROC step only.

See Also

Statement Options:

EQUALS option for the PROC SORT statement in *Base SAS Procedures Guide*.

System Options:

"THREADS System Option" on page 2037

SORTSIZE= System Option

Specifies the amount of memory that is available to the SORT procedure.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Sort: Procedure options
System administration: Memory

PROC OPTIONS GROUP= MEMORY
SORT

See: SORTSIZE= System Option in the documentation for your operating environment.

Syntax

SORTSIZE=*n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the amount of memory in terms of 1 (byte); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **`4000`** specifies 4,000 bytes and a value of **`2m`** specifies 2,097,152 bytes. If *n*=0, the sort utility uses its default. Valid values for SORTSIZE range from 0 to 9,223,372,036,854,775,807.

***hex*X**
specifies the amount of memory as a hexadecimal number. This number must begin with a number (0-9), followed by an X. For example, **`0fffx`** specifies 4095 bytes of memory.

MIN
specifies the minimum amount of memory available.

MAX
specifies the maximum amount of memory available.

Operating Environment Information: Values for MIN and MAX will vary, depending on your operating environment. For details, see the SAS documentation for your operating environment △

Details

Generally, the value of the SORTSIZE= system option should be less than the physical memory available to your process. If the SORT procedure needs more memory than you specify, the system creates a temporary utility file.

PERFORMANCE NOTE: Proper specification of SORTSIZE= can improve sort performance by restricting the swapping of memory that is controlled by the operating environment.

See Also

System Option:

"SUMSIZE= System Option" on page 2017

"The SORT procedure" in the SAS documentation for your operating environment

SORTVALIDATE System Option

Specifies whether the SORT procedure verifies if a data set is sorted according to the variables in the BY statement when a user-specified sort order is denoted in the sort indicator.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Sort: Procedure options

PROC OPTIONS GROUP= SORT

Syntax

SORTVALIDATE | NOSORTVALIDATE

Syntax Description

SORTVALIDATE
specifies that the SORT procedure verifies if the observations in the data set are sorted by the variables specified in the BY statement.

NOSORTVALIDATE
specifies that the SORT procedure is not to verify if the observations in the data set are sorted. This is the default.

Details

You can use the SORTVALIDATE system option to specify whether the SORT procedure validates that a data set is sorted correctly when the data set sort indicator shows a user-specified sort order. The user can specify a sort order by using the SORTEDBY= data set option in a DATA statement or by using the SORTEDBY= option in the DATASETS procedure MODIFY statement. When the sort indicator is set by a user, SAS cannot be absolutely certain that a data set is sorted according to the variables in the BY statement.

If the SORTVALIDATE system option is set and the data set sort indicator was set by a user, the SORT procedure performs a sequence check on each observation to ensure that the data set is sorted according to the variables in the BY statement. If the data set is not sorted correctly, SAS sorts the data set.

At the end of a successful sequence check or at the end of a sort, the SORT procedure sets the **`Validated`** sort information to Yes. If a sort is performed, the SORT procedure updates the **`Sortedby`** sort information to the variables that are specified in the BY statement.

If an output data set is specified, the **`Validated`** sort information in the output data set is set to Yes. If no sort is necessary, the data set is copied to the output data set.

See Also

Data Set Option:

"SORTEDBY= Data Set Option" on page 57

Procedures:

"The DATASETS Procedure" in the *Base SAS Procedures Guide*

"The SORT Procedure" in the *Base SAS Procedures Guide*

"Sorted Data Sets" in *SAS Language Reference: Concepts*

SOURCE System Option

Specifies whether SAS writes source statements to the SAS log.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

Syntax

SOURCE | NOSOURCE

Syntax Description

SOURCE
specifies to write SAS source statements to the SAS log.

NOSOURCE
specifies not to write SAS source statements to the SAS log.

Details

The SOURCE system option does not affect whether statements from a file read with %INCLUDE or from an autocall macro are printed in the SAS log.

Note: SOURCE must be in effect when you execute SAS programs that you want to send to SAS for problem determination and resolution. △

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

SOURCE2 System Option

Specifies whether SAS writes secondary source statements from included files to the SAS log.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SAS log

PROC OPTIONS GROUP= LOGCONTROL

Syntax

SOURCE2 | NOSOURCE2

Syntax Description

SOURCE2
specifies to write to the SAS log secondary source statements from files that have been included by %INCLUDE statements.

NOSOURCE2
specifies not to write secondary source statements to the SAS log.

Details

Note: SOURCE2 must be in effect when you execute SAS programs that you want to send to SAS for problem determination and resolution. △

See Also

"The SAS Log" in *SAS Language Reference: Concepts*

SPOOL System Option

Specifies whether SAS statements are written to a utility data set in the WORK data library.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

SPOOL | NOSPOOL

Syntax Description

SPOOL
specifies that SAS write statements to a utility data set in the WORK data library for later use by a %INCLUDE or %LIST statement, or by the RECALL command, within a windowing environment.

NOSPOOL
specifies that SAS does not write statements to a utility data set. Specifying NOSPOOL accelerates execution time, but you cannot use the %INCLUDE and %LIST statements to resubmit SAS statements that were executed earlier in the session.

Examples

Specifying SPOOL is especially helpful in interactive line mode because you can resubmit a line or lines of code by referring to the line numbers. Here is an example of code including line numbers:

```
00001  data test;
00002     input w x y z;
00003     datalines;
00004  411.365 101.945 323.782 512.398
00005  ;
```

If SPOOL is in effect, you can resubmit line number 1 by submitting this statement:

```
%inc 1;
```

You can also resubmit a range of lines by placing a colon (:) or dash (-) between the line numbers. For example, these statements resubmit lines 1 through 3 and 4 through 5 of the above example:

```
%inc 1:3;
%inc 4-5;
```

SQLCONSTDATETIME System Option

Specifies whether the SQL procedure replaces references to the DATE, TIME, DATETIME, and TODAY functions in a query with their equivalent constant values before the query executes.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Files: SAS Files
System administration: SQL
PROC OPTIONS GROUP= SASFILES
SQL

Syntax

SQLCONSTDATETIME | NOSQLCONSTDATETIME

Syntax Description

SQLCONSTDATETIME
specifies that the SQL procedure is to replace references to the DATE, TIME, DATETIME, and TODAY functions with their equivalent numeric constant values.

NOSQLCONSTDATETIME
specifies that the SQL procedure is not to replace references to the DATE, TIME, DATETIME, and TODAY functions with their equivalent numeric constant values.

Details

When the SQLCONSTDATETIME system option is set, the SQL procedure evaluates the DATE, TIME, DATETIME, and TODAY functions in a query once, and uses those values throughout the query. Computing these values once ensures consistency of results when the functions are used multiple times in a query or when the query executes the functions close to a date or time boundary.

When the NOSQLCONSTDATETIME system option is set, the SQL procedure evaluates these functions in a query each time it processes an observation.

If both the SQLREDUCEPUT system option and the SQLCONSTDATETIME system option are specified, the SQL procedure replaces the DATE, TIME, DATETIME, and TODAY functions with their respective values in order to determine the PUT function value before the query executes:

```
select x from &lib..c where (put(bday, date9.) = put(today(), date9.));
```

Note: The value that is specified in the SQLCONSTDATETIME system option is in effect for all SQL procedure statements, unless the CONSTDATETIME option in the PROC SQL statement is set. The value of the CONSTDATETIME option takes precedence over the SQLCONSTDATETIME system option. However, changing the value of the CONSTDATETIME option does not change the value of the SQLCONSTDATETIME system option. △

See Also

System option:
"SQLREDUCEPUT= System Option" on page 2006
PROC SQL statement CONSTDATETIME option in *Base SAS Procedures Guide*
Improving Query Performance in *SAS SQL Procedure User's Guide*

SQLREDUCEPUT= System Option

For the SQL procedure, specifies the engine type that a query uses for which optimization is performed by replacing a PUT function in a query with a logically equivalent expression.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
System administration: SQL
System administration: Performance

PROC OPTIONS GROUP= SASFILES
SQL
PERFORMANCE

Syntax

SQLREDUCEPUT= ALL | NONE | DBMS | BASE

Syntax Description

ALL
specifies that optimization is performed on all PUT functions regardless of any engine that is used by the query to access the data.

NONE
specifies that no optimization is to be performed.

DBMS
specifies that optimization is performed on all PUT functions whose query is performed by a SAS/ACCESS engine. This is the default.

Requirement: The first argument to the PUT function must be a variable obtained by a table that is accessed using a SAS/ACCESS engine.

BASE
specifies that optimization is performed on all PUT functions whose query is performed by a SAS/ACCESS engine or a Base SAS engine.

Details

By using the SQLREDUCEPUT= system option, you can specify that SAS reduces the PUT function as much as possible before the query is processed. If the query also contains a WHERE clause, the evaluation of the WHERE clause is simplified. The following SELECT statements are examples of queries that would be reduced if this option was set to any value other than **none**:

```
select x, y from &lib..b where (PUT(x, abc.) in ('yes', 'no'));
select x from &lib..a where (PUT(x, udfmt.) = trim(left('small')));
```

If both the SQLREDUCEPUT system option and the SQLCONSTDATETIME system option are specified, the SQL procedure replaces the DATE, TIME, DATETIME, and TODAY functions with their respective values to determine the PUT function value before the query executes. The following two SELECT clauses show the original and optimized queries:

```
select x from &lib..c where (put(bday, date9.) = put(today(), date9.));
```

would be reduced to

```
select x from &lib..c where (put(bday, date9.) = "01Jun2008");
```

If a query does not contain the PUT function, optimization is not performed.

Note: The value that is specified in the SQLREDUCEPUT system option is in effect for all SQL procedure statements, unless the REDUCEPUT option in the PROC SQL statement is set. The value of the REDUCEPUT option takes precedence over the SQLREDUCEPUT system option. However, changing the value of the REDUCEPUT option does not change the value of the SQLREDUCEPUT system option. △

See Also

System option:

"SQLCONSTDATETIME System Option" on page 2005
"SQLREDUCEPUTOBS= System Option" on page 2008

"PROC SQL Statement REDUCEPUT option" in the *Base SAS Procedures Guide*

"Improving Query Performance" in the *SAS SQL Procedure User's Guide*

SQLREDUCEPUTOBS= System Option

For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the minimum number of observations that must be in a table in order for PROC SQL to consider optimizing the PUT function in a query.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
System administration: SQL
System administration: Performance

Interaction: If the SQLREDUCEPUT= system option is set to NONE, conditions for both the SQLREDUCEPUTOBS= and SQLREDUCEPUTVALUES= system options must be met in order for the SQL procedure to consider optimizing the PUT function.

PROC OPTIONS GROUP= SASFILES
SQL
PERFORMANCE

Syntax

SQLREDUCEPUTOBS=*n* | *n*K | *n*M | *n*G | *n*T |*hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the number of observations that must be in a table before the SQL procedure considers to optimize the PUT function. *number-of-observations* is an integer that can be allocated in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **8** specifies eight buffers, and a value of **3k** specifies 3,072 buffers.

Default: 0, which indicates that there is no minimum number of observations in a table required for the SQL procedure to optimize the PUT function.

Range: $0 - 2^{63}-1$, or approximately 9.2 quintillion

***hex*X**
specifies the number of observations that must be in a table before the SQL procedure considers to optimize the PUT function as a hexadecimal value. You must

specify the value beginning with a number (0–9), followed by an X. For example, the value **`2dx`** specifies 45 buffers.

MIN
sets the number of observations that must be in a table before the SQL procedure considers to optimize the PUT function to 0. A value of 0 indicates that there is no minimum number of observations required. This is the default.

MAX
sets the maximum number of observations that must be in a table before the SQL procedure considers to optimize the PUT function to 2^{63}–1, or approximately 9.2 quintillion.

Details

For databases that allow implicit pass-through when the row count for a table is not known, the SQL procedure allows the optimization in order for the query to be executed by the database. When the SQLREDUCEPUT= system option is set to NONE, the SQL procedure considers the value of both the SQLREDUCEPUTVALUES= and SQLREDUCEPUTOBS= system options and determines whether to optimize the PUT function.

For databases that do not allow implicit pass-through, the SQL procedure does not perform the optimization, and more of the query is performed by SAS.

See Also

System options:

"SQLREDUCEPUT= System Option" on page 2006

"Improving Query Performance" in the *SAS SQL Procedure User's Guide*

SQLREDUCEPUTVALUES= System Option

For the SQL procedure when the SQLREDUCEPUT= system option is set to NONE, specifies the maximum number of SAS format values that can exist in a PUT function expression in order for PROC SQL to consider optimizing the PUT function in a query.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
System administration: SQL
System administration: Performance

Interaction: If the SQLREDUCEPUT= system option is set to NONE, conditions for both the SQLREDUCEPUTVALUES= and SQLREDUCEPUTOBS= system options must be met in order for the SQL procedure to consider optimizing the PUT function.

PROC OPTIONS GROUP= SASFILES
SQL
PERFORMANCE

Syntax

SQLREDUCEPUTVALUES= *n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the number of SAS format values that can exist in a PUT function expression, where *n* is an integer that can be allocated in multiples of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). For example, a value of **8** specifies eight buffers, and a value of **3k** specifies 3,072 buffers.

Default: 0, which indicates that there is no minimum number of SAS format values that can exist in a PUT function expression.

Range: 0–5,000

Interaction: If the number of format values in a PUT function expression is greater than this value, the SQL procedure does not optimize the PUT function.

***hex*X**
specifies the number of SAS format values that can exist in a PUT function expression. You must specify the value beginning with a number (0–9), followed by an X. For example, the value **2dx** specifies 45 buffers.

MIN
sets the number of SAS format values that can exist in a PUT function expression to 0. A value of 0 indicates that there is no minimum number of SAS format values that are required. This is the default.

MAX
sets the maximum number of SAS format values that can exist in a PUT function expression to 5,000.

Details

Some formats, especially user-defined formats, can contain many format values. Depending on the number of matches for a given PUT function expression, the resulting expression can list many format values. If the number of format values becomes too large, the query performance can degrade. When the SQLREDUCEPUT= system option is set to NONE, the SQL procedure considers the value of both the SQLREDUCEPUTVALUES= and SQLREDUCEPUTOBS= system options and determines whether to optimize the PUT function.

See Also

System options:
"SQLREDUCEPUT= System Option" on page 2006
"SQLREDUCEPUTOBS= System Option" on page 2008
"Improving Query Performance" in the *SAS SQL Procedure User's Guide*

SQLREMERGE System Option

Specifies whether the SQL procedure can process queries that use remerging of data.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
System administration: SQL

PROC OPTIONS GROUP= SASFILES
SQL

Syntax

SQLREMERGE | NOSQLREMERGE

Syntax Description

SQLREMERGE
specifies that the SQL procedure can process queries that use remerging of data.

NOSQLREMERGE
specifies that the SQL procedure cannot process queries that use remerging of data.

Details

The remerge feature of the SQL procedure makes two passes through a table, using data in the second pass that was created in the first pass, in order to complete a query. When the NOSQLREMERGE system option is set, the SQL procedure cannot process remerging of data. If remerging is attempted when the NOSQLREMERGE option is set, an error is written to the SAS log.

See Also

PROC SQL statement REMERGE option in *Base SAS Procedures Guide*

Remerging Data in the summary-function component of the SQL procedure in *Base SAS Procedures Guide*

"Improving Query Performance" in the *SAS SQL Procedure User's Guide*

SQLUNDOPOLICY= System Option

Specifies whether the SQL procedure keeps or discards updated data if errors occur while the data is being updated.

Valid in: configuration file, SAS invocation, Options statement

Category: Files: SAS Files
System administration: SQL

PROC OPTIONS GROUP= SASFILES
SQL

Syntax

SQLUNDOPOLICY=NONE | OPTIONAL | REQUIRED

Syntax Description

NONE
specifies to keep changes that are made by the INSERT and UPDATE statements.

OPTIONAL
specifies to reverse changes that are made by the INSERT and UPDATE statements as long as reversing the changes is reliable.

REQUIRED
specifies to undo all changes that are made by the INSERT and UPDATE statements, up to the point of the error. This is the default.

CAUTION:
Some UNDO operations cannot reliably reverse changes. In some situations, reversing the effects of the INSERT and UPDATE statements cannot be done reliably. When operations cannot be reversed, the SQL procedure issues an error message and does not execute the statement. For example, when a program uses a SAS/ACCESS view, or when a SAS data set is accessed through a SAS/SHARE server and is opened with the data set option CNTLLEV=RECORD, changes cannot be reliably reversed. △

CAUTION:
Some UNDO operations might not reverse changes. In situations where multiple transactions are made to the same record, PROC SQL might not reverse a change; it will issue an error message instead. For example, if an error occurs during an insert, PROC SQL can delete a record that another user updated. In that case, the UNDO statement is not executed, and an error message is issued. △

Details

The value that is specified in the SQLUNDOPOLICY= system option is in effect for all SQL procedure statements, unless the UNDO_POLICY option in the PROC SQL statement is set. The value of the UNDO_POLICY option takes precedence over the SQLUNDOPOLICY= system option. The RESET statement can also be used to set or reset the UNDO_POLICY option. However, changing the value of the UNDO_POLICY option does not change the value of the SQLUNDOPOLICY= system option. Once the procedure completes, the undo policy reverts to the value of the SQLUNDOPOLICY= system option.

If you are updating a data set using the SPD Engine, you can significantly improve processing performance by setting SQLUNDOPOLICY=NONE. However, ensure that NONE is an appropriate setting for your application.

See Also

PROC SQL Statement UNDO_POLICY option in the *Base SAS Procedures Guide*

STARTLIB System Option

Specifies whether SAS assigns user-defined permanent librefs when SAS starts.

Valid in: configuration file, SAS invocation

Category: Files: External files

PROC OPTIONS GROUP= EXTFILES

Syntax

STARTLIB | NOSTARTLIB

Syntax Description

STARTLIB
specifies that when SAS starts, SAS assigns user-defined permanent librefs. STARTLIB is the default for the windowing environment.

NOSTARTLIB
specifies that SAS is not to assign user-defined permanent librefs when SAS starts. NOSTARTLIB is the default for batch mode, interactive line mode, and noninteractive mode.

Details

You assign a permanent libref only in the windowing environment by using the New Library window and by selecting the **`Enable at startup`** check box. SAS stores the permanent libref in the SAS registry. To open the New Library window, right-mouse click **`Libraries`** in the Explorer window and select **`New`**. Alternatively, type DMLIBASSIGN in the command box.

In the windowing environment, SAS automatically assigns permanent librefs when SAS starts because STARTLIB is the default.

In all other execution modes (batch, interactive line, and noninteractive), SAS assigns permanent librefs only when you start SAS with the STARTLIB option specified either on the command line or in the configuration file.

STEPCHKPT System Option

Specifies whether checkpoint-restart data is to be recorded for a batch program.

Valid in: configuration file, SAS invocation

Category: Environment control: Error handling

Requirement: can be used only in batch mode

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

STEPCHKPT | NOSTEPCHKPT

Syntax Description

STEPCHKPT
enables checkpoint mode, which specifies to record checkpoint-restart data.

NOSTEPCHKPT
disables checkpoint mode, which specifies not to record checkpoint-restart data. This is the default.

Details

Using the STEPCHKPT system option puts SAS in checkpoint mode for SAS programs that run in batch. Each time a DATA step or PROC step executes, SAS records data in a checkpoint-restart library. If a program terminates without completing, the program can be resubmitted, beginning with the step that was executing when the program terminated.

To ensure that the checkpoint-restart data is accurate, when you specify the STEPCHKPT option, also specify the ERRORCHECK STRICT option and set the ERRORABEND option so that SAS terminates for most errors.

Checkpoint mode is not valid for batch programs that contain the DM statement, which submits commands to SAS. If checkpoint mode is enabled and SAS encounters a DM statement, checkpoint mode is disabled and the checkpoint catalog entry is deleted.

See Also

System Options:

"STEPCHKPTLIB= System Option" on page 2014
"STEPRESTART System Option" on page 2016
"ERRORABEND System Option" on page 1902
"ERRORCHECK= System Option" on page 1904

Statement:

"CHECKPOINT EXECUTE_ALWAYS Statement" on page 1462

"Restarting Batch Programs" in *SAS Language Reference: Concepts*

STEPCHKPTLIB= System Option

Specifies the libref of the library where checkpoint-restart data is saved.

Valid in: configuration file, SAS invocation
Category: Environment control: Error handling
Requirement: can be used only in batch mode
PROC OPTIONS GROUP= ERRORHANDLING

Syntax

STEPCHKPTLIB=*libref*

Syntax Description

libref
specifies the libref that identifies the library where the checkpoint-restart data is saved.

Default: Work

Requirement: The LIBNAME statement that identifies the checkpoint-restart library must use the BASE engine and be the first statement in the batch program.

Details

When the STEPCHKPT system option is specified, checkpoint-restart data for batch programs is saved in the libref that is specified in the STEPCHKPTLIB= system option. If no libref is specified, SAS uses the Work library to save checkpoint data. The LIBNAME statement that defines the libref must be the first statement in the batch program.

If the Work library is used to save checkpoint data, the NOWORKTERM and NOWORKINIT system options must be specified so that the checkpoint-restart data is available when the batch program is resubmitted. These two options ensure that the Work library is saved when SAS ends and is restored when SAS starts. If the NOWORKTERM option is not specified, the Work library is deleted at the end of the SAS session and the checkpoint-restart data is lost. If the NOWORKINIT option is not specified, a new Work library is created when SAS starts, and again the checkpoint-restart data is lost.

The STEPCHKPTLIB= option must be specified for any SAS session that accesses checkpoint-restart data that is not saved to the Work library.

See Also

System Options:

"STEPCHKPT System Option" on page 2013
"STEPRESTART System Option" on page 2016
"WORKINIT System Option" on page 2056
"WORKTERM System Option" on page 2057

Statement:

"CHECKPOINT EXECUTE_ALWAYS Statement" on page 1462

"Restarting Batch Programs" in *SAS Language Reference: Concepts*

STEPRESTART System Option

Specifies whether to execute a batch program by using checkpoint-restart data.

Valid in: configuration file, SAS invocation
Category: Environment control: Error handling
Requirement: can be used only in batch mode
PROC OPTIONS GROUP= ERRORHANDLING

Syntax

STEPRESTART | NOSTEPRESTART

Syntax Description

STEPRESTART
enables restart mode, which specifies to execute the batch program by using the checkpoint-restart data.

NOSTEPRESTART
disables restart mode, which specifies not to execute the batch program using checkpoint-restart data.

Details

You specify the STEPRESTART option when you want to resubmit a batch program that ran in checkpoint mode and terminated before it completed. When you resubmit the batch program, SAS determines from the checkpoint data which DATA step or PROC step was executing when the program terminated, and resumes executing the batch program by using that DATA or PROC step.

See Also

System Options:
"STEPCHKPT System Option" on page 2013
"STEPCHKPTLIB= System Option" on page 2014
Statement:
"CHECKPOINT EXECUTE_ALWAYS Statement" on page 1462
"Restarting Batch Programs" in *SAS Language Reference: Concepts*

SUMSIZE= System Option

Specifies a limit on the amount of memory that is available for data summarization procedures when class variables are active.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: System administration: Memory

PROC OPTIONS GROUP= MEMORY

Syntax

SUMSIZE=*n* | *n*K | *n*M | *n*G | *n*T | *hex*X | MIN | MAX

Syntax Description

***n* | *n*K | *n*M | *n*G | *n*T**
specifies the amount of memory in terms of 1 (bytes); 1,024 (kilobytes); 1,048,576 (megabytes); 1,073,741,824 (gigabytes); or 1,099,511,627,776 (terabytes). When *n*=0, the default value, the amount of memory is determined by values of the MEMSIZE option and the REALMEMSIZE option. Valid values for SUMSIZE range from 0 to $2^{(n-1)}$ where *n* is the data width in bits (32 or 64) of the operating system.

***hex*X**
specifies the amount of memory as a hexadecimal number. You must specify the value beginning with a number (0–9), followed by an X. For example, a value of **`0fffx`** specifies 4,095 bytes of memory.

MIN
specifies the minimum amount of memory available.

MAX
specifies the maximum amount of memory available.

Details

The SUMSIZE= system option affects the MEANS, OLAP, REPORT, SUMMARY, SURVEYFREQ, SURVEYLOGISTIC, SURVEYMEANS, and TABULATE procedures.

Proper specification of SUMSIZE= can improve procedure performance by restricting the swapping of memory that is controlled by the operating environment.

Generally, the value of the SUMSIZE= system option should be less than the physical memory available to your process. If the procedure you are using needs more memory than you specify, the system creates a temporary utility file.

If the value of SUMSIZE is greater than the values of the MEMSIZE option and the REALMEMSIZE option, SAS uses the values of the MEMSIZE option and REALMEMSIZE option.

See Also

System Options:

"SORTSIZE= System Option" on page 2000

"MEMSIZE System Option" in the documentation for your operating environment.

"REALMEMSIZE System Option" in the documentation for your operating environment.

SVGCONTROLBUTTONS

Specifies whether to display the paging control buttons and an index in a multipage SVG document.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Syntax

SVGCONTROLBUTTONS | NOSVGCONTROLBUTTONS

Syntax Description

SVGCONTROLBUTTONS
specifies to display the paging control buttons in the SVG document.

NOSVGCONTROLBUTTONS
specifies not to display the paging control buttons in the SVG document. This is the default.

Details

When SVGCONTROLBUTTONS is specified, the size of the SVG is increased to accommodate the script that controls paging in the SVG document.

The SVGView printer sets the option to SVGCONTROLBUTTONS.

SVGHEIGHT= System Option

Specifies the height of the viewport unless the SVG output is embedded in another SVG output; specifies the value of the height attribute of the outermost <svg> element in the SVG file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Restriction: The SVGHEIGHT= option sets the **height** attribute only on the outermost **<svg>** element.

Syntax

SVGHEIGHT= *number-of-units*<*unit-of-measure*> | "" | ''

Syntax Description

number-of- units
specifies the height as a number of *unit-of-measure*.

Requirement: *number-of- units* must be a positive integer value.

Interaction: If *number-of- units* is a negative number, the SVG document is not rendered by the browser.

unit-of-measure
specifies the unit of measurement, which can be one of the following:

%	percentage
cm	centimeters
em	the height of the element's font
ex	the height of the letter x
in	inches
mm	millimeters
pc	picas
pt	points
px	pixels

Default: px

"" | ''
specifies to reset the height to the default value of 600 pixels.

Requirement: Use two double quotation marks or two single quotation marks with no space between them.

Details

For embedded **<svg>** elements, the SVGHEIGHT= option specifies the height of the rectangular region into which the **<svg>** element is placed. The SVG output is scaled to fit the viewBox if SVGHEIGHT="100%".

If the SVGHEIGHT= option is not specified, the height attribute on the **<svg>** element is not set, which effectively provides full scalability by using a height of 100%.

The value for the SVGHEIGHT= option can be specified using no delimiters, enclosed in single or double quotation marks, or enclosed in parentheses.

Examples

The following OPTIONS statement specifies to size the SVG output to portrait letter-sized and to scale the output to 100% of the viewport:

```
options printerpath=svg orientation=portrait svgheight="100%" svgwidth="100%"
        papersize=letter;
```

By using these option values, SAS creates the following **<svg>** element:

```
<svg> xmlns="http://www.w3.org/2000/svg"
      xmlns:xlink="http://www.w3.org/1999/xlink"
      xml:space="preserve"
      onload='Init(evt)' version="1.1"
      width="100%" height="100%"
      viewBox="0 0 850 1100"
</svg>
```

The value of "100%" in the SVGHEIGHT= option specifies to scale the SVG output height to 100% of the viewport, which is based on the value of the PAPERSIZE= option. The paper size is letter in the portrait orientation, which has a height of 11" at 100 dpi.

See Also

System options:

- "SVGCONTROLBUTTONS" on page 2018
- "SVGPRESERVEASPECTRATIO= System Option" on page 2021
- "SVGTITLE= System Option" on page 2023
- "SVGVIEWBOX= System Option" on page 2024
- "SVGWIDTH= System Option" on page 2026
- "SVGX= System Option" on page 2028
- "SVGY= System Option" on page 2029

"Using SAS System Options" on page 1823

The SAS Registry in *SAS Language Reference: Concepts*

Creating Scalable Vector Graphics Using Universal Printing in *SAS Language Reference: Concepts*

SVGPRESERVEASPECTRATIO= System Option

Specifies whether to force uniform scaling of SVG output; specifies the preserveAspectRatio attribute on the outermost <svg> element.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Restriction: The SVGPRESERVEASPECTRATIO= option sets the **`preserveAspectRatio`** attribute only on the outermost **`<svg>`** element.

Syntax

SVGPRESERVEASPECTRATIO=*align* | *meetOrSlice* | **NONE** | ""

SVGPRESERVEASPECTRATIO="*align meetOrSlice*"

Syntax Description

align
specifies to force uniform scaling by specifying the alignment method to use. The value for *align* can be one of the following:

`xMinYMin` specifies to force uniform scaling by using the following alignment:
- Align the <min–x> of the element's viewBox with the smallest X value of the viewport.
- Align the <min–y> of the element's viewBox with the smallest Y value of the viewport.

`xMidYMin` specifies to force uniform scaling by using the following alignment:
- Align the midpoint X value of the element's viewBox with the midpoint X value of the viewport.
- Align the <min–y> of the element's viewBox with the smallest Y value of the viewport.

`xMaxYMin` specifies to force uniform scaling by using the following alignment:
- Align the <min–x>+<width> of the element's viewBox with the maximum X value of the viewport.
- Align the <min–y> of the element's viewBox with the smallest Y value of the viewport.

`xMinYMid` specifies to force uniform scaling by using the following alignment:
- Align the <min–x> of the element's viewBox with the smallest X value of the viewport.
- Align the midpoint Y value of the element's viewBox with the midpoint Y value of the viewport.

`xMidYMid` specifies to force uniform scaling by using the following alignment:

Align the midpoint X value of the element's viewBox with the midpoint X value of the viewport.

Align the midpoint Y value of the element's viewBox with the midpoint Y value of the viewport. This is the default.

xMaxYMid specifies to force uniform scaling by using the following alignment:

Align the <min–x>+<width> of the element's viewBox with the maximum X value of the viewport.

Align the midpoint Y value of the element's viewBox with the midpoint Y value of the viewport.

xMinYMax specifies to force uniform scaling by using the following alignment:

Align the <min–x> of the element's viewBox with the smallest X value of the viewport.

Align the <min–y>+<height> of the element's viewBox with the maximum Y value of the viewport.

xMidYMax specifies to force uniform scaling by using the following alignment:

Align the midpoint X value of the element's viewBox with the midpoint X value of the viewport.

Align the <min–y>+<height> of the element's viewBox with the maximum Y value of the viewport.

xMaxYMax specifies to force uniform scaling by using the following alignment:

Align the <min–x>+<width> of the element's viewBox with the maximum X value of the viewport.

Align the <min–y>+<height> of the element's viewBox with the maximum Y value of the viewport.

meetOrSlice
specifies to preserve the aspect ratio and how the viewBox displays. The following values are valid for *meetOrSlice*:

meet specifies to scale the SVG graphic as follows:

- preserve the aspect ratio
- make the entire viewBox visible within the viewport
- scale up the viewBox as much as possible while meeting other criteria

If the aspect ratio of the graphic does not match the viewport, some of the viewport will extend beyond the bounds of the viewBox.

slice specifies to scale the SVG graphic as follows:

- preserve the aspect ratio
- cover the entire viewBox with the viewport
- scale down the viewBox as much as possible while meeting other criteria

If the aspect ratio of the viewBox does not match the viewport, some of the viewBox will extend the bounds of the viewport.

NONE
specifies not to force uniform scaling and to scale the SVG output nonuniformly so that the element's bounding box exactly matches the viewport rectangle.

""

specifies to reset the **preserveAspectRatio** attribute of the **<svg>** element to the default value of **xMidYMid meet**.

Requirement: Use two double quotation marks with no space between them.

Details

When the value of the SVGPRESERVEASPECTRATIO= option includes both *align* and *meetOrSlice*, you can delimit the value by using single or double quotation marks or parentheses.

The **preserveAspectRatio** attribute applies only when a value is provided for the viewBox on the same **<svg>** element. If the **viewBox** attribute is not provided, the **preserveAspectRatio** attribute is ignored.

Examples

The following OPTIONS statements are examples of using the SVGPRESERVEASPECTRATIO= system option:

```
options svgpreserveaspectratio=xMinYMax;
options svgpreserveaspectratio="xMinYMin meet";
options svgpreserveaspectratio=(xMinYMin meet);
options svgpreserveaspectratio="";
```

See Also

System options:

"SVGCONTROLBUTTONS" on page 2018
"SVGHEIGHT= System Option" on page 2019
"SVGTITLE= System Option" on page 2023
"SVGWIDTH= System Option" on page 2026
"SVGVIEWBOX= System Option" on page 2024
"SVGX= System Option" on page 2028
"SVGY= System Option" on page 2029

Creating Scalable Vector Graphics Using Universal Printing in *SAS Language Reference: Concepts*

SVGTITLE= System Option

Specifies the title in the title bar of the SVG output; specifies the value of the <title> element in the SVG file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Syntax

SVGTITLE="*title*" | "" | ''

Syntax Description

"*title*"
specifies the title of the SVG.

"" | ''
specifies to reset the title to empty.

Requirement: Use two double quotation marks or two single quotation marks with no space between them.

Details

If the SVGTITLE option is not specified, the title bar of the SVG output displays the filename of the SVG output.

The value for the SVGTITLE= option must be enclosed in single or double quotation marks, or enclosed in parentheses.

See Also

System options:

"SVGCONTROLBUTTONS" on page 2018
"SVGHEIGHT= System Option" on page 2019
"SVGPRESERVEASPECTRATIO= System Option" on page 2021
"SVGWIDTH= System Option" on page 2026
"SVGVIEWBOX= System Option" on page 2024
"SVGX= System Option" on page 2028
"SVGY= System Option" on page 2029

Creating Scalable Vector Graphics Using Universal Printing in *SAS Language Reference: Concepts*

SVGVIEWBOX= System Option

Specifies the coordinates, width, and height that are used to set the viewBox attribute on the outermost <svg> element, which enables SVG output to scale to the viewport.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Restriction: The SVGVIEWBOX= option sets the **`viewBox`** attribute only on the outermost **`<svg>`** element.

Syntax

SVGVIEWBOX="*min-x min-y width height*" | none | "" | ”

Syntax Description

min–x
specifies the beginning x coordinate of the viewBox, in user units.

Requirement: *min-x* can be 0, or a positive or a negative integer value.

min–y
specifies the beginning y coordinate of the viewBox, in user units.

Requirement: *min–y* can be 0, or a positive or negative integer value.

width
specifies the width of the viewBox, in user units.

Requirement: *width* must be a positive integer value.

height
specifies the height of the viewBox, in user units.

Requirement: *height* must be a positive integer value.

none
specifies that no **viewBox** attribute is to be specified on the outermost **<svg>** element, which will effectively create a static SVG document.

"" | ”
specifies to reset the width and height of the viewBox to the width and height of the paper size for the SVG printer.

Requirement: Use two double quotation marks or two single quotation marks with no space between them.

Details

When the **viewBox** attribute is specified, the SVG output is scaled to be rendered in the viewport and the current coordinate system is updated to be the dimensions that are specified by the **viewBox** attribute. If it is not specified, the **viewBox** attribute on the outermost **<svg>** element sets the height and width arguments of the viewBox attribute to the paper height and paper width as defined by the PAPERSIZE= system option.

The coordinates, width, and height of the **viewBox** attribute should be mapped to the coordinates, width, and height of the viewport, taking into account the values of the **preserveAspectRatio** attribute.

The value for the SVGVIEWBOX= option must be enclosed in single or double quotation marks, or enclosed in parentheses.

You can use a negative value for *min-x* and *min-y* to place the SVG document in the output. A negative value of *min-x* shifts the output to the right. A negative value of *min-y* shifts the placement of the output downward.

Examples

The following OPTIONS statement specifies to scale the output to a width of 100 user units and a height of 200 user units:

```
options printerpath=svg svgviewbox="0 0 100 200" dev=sasprtc;
```

By using these option values, SAS creates the following **<svg>** element:

```
<svg> xmlns="http://www.w3.org/2000/svg"
      xmlns:xlink="http://www.w3.org/1999/xlink"
      xml:space="preserve"
      onload='Init(evt)' version="1.1"
      viewBox="0 0 100 200"
</svg>
```

See Also

System options:

"SVGCONTROLBUTTONS" on page 2018
"SVGHEIGHT= System Option" on page 2019
"SVGPRESERVEASPECTRATIO= System Option" on page 2021
"SVGTITLE= System Option" on page 2023
"SVGWIDTH= System Option" on page 2026
"SVGX= System Option" on page 2028
"SVGY= System Option" on page 2029

Creating Scalable Vector Graphics Using Universal Printing in *SAS Language Reference: Concepts*

SVGWIDTH= System Option

Specifies the width of the viewport unless the SVG output is embedded in another SVG output; specifies the value of the width attribute in the outermost <svg> element in the SVG file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Restriction: The SVGWIDTH= option sets the **width** attribute only on the outermost **<svg>** element.

Syntax

SVGWIDTH= *number-of-units*<*unit-of-measure*> | "" | ''

Syntax Description

number-of-units
specifies the width as a number of *unit-of-measure*.

Requirement: *number-of- units* must be a positive integer value.

Interaction: If *number-of- units* is a negative number, the SVG document is not rendered by the browser.

unit-of-measure
specifies the unit of measurement, which can be one of the following:

%	percentage
cm	centimeters
em	the height of the element's font
ex	the height of the letter x
in	inches
mm	millimeters
pc	picas
pt	points
px	pixels

Default: px

"" | ''
specifies to reset the width to the default value of 800 pixels.

Requirement: Use two double quotation marks or two single quotation marks with no space between them.

Details

For embedded **<svg>** elements, the SVGWIDTH= option specifies the width of the rectangular region into which the **<svg>** element is placed. The SVG output is scaled to fit the viewBox if SVGWIDTH="100%".

If the SVGWIDTH= option is not specified, the width attribute on the **<svg>** element is not set, which effectively provides full scalability by using a width of 100%.

The value for the SVGWIDTH= option can be specified without delimiters, enclosed in single or double quotation marks, or enclosed in parentheses.

Examples

The following OPTIONS statement specifies to size the SVG output to portrait letter-sized and to scale the output to 100% of the viewport:

```
options printerpath=svg orientation=portrait svgheight="100%" svgwidth="100%"
        papersize=letter;
```

By using these option values, SAS creates the following **<svg>** element:

```
<svg> xmlns="http://www.w3.org/2000/svg"
      xmlns:xlink="http://www.w3.org/1999/xlink"
      xml:space="preserve"
      onload='Init(evt)' version="1.1"
      width="100%" height="100%"
      viewBox="0 0 850 1100"
</svg>
```

The value of "100%" in the SVGWIDTH= option specifies to scale the SVG output width to 100% of the viewport, which is based on the value of the PAPERSIZE= option. The paper size is letter in the portrait orientation, which has a width of 8.5" at 96 dpi.

See Also

System options:

"SVGCONTROLBUTTONS" on page 2018
"SVGHEIGHT= System Option" on page 2019
"SVGPRESERVEASPECTRATIO= System Option" on page 2021
"SVGTITLE= System Option" on page 2023
"SVGVIEWBOX= System Option" on page 2024
"SVGX= System Option" on page 2028
"SVGY= System Option" on page 2029
"Using SAS System Options" on page 1823

The SAS Registry in *SAS Language Reference: Concepts*

Creating Scalable Vector Graphics Using Universal Printing in *SAS Language Reference: Concepts*

SVGX= System Option

Specifies the x-axis coordinate of one corner of the rectangular region into which an embedded <svg> element is placed; specifies the x attribute in the outermost <svg> element in an SVG file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Restriction: The SVGX= option sets the **x** attribute only on the outermost **<svg>** element.

Syntax

SVGX= *number-of-units*<*unit-of-measure*> | "" | ''

Syntax Description

number-of-units
specifies the x-axis coordinate as a number of *unit-of-measure*.

unit-of-measure
specifies the unit of measurement, which can be one of the following:

%	percentage
cm	centimeters
em	the height of the element's font
ex	the height of the letter x
in	inches
mm	millimeters

pc	picas
pt	points
px	pixels

Default: px

"" | ''
specifies to reset the **x** attribute to 0 on the **<svg>** element and the x-axis coordinate for embedded SVG to 0.

Requirement: Use two double quotation marks or two single quotation marks with no space between them.

Details

If the SVGX= option is not set, the **x** attribute on the **<svg>** element effectively has a value of 0 and no x-axis coordinate is set for embedded SVG output.

The value for the SVGX= option can be specified without delimiters, enclosed in single or double quotation marks, or enclosed in parentheses.

The **x** attribute on the outermost **<svg>** element has no effect on SVG documents that are produced by SAS. You can use the SVGX= system option to specify the x-axis coordinate if the SVG document is processed outside of SAS.

See Also

System options:

"SVGCONTROLBUTTONS" on page 2018
"SVGHEIGHT= System Option" on page 2019
"SVGPRESERVEASPECTRATIO= System Option" on page 2021
"SVGTITLE= System Option" on page 2023
"SVGWIDTH= System Option" on page 2026
"SVGVIEWBOX= System Option" on page 2024
"SVGY= System Option" on page 2029

Creating Scalable Vector Graphics Using Universal Printing in *SAS Language Reference: Concepts*

SVGY= System Option

Specifies the y-axis coordinate of one corner of the rectangular region into which an embedded <svg> element is placed; specifies the y attribute in the outermost <svg> element in an SVG file.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: SVG

PROC OPTIONS GROUP= SVG

Restriction: The SVGY= option sets the **y** attribute only on the outermost **<svg>** element.

Syntax

SVGY= *number-of-units<unit-of-measure>* | "" | ''

Syntax Description

number-of-units
specifies the y-axis coordinate as a number of *unit-of-measure*.

unit-of-measure
specifies the unit of measurement, which can be one of the following:

%	percentage
cm	centimeters
em	the height of the element's font
ex	the height of the letter x
in	inches
mm	millimeters
pc	picas
pt	points
px	pixels

Default: px

"" | ''
specifies to reset the **y** attribute on the **<svg>** element and the y-axis coordinate for embedded SVG output to 0.

Requirement: Use two double quotation marks or two single quotation marks with no space between them.

Details

If the SVGY= option is not set, the **y** attribute on the **<svg>** element effectively has a value of 0 and no y-axis coordinate is set for embedded SVG output.

The value for the SVGY= option can be specified without delimiters, enclosed in single or double quotation marks, or enclosed in parentheses.

The **y** attribute on the outermost **<svg>** element has no effect on SVG documents that are produced by SAS. You can use the SVGY= system option to specify the y-axis coordinate if the SVG document is processed outside of SAS.

See Also

System options:

"SVGCONTROLBUTTONS" on page 2018
"SVGHEIGHT= System Option" on page 2019
"SVGPRESERVEASPECTRATIO= System Option" on page 2021
"SVGTITLE= System Option" on page 2023
"SVGWIDTH= System Option" on page 2026
"SVGVIEWBOX= System Option" on page 2024
"SVGX= System Option" on page 2028

Creating Scalable Vector Graphics Using Universal Printing in *SAS Language Reference: Concepts*

SYNTAXCHECK System Option

In non-interactive or batch SAS sessions, specifies whether to enable syntax check mode for multiple steps.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Environment control: Error handling
PROC OPTIONS GROUP= ERRORHANDLING

Syntax

SYNTAXCHECK | NOSYNTAXCHECK

Syntax Description

SYNTAXCHECK
enables syntax check mode for statements that are submitted within a non-interactive or batch SAS session.

NOSYNTAXCHECK
does not enable syntax check mode for statements that are submitted within a non-interactive or batch SAS session.

CAUTION:
Setting NOSYNTAXCHECK might cause a loss of data. Manipulating and deleting data by using untested code might result in a loss of data if your code contains invalid syntax. Be sure to test code completely before placing it in a production environment. △

Details

If a syntax or semantic error occurs in a DATA step after the SYNTAXCHECK option is set, then SAS enters syntax check mode, which remains in effect from the point where SAS encountered the error to the end of the code that was submitted. After SAS enters syntax mode, all subsequent DATA step statements and PROC step statements are validated.

While in syntax check mode, only limited processing is performed. For a detailed explanation of syntax check mode, see "Syntax Check Mode" in the section "Error Processing in SAS" in *SAS Language Reference: Concepts*.

Place the OPTIONS statement that enables SYNTAXCHECK before the step for which you want it to take effect. If you place the OPTIONS statement inside a step, then SYNTAXCHECK will not take effect until the beginning of the next step.

NOSYNTAXCHECK enables continuous processing of statements regardless of syntax error conditions.

SYNTAXCHECK is ignored in the SAS windowing environment and in SAS line-mode sessions.

Comparisons

You use the SYNTAXCHECK system option to validate syntax in a non-interactive or a batch SAS session. You use the DMSSYNCHK system option to validate syntax in an interactive session by using the SAS windowing environment.

The ERRORCHECK= option can be set to enable or disable syntax check mode for the LIBNAME statement, the FILENAME statement, the %INCLUDE statement, and the LOCK statement in SAS/SHARE. If you specify the NOSYNTAXCHECK option and the ERRORCHECK=STRICT option, then SAS does not enter syntax check mode when an error occurs.

See Also

System Options:

"DMSSYNCHK System Option" on page 1890

"ERRORCHECK= System Option" on page 1904

"Error Processing in SAS" in the section "Error Processing and Debugging" in *SAS Language Reference: Concepts*

SYSPRINTFONT= System Option

Specifies the default font to use for printing, which can be overridden by explicitly specifying a font and an ODS style.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: Procedure output

PROC OPTIONS GROUP= LISTCONTROL

See: SYSPRINTFONT= System Option in the documentation for your operating environment.

Syntax

SYSPRINTFONT=("*face-name*" <*weight*> <*style*> <*character-set*> <*point-size*> <NAMED "*printer-name* " | UPRINT="*printer-name* " | DEFAULT | ALL>)

Syntax Description

"face-name"

specifies the name of the font face to use for printing.

Requirement: If *face-name* consists of more than one word, you must be enclose the value in single or double quotation marks. The quotation marks are stored with the face-name.

Requirement: When you use the SYSPRINTFONT= option with multiple arguments, you must enclose the arguments in parentheses.

Interaction: When you specify UPRINT=*printer-name*, *face-name* must be a valid font for *printer-name*.

weight

specifies the weight of the font, such as BOLD. A list of valid values for your specified printer appears in the SAS: Printer Properties window.

Default: NORMAL

style
specifies the style of the font, such as Italic. A list of valid values for your specified printer appears in the SAS: Printer Properties window.

Default: REGULAR

character-set
specifies the character set to use for printing.

Default: If the font does not support the specified character set, the default character set is used. If the default character set is not supported by the font, the font's default character set is used.

Range: Valid values are listed in the SAS: Printer Properties window, under the **`Font`** tab.

point-size
specifies the point size to use for printing. If you omit this argument, SAS uses the default.

Requirement: *Point-size* must be an integer. It must also be placed after the *face-name*, *weight*, *style*, and *character-set* arguments.

NAMED "*printer-name*"
specifies a printer in the Windows operating environment to which these settings apply.

Restriction: This argument is valid only for printers in the Windows operating environment. To specify a Universal Printer, use the UPRINT=argument.

Requirement: The *printer-name* must exactly match the name shown in the Print Setup dialog box (except that the printer name is not case sensitive).

Requirement: If the printer is more than one word, the *printer-name* must be enclosed in double quotation marks. The quotation marks are stored with the printer-name.

UPRINT="*printer-name*"
specifies a Universal Printer to which these settings apply.

Restriction: This argument is valid only for printers that are listed in the SAS Registry.

Requirement: The *printer-name* must match exactly the name shown in the Print Setup dialog box (except that the printer name is not case sensitive).

Requirement: If the *printer-name* is more than one word, it must be enclosed in single or double quotation marks. The quotation marks are stored with the printer-name.

DEFAULT | ALL
specifies whether the font settings apply to the default printer or to all printers:

DEFAULT
specifies that the font settings apply to the current default printer that is specified by the SYSPRINT= system option.

ALL
specifies that the font settings apply to all installed printers.

Details

The SYSPRINTFONT= system option sets the font to use when printing to the current default printer, to a specified printer or to all printers.

In some cases, you might need to specify the font from a SAS program. In this case, you might want to view the SAS: Printer Properties window for allowable names, styles weights, and sizes for your fonts. For examples of how to apply the SYSPRINTFONT= option in a SAS program, see "Examples" on page 2034.

If you specified SYSPRINTFONT= with DEFAULT or without a keyword and later use the Print Setup dialog box to change the current default printer, then the font used with the current default printer will be the font that was specified with SYSPRINTFONT, if the specified font exists on the printer. If the current printer does not support the specified font, the printer's default font is used.

The following fonts are widely supported:

- Helvetica
- Times
- Courier
- Symbol

By specifying one of these fonts in a SAS program, you can usually avoid returning an error. If that particular font is not supported, a similar-looking font prints in its place.

All Universal printers and many SAS/GRAPH devices use the FreeType engine to render TrueType fonts. For more information, see Using TrueType Fonts with Universal Printing and SAS/GRAPH Devices in *SAS Language Reference: Concepts*.

Note: As an alternative to using the SYSPRINTFONT= system option, you can set fonts with the SAS: Printer Properties window, under the **Font** tab. From the drop-down menu select **File ▶ Print Setup ▶ Properties ▶ Font**. Using a dialog box is fast and easy because you choose your font, style, weight, size, and character set from a list of options that your selected printer supports. △

Examples

Specifying a Font to the Default Printer

This example specifies the 12–point Times font on the default printer:

```
options sysprintfont=("times" 12);
```

Specifying a Font to a Named Windows Printer

This example specifies to use Courier on the printer named HP LaserJet IIIsi Postscript. Specify the printer name in the same way that it is specified in the SAS Print Setup dialog box:

```
options sysprintfont= ("courier" named "hp laserjet 111s, postscript");
```

Specifying a Font to a Universal Printer, on the SAS command line

This example specifies the Albany AMT font for the PDF Universal Printer::

```
sysprintfont=('courier' 11 uprint='PDF')
```

TERMINAL System Option

Specifies whether to associate a terminal with a SAS session.

Valid in: configuration file, SAS invocation

Category: Environment control: Initialization and operation

PROC OPTIONS GROUP= EXECMODES

Syntax

TERMINAL | NOTERMINAL

Syntax Description

TERMINAL
specifies that SAS evaluate the execution environment and if a physical display is not available for an interactive environment, sets the option to NOTERMINAL. Specify TERMINAL when you use the SAS windowing environment.

NOTERMINAL
specifies that SAS not evaluate the execution environment.

Details

SAS defaults to the appropriate setting for the TERMINAL system option based on whether the session is invoked in the foreground or the background. If NOTERMINAL is specified, dialog boxes are not displayed.

The TERMINAL option is normally used with the following execution modes:

- SAS windowing environment mode
- interactive line mode
- noninteractive mode.

TERMSTMT= System Option

Specifies the SAS statements to execute when SAS terminates.

Valid in: configuration file, SAS invocation

Category: Environment control: Initialization and operation

PROC OPTIONS GROUP= EXECMODES

Syntax

TERMSTMT=*'statement(s)'*

Syntax Description

'statement(s)'
is one or more SAS statements.

Maximum length: 2,048 characters

Operating Environment Information: In some operating system environments there is a limit to the size of the value for TERMSTMT=. To circumvent this limitation, you can use the %INCLUDE statement. △

Details

TERMSTMT= is fully supported in batch mode. In interactive modes, TERMSTMT= is executed only when you submit the ENDSAS statement from an editor window to terminate the SAS session. Terminating SAS by any other means in interactive mode results in TERMSTMT= not being executed.

An alternate method for specifying TERMSTMT= is to put a %INCLUDE statement at the end of a batch file or to submit a %INCLUDE statement before terminating the SAS session in interactive mode.

Comparisons

TERMSTMT= specifies the SAS statements to be executed at SAS termination, and INITSTMT= specifies the SAS statements to be executed at SAS initialization.

See Also

System Option:

"INITSTMT= System Option" on page 1928

Statement:

"%INCLUDE Statement" on page 1584

TEXTURELOC= System Option

Specifies the location of textures and images that are used by ODS styles.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

TEXTURELOC=*location*

Syntax Description

location
specifies the location of textures and images used by ODS styles. *Location* can refer either to the physical name of the directory or to a URL reference to the directory.

Requirement: If *location* is not a fileref, then you must enclose the value in quotation marks.

Restriction: Only one location is allowed per statement.

Requirement: The files in the directory must be in the form of gif, jpeg, or bitmap.

Requirement: *Location* must refer to a directory.

See Also

"Dictionary of ODS Language Statements" in *SAS Output Delivery System: User's Guide*.

THREADS System Option

Specifies that SAS use threaded processing if it is available.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: System administration: Performance

PROC OPTIONS GROUP= PERFORMANCE

Syntax

THREADS | NOTHREADS

Syntax Description

THREADS
specifies to use threaded processing for SAS applications that support it.

Interaction If THREADS is specified either as a SAS system option or in PROC SORT and another program has the input SAS data set open for reading, writing, or updating using the SPD engine, then the procedure might fail and write a subsequent message to the SAS log.

NOTHREADS
specifies not to use threaded processing for running SAS applications that support it.

Interaction: When you specify NOTHREADS, CPUCOUNT= is ignored unless you specify a procedure option that overrides the NOTHREADS system option.

Details

The THREADS system option enables some legacy SAS processes that are thread-enabled to take advantage of multiple CPUs by threading the processing and I/O

operations. Threading the processing and I/O operations achieves a degree of parallelism that generally reduces the real time to completion for a given operation at the possible cost of additional CPU resources. In SAS 9 and SAS 9.1, the thread-enabled processes include

- Base SAS engine indexing
- Base SAS procedures: SORT, SUMMARY, MEANS, REPORT, TABULATE, and SQL
- SAS/STAT procedures: GLM, LOESS, REG, ROBUSTREG.

For example, in some cases, processing small data sets, SAS might determine to use a single-threaded operation.

Set this option to NOTHREADS to achieve SAS behavior most compatible with releases before to SAS 9, if you find that threading does not improve performance or if threading might be related to an unexplainable problem. See the specific documentation for each product to determine whether it has functionality that is enabled by the THREADS option.

Comparisons

The system option THREADS determines when threaded processing is in effect. The SAS system option CPUCOUNT= suggests how many system CPUs are available for use by thread-enabled SAS procedures.

See Also

System Option:

"CPUCOUNT= System Option" on page 1875

"UTILLOC= System Option" on page 2043

"Support for Parallel Processing" in *SAS Language Reference: Concepts*.

TOOLSMENU System Option

Specifies whether the Tools menu is included in SAS windows.

Default: TOOLSMENU

Valid in: configuration file, SAS invocation

Category: Environment control: Display

PROC OPTIONS GROUP= ENVDISPLAY

Syntax

TOOLSMENU | NOTOOLSMENU

Syntax Description

TOOLSMENU
specifies that the Tools menu is included in SAS windows.

NOTOOLSMENU
specifies that the Tools menu is not included in SAS windows.

TOPMARGIN= System Option

Specifies the print margin at the top of the page for output directed to an ODS printer destination.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Log and procedure output control: ODS Printing

PROC OPTIONS GROUP= ODSPRINT

Syntax

TOPMARGIN= *margin-size<margin-unit>*

Syntax Description

margin-size
specifies the size of the margin.

Restriction: The bottom margin should be small enough so that the top margin plus the bottom margin is less than the height of the paper.

Interactions: Changing the value of this option might result in changes to the value of the PAGESIZE= system option.

<margin-unit>
specifies the units for margin-size. The margin-unit can be *in* for inches or *cm* for centimeters. <*margin-unit*> is saved as part of the value of the TOPMARGIN system option.

Default: inches

Details

All margins have a minimum that is dependent on the printer and the paper size. The default value of the TOPMARGIN system option is **`0.00 in`**.

Operating Environment Information: Most SAS system options are initialized with default settings when SAS is invoked. However, the default settings and option values for some SAS system options might vary both by operating environment and by site. For details, see the SAS documentation for your operating environment. △

For additional information about declaring an ODS printer destination, see the ODS statements in the *SAS Output Delivery System: User's Guide*.

See Also

System Options:

"BOTTOMMARGIN= System Option" on page 1850
"LEFTMARGIN= System Option" on page 1934
"RIGHTMARGIN= System Option" on page 1984

TRAINLOC= System Option

Specifies the URL for SAS online training courses.

Valid in: configuration file, SAS invocation
Category: Environment control: Files
PROC OPTIONS GROUP= ENVFILES

Syntax

TRAINLOC=*"base-URL"*

Syntax Description

base-URL
specifies the address where the SAS online training courses are located.

Details

The TRAINLOC= system option specifies the base location (typically a URL) of SAS online training courses. These online training courses are typically accessed from an intranet server or a local CD-ROM.

Examples

Some examples of the *base-URL* are:

- "*file://e:\onlintut*"
- "*http://server.abc.com/SAS/sastrain*"

UNIVERSALPRINT System Option

Specifies whether to enable Universal Printing services.

Valid in: configuration file, SAS invocation

Category: Log and procedure output control: ODS Printing
PROC OPTIONS GROUP= ODSPRINT

Syntax

UNIVERSALPRINT | NOUNIVERSALPRINT

Syntax Description

UNIVERSALPRINT
routes all printing through the Universal Print services.
Alias: UPRINT

NOUNIVERSALPRINT
disables printing through the Universal Print services.
Alias: NOUPRINT

Details

Universal Printing services provides interactive and batch printing capabilities to SAS applications and procedures. The ODS PRINTER destination uses Universal Print services whenever the UNIVERSALPRINT option is enabled.

See Also

"Printing with SAS" in *SAS Language Reference: Concepts*

UPRINTCOMPRESSION System Option

Specifies whether to enable compression of file created by some Universal Printer and SAS/GRAPH devices.

Alias: UPC | NOUPC
Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window
Category: Log and procedure output control: ODS Printing
PROC OPTIONS GROUP= ODSPRINT

Syntax

UPRINTCOMPRESSION | NOUPRINTCOMPRESSION

Syntax Description

UPRINTCOMPRESSION
specifies to enable compression of files created by some Universal Printers and some SAS/GRAPH devices. This is the default.

NOUPRINTCOMPRESSION
specifies to disable compression of files created by some Universal Printers and some SAS/GRAPH devices.

Details

The following table lists the Universal Printers and the SAS/GRAPH devices that are affected by the UPRINTCOMPRESSION system option:

Universal Printers	SAS/GRAPH Device Drivers
PCL5, PCL5C, PCL5E	PCL5, PCL5C, PCL5E
PDF	PDF, PDFA, PDFC
SVGZ	SVGZ
PS	SASPRTC, SASPRTG, SASPRTM
	UEPS, UPSC, UPCL5, UPCL5C, UPCL5E, UPDF, UPSL, UPSLC

When NOUPRINTCOMPRESSION is set, the DEFLATION= option is ignored.

The ODS PRINTER statement option, COMPRESS=, takes precedence over the UPRINTCOMPRESSION system option.

See Also

System options:

"DEFLATION= System Option" on page 1880

Statements:

"ODS PRINTER Statement" in the *SAS Output Delivery System: User's Guide*

USER= System Option

Specifies the default permanent SAS library.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: USER= System Option in the documentation for your operating environment.

Syntax

USER= *library-specification*

Syntax Description

library-specification
specifies the libref or physical name of a SAS library.

Details

If this option is specified, you can use one-level names to reference permanent SAS files in SAS statements. However, if USER=WORK is specified, SAS assumes that files referenced with one-level names refer to temporary work files.

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

UTILLOC= System Option

Specifies one or more file system locations in which applications can store utility files.

Valid in: configuration file and SAS invocation

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

See: UTILLOC= System Option in the documentation for your operating environment.

Syntax

UTILLOC= WORK | *location* | (*location-1*... *location-n*)

Syntax Description

WORK
specifies that SAS creates utility files in the same directory as the Work library. This is the default.

location
specifies the location of an existing directory for utility files that are created by applications. Enclose *location* in single or double quotation marks when the location contains spaces.

Operating Environment Information: On z/OS each *location* is a list of DCB and SMS options to be used when creating utility files. △

(*location-1 ... location-n*)
specifies a list of existing directories that can be accessed in parallel for utility files that are created by applications. A single utility file cannot span locations. Enclose a location in single or double quotation marks when the location contains spaces. Any location that does not exist is deleted from the value of the UTILLOC= system option.

Operating Environment Information: On z/OS, each *location* is a list of DCB and SMS options to be used when creating utility files. △

Requirement: If you have more than one location, then you must enclose the list of locations in parentheses.

Details

Thread-enabled SAS applications are able to create temporary utility files that can be accessed in parallel by separate threads.

For the SORT procedure, the UTILLOC= system option affects the placement of the utility files only if the multi-threaded SAS sort is used. The multi-threaded SAS sort can be invoked when the THREAD system option is specified and the value of the CPUCOUNT= system option is greater than 1. The multi-threaded SAS sort can also be invoked when you specify the THREADS option in the PROC SORT statement. The multi-threaded sort stores all temporary data in a single utility file within one of the locations that are specified by the UTILLOC= system option. The size of this utility file is proportional to the amount of data that is read from the input data set. A second utility file of the same size can be created in another of these locations when the amount of data that is read from the input data set is large or the amount of memory that is available to the SORT procedure is small.

See Also

System Option:

"CPUCOUNT= System Option" on page 1875

"THREADS System Option" on page 2037

The SORT Procedure in *Base SAS Procedures Guide*

"Support for Parallel Processing" in *SAS Language Reference: Concepts*.

UUIDCOUNT= System Option

Specifies the number of UUIDs to acquire from the UUID Generator Daemon.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

Syntax

UUIDCOUNT= *n* | MIN | MAX

Syntax Description

n
specifies the number of UUIDs to acquire. Zero indicates that the UUID Generator Daemon is not required.

Range: 0–1000

Default: 100

MIN | MAX

MIN specifies that the number of UUIDs to acquire is zero, indicating that the UUID Generator Daemon is not required.

MAX specifies that 1000 UUIDs at a time should be acquired from the UUID Generator Daemon.

Details

If a SAS application will generate a large number of UUIDs, this value can be adjusted at any time during a SAS session to reduce the number of times that the SAS session would have to contact the SAS UUID Generator Daemon.

See Also

System Option:

"UUIDGENDHOST= System Option" on page 2045

Function:

"UUIDGEN Function" on page 1181

"Universal Unique Identifiers and the Object Spawner" in *SAS Language Reference: Concepts*

UUIDGENDHOST= System Option

Identifies the host and port or the LDAP URL that the UUID Generator Daemon runs on.

Valid in: configuration file, SAS invocation

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

Syntax

UUIDGENDHOST= *'host-string'*

Syntax Description

'host-string'
is either of the form hostname:port or an LDAP URL. The value must be in one string. Enclose an LDAP URL string with quotation marks.

Details

SAS does not guarantee that all UUIDs are unique. Use the SAS UUID Generator Daemon (UUIDGEN) to ensure unique UUIDs.

Examples

- Specifying hostname:port as the *'host-string'*:

```
sas -UUIDGENDHOST 'myhost.com:5306'
```

or

```
sas UUIDGENDHOST= 'myhost.com:5306'
```

- Specifying an LDAP URL as the *'host-string'*:

```
"ldap://ldap-hostname/sasspawner-distinguished-name"
```

- A more detailed example of an LDAP URL as the *'host-string'*:

```
"ldap://ldaphost/sasSpawnercn=UUIDGEND,sascomponent=sasServer,
cn=ABC,o=ABC Inc,c=US"
```

- Specifying your binddn and password, if your LDAP server is secure:

```
"ldap://ldap-hostname/sasSpawner-distinguished-name????
bindname=binddn,password=bind-password"
```

- An example with a bindname value and a password value:

```
"ldap://ldaphost/
sasSpawnercn=UUIDGEND,sascomponent=sasServer,cn=ABC,o=ABC Inc,c=US
????bindname=cn=me%2co=ABC Inc %2cc=US,
password=itsme"
```

Note: When specifying your bindname and password, commas that are a part of your bindname and your password must be replaced with the string **"%2c"**. In the previous example, the bindname is as follows:

```
cn=me,o=ABC Inc,c=US
```

△

See Also

System Option:

"UUIDCOUNT= System Option" on page 2044

Function:

"UUIDGEN Function" on page 1181

V6CREATEUPDATE= System Option

Specifies the type of message to write to the SAS log when Version 6 data sets are created or updated.

Valid in: configuration file, SAS invocation

Category: Environment control: Files

PROC OPTIONS GROUP= SASFILES

Syntax

V6CREATEUPDATE = ERROR | NOTE | WARNING | IGNORE

Syntax Description

ERROR
specifies that an ERROR is written to the SAS log when the V6 engine is used to open a SAS data set for creation or update. The attempt to create or update a SAS data set in Version 6 format will fail. Reading Version 6 data sets will not generate an error.

NOTE
specifies that a NOTE is written to the SAS log when the V6 engine is used; all other processing occurs normally.

WARNING
specifies that a WARNING is written to the SAS log when the V6 engine is used; all other processing occurs normally.

IGNORE
disables the V6CREATEUPDATE= system option. Nothing is written to the SAS log when the V6 engine is used.

VALIDFMTNAME= System Option

Specifies the maximum size (32 characters or 8 characters) that user-created format and informat names can be before an error or warning is issued.

Default: LONG

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

VALIDFMTNAME=LONG | FAIL | WARN

Syntax Description

LONG
specifies that format and informat names can be up to 32 alphanumeric characters. This is the default.

FAIL
specifies that creating a format or informat name that is longer than eight characters results in an error message.

Tip: Specify this setting for using formats and informats that are valid in both SAS 9 and previous releases of SAS.

Interaction: If you explicitly specify the V7 or V8 Base SAS engine, such as in a LIBNAME statement, then SAS automatically uses the VALIDFMTNAME=FAIL behavior for data sets that are associated with those engines.

WARN
specifies that creating a format or informat name that is longer than eight characters results in a warning message to remind you that the format or informat cannot be used with releases before to SAS 9.

Details

SAS 9 enables you to define format and informat names up to 32 characters. Previous releases were limited to eight characters. The VALIDFMTNAME= system option applies to format and informat names in both data sets and format catalogs. VALIDFMTNAME= does not control the length of format and informat names. It only controls the length of format and informat names that you associate with variables when you create a SAS data set.

If a SAS data set has a variable with a long format or informat name, which means that a release before SAS 9 cannot read it, then you can remove the long name so that the data set can be accessed by an earlier release. However, in order to retain the format attribute of the variable, an identical format with a short name would have to be applied to the variable.

Note: After you create a format or informat using a name that is longer than eight characters, if you rename it using eight or fewer characters, a release before SAS 9

cannot use the format or informat. You must recreate the format or informat using the shorter name. △

See Also

For more information about SAS names, see "Names in the SAS Language" Names in the SAS Language in *SAS Language Reference: Concepts*.

For information about defining formats and informats, see "The FORMAT Procedure" in *Base SAS Procedures Guide*.

For information about compatibility issues, see "SAS 9.1 Compatibility with SAS Files From Earlier Releases" in *SAS Language Reference: Concepts*.

VALIDVARNAME= System Option

Specifies the rules for valid SAS variable names that can be created and processed during a SAS session.

Default: V7

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files

PROC OPTIONS GROUP= SASFILES

Syntax

VALIDVARNAME=V7 | UPCASE | ANY

Syntax Description

V7

specifies that variable names must follow these rules:

- can be up to 32 characters in length.
- must begin with a letter of the Latin alphabet (A - Z, a - z) or the underscore character. Subsequent characters can be letters of the Latin alphabet, numerals, or underscores.
- cannot contain blanks.
- cannot contain special characters except for the underscore.
- can contain mixed-case letters. SAS stores and writes the variable name in the same case that is used in the first reference to the variable. However, when SAS processes a variable name, SAS internally converts it to uppercase. You cannot, therefore, use the same variable name with a different combination of uppercase and lowercase letters to represent different variables. For example, **`cat`**, **`Cat`**, and **`CAT`** all represent the same variable.
- cannot be assigned the names of special SAS automatic variables (such as _N_ and _ERROR_) or variable list names (such as _NUMERIC_, _CHARACTER_, and _ALL_).

UPCASE
specifies that the variable name follows the same rules as V7, except that the variable name is uppercase, as in earlier versions of SAS.

ANY
specifies that SAS variable names must follow these rules:

- can be up to 32 characters in length
- can be special and multi-byte characters not to exceed 32 bytes.
- cannot contain any null bytes
- leading blanks are preserved, but trailing blanks are ignored
- name must contain at least one character. An all blank name is not permitted.
- can begin with or contain any characters, including blanks

  Note: If you use any characters other than the ones that are valid when the VALIDVARNAME system option is set to V7 (letters of the Latin alphabet, numerals, or underscores), then you must express the variable name as a *name literal* and you must set VALIDVARNAME=ANY. See "SAS Name Literals" and "Avoiding Errors When Using Name Literals" in *SAS Language Reference: Concepts*.

  If you use either the percent sign (%) or the ampersand (&), then you must use single quotation marks in the name literal in order to avoid interaction with the SAS Macro Facility. △
- can contain mixed-case letters. SAS stores and writes the variable name in the same case that is used in the first reference to the variable. However, when SAS processes a variable name, SAS internally converts it to uppercase. You cannot, therefore, use the same variable name with a different combination of uppercase and lowercase letters to represent different variables. For example, **`cat`**, **`Cat`**, and **`CAT`** all represent the same variable.

Warning: The intent of the VALIDVARNAME=ANY option is to enable compatibility with other DBMS variable (column) naming conventions, such as allowing embedded blanks and national characters. Throughout SAS, using the name literal syntax with variable names that exceed the 32–byte limit or have excessive embedded quotation marks might cause unexpected results.

See Also

"Rules for Words and Names in the SAS Language" in *SAS Language Reference: Concepts*

VARLENCHK= System Option

Specifies the type of message to write to the SAS log when the input data set is read using the SET, MERGE, UPDATE, or MODIFY statements.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Files: SAS Files
PROC OPTIONS GROUP= SASFILES

Syntax

VARLENCHK=NOWARN | WARN | ERROR

Syntax Description

NOWARN
specifies that no warning message is issued when the length of a variable that is being read is larger than the length that is defined for the variable.

WARN
specifies that a warning is issued when the length of a variable that is being read is larger than the length that is defined for the variable. This is the default.

ERROR
specifies that an error message is issued when the length of a variable that is being read is larger than the length that is defined for the variable.

Details

After a variable is defined, the length of a variable can be changed only by a LENGTH statement. If a variable is read by the SET, MERGE, UPDATE, or MODFIY statements and the length of the variable is longer than a variable of the same name, SAS issues a warning message and uses the shorter, original length of the variable. By using the shorter length, data will not be truncated.

When you intentionally truncate data, perhaps to remove unnecessary blanks from character variables, SAS issues a warning message that might not be useful to you. To make it so that SAS does not issue the warning message or set a nonzero return code, you can set the VARLENCHK= system option to NOWARN. When VARLENCHK=NOWARN, SAS does not issue a warning message and sets the return code to SYSRC=0.

Alternatively, if you set VARLENCHK=ERROR and the length of a variable that is being read is larger than the length that is defined for the variable, SAS issues an error and sets the return code SYSRC=8.

Examples

Example 1: SAS Issues a Warning Message Merging Two Data Sets with Different Variable Lengths This example merges two data sets, the sashelp.class data set and the exam_schedule data set. The length of the variable Name is set to 8 by the first SET statement, **`set sashelp.class;`**. The exam_schedule data set sets the length of Name to 10. When exam_schedule is read in the second SET statement, **`set exam_schedule key=Name;`**, SAS issues a warning message because the length of Name in the exam_schedule data set is longer than the length of Name in the sashelp.class data set, and data might have been truncated.

```
/& Create the exam_schedule data set. */

data exam_schedule(index=(Name));
```

```
  input Name $10. +1 Exam_Date mmddyy10.;
  format Exam_Date mmddyy10.;
datalines;
Carol      06/09/2008
Hui         06/09/2008
Janet      06/09/2008
Geoffrey 06/09/2008
John        06/09/2008
Joyce       06/09/2008
Helga      06/09/2008
Mary       06/09/2008
Roberto   06/09/2008
Ronald     06/09/2008
Barbara    06/10/2008
Louise     06/10/2008
Alfred     06/11/2008
Alice      06/11/2008
Henri      06/11/2008
James      06/11/2008
Philip     06/11/2008
Tomas     06/11/2008
William    06/11/2008


/* Merge the data sets sashelp.class and exam_schedule  */

data exams;
  set sashelp.class;
  set exam_schedule key=Name;
run;
```

The following SAS log shows the warning message:

Output 7.11 The Warning Message in the SAS Log

```
1    data exam_schedule(index=(Name));
2      input Name $10. +1 Exam_Date mmddyy10.;
3      format Exam_Date mmddyy10.;
4    datalines;

NOTE: The data set WORK.EXAM_SCHEDULE has 20 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           4.32 seconds
      cpu time            0.24 seconds


25   ;
26
27   data exams;
28     set sashelp.class;
29     set exam_schedule key=Name;
30   run;

WARNING: Multiple lengths were specified for the variable Name by input data set(s). This may cause truncation of data.
NOTE: There were 19 observations read from the data set SASHELP.CLASS.
NOTE: The data set WORK.EXAMS has 19 observations and 6 variables.
NOTE: DATA statement used (Total process time):
      real time           0.51 seconds
      cpu time            0.00 seconds
```

Example 2: Turn Off the Warning Message and Use the LENGTH Statement to Match Variable Lengths In order to merge the two data sets, sashelp.class and exam_schedule, you can examine the values of Name in exam_schedule. You can see that there are no values that are greater than 8 and that you can change the length of Name without losing data.

To change the length of the variable Name, you use a LENGTH= statement in a DATA step before the **set exam_schedule;** statement. If the value of VARLENCHK is WARN (the default), SAS issues the warning message that the value of Name is truncated when it is read from work.exam_schedule. Because you know that data is not lost, you might want to turn the warning message off:

```
options varlenchk=nowarn;
data exam_schedule(index=(Name));
  length Name $ 8;
  set exam_schedule;
run;
```

The following is the SAS log output:

```
37   options varlenchk=nowarn;

38   options varlenchk=nowarn;
39   data exam_schedule(index=(Name));
40     length Name $ 8;
41     set exam_schedule;
42   run;

NOTE: There were 20 observations read from the data set WORK.EXAM_SCHEDULE.
NOTE: The data set WORK.EXAM_SCHEDULE has 20 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds
```

See Also

Looking at Sources of Common Problems in the "Combining SAS Data Sets: Basic Concepts" section of *SAS Language Reference: Concepts*

VIEWMENU System Option

Specifies whether the View menu is included in SAS windows.

Default: VIEWMENU

Valid in: configuration file, SAS invocation

Category: Environment control: Display

PROC OPTIONS GROUP= ENVDISPLAY

Syntax

VIEWMENU | NOVIEWMENU

Syntax Description

VIEWMENU
specifies that the View menu is included in SAS windows.

NOVIEWMENU
specifies that the View menu is not included in SAS windows.

VNFERR System Option

Specifies whether SAS issues an error or warning when a BY variable exists in one data set but not another data set when processing the SET, MERGE, UPDATE, or MODIFY statements.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Error handling

PROC OPTIONS GROUP= ERRORHANDLING

Syntax

VNFERR | NOVNFERR

Syntax Description

VNFERR
specifies that SAS issue an error when a BY variable exists in one data set but not in another data set when processing the SET, MERGE, UPDATE, or MODIFY statements. When the error occurs, SAS enters into syntax-check mode.

NOVNFERR
specifies that SAS issue a warning when a BY variable exists in one data set but not in another data set when processing the SET, MERGE, UPDATE, or MODIFY statements. When the warning occurs, SAS does not enter into syntax-check mode.

Details

Operating Environment Information: Under z/OS, SAS also issues an error or a warning when the data set specified by DDNAME points to a DUMMY library. △

Comparisons

- VNFERR is similar to the BYERR system option, which issues an error and enters into syntax-check mode if the SORT procedure attempts to sort a _NULL_ data set.
- VNFERR is similar to the DSNFERR system option, which issues an error when a SAS data set is not found.

See Also

System Options:

"BYERR System Option" on page 1855

"DSNFERR System Option" on page 1892

Syntax Check Mode in *SAS Language Reference: Concepts*

WORK= System Option

Specifies the WORK data library.

Valid in: configuration file, SAS invocation

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

See: WORK= System Option in the documentation for your operating environment.

Syntax

WORK=*library-specification*

Syntax Description

library-specification
specifies the libref or physical name of the storage space where all data sets with one-level names are stored. This library must exist.

Operating Environment Information: A valid library specification and its syntax are specific to your operating environment. On the command line or in a configuration file, the syntax is specific to your operating environment. For details, see the SAS documentation for your operating environment. △

Details

This library is deleted at the end of your SAS session by default. To prevent the files from being deleted, specify the NOWORKTERM system option.

See Also

System Option:

"WORKTERM System Option" on page 2057

WORKINIT System Option

Specifies whether to initialize the WORK library at SAS invocation.

Valid in: configuration file, SAS invocation
Category: Environment control: Files
PROC OPTIONS GROUP= ENVFILES

Syntax

WORKINIT | NOWORKINIT

Syntax Description

WORKINIT
erases files that exist from a previous SAS session in an existing WORK library at SAS invocation.

NOWORKINIT
does not erase files from the WORK library at SAS invocation.

Comparisons

The WORKINIT system option initializes the WORK data library and erases all files from a previous SAS session at SAS invocation. The WORKTERM system option controls whether SAS erases WORK files at the end of a SAS session.

See Also

System Option:
"WORKTERM System Option" on page 2057

Operating Environment Information: WORKINIT has behavior and functions specific to the UNIX operating environment. For details, see the SAS documentation for the UNIX operating environment. △

WORKTERM System Option

Specifies whether to erase the WORK files when SAS terminates.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Environment control: Files

PROC OPTIONS GROUP= ENVFILES

Syntax

WORKTERM | NOWORKTERM

Syntax Description

WORKTERM
erases the WORK files at the termination of a SAS session.

NOWORKTERM
does not erase the WORK files.

Details

Although NOWORKTERM prevents the WORK data sets from being deleted, it has no effect on initialization of the WORK library by SAS. SAS normally initializes the WORK library at the start of each session, which effectively destroys any pre-existing information.

Comparisons

Use the NOWORKINIT system option to prevent SAS from erasing existing WORK files on invocation. Use the NOWORKTERM system option to prevent SAS from erasing existing WORK files on termination.

See Also

System Option:

"WORKINIT System Option" on page 2056

YEARCUTOFF= System Option

Specifies the first year of a 100-year span that is used by date informats and functions to read a two–digit year.

Valid in: configuration file, SAS invocation, OPTIONS statement, SAS System Options window

Category: Input control: Data Processing

PROC OPTIONS GROUP= INPUTCONTROL

Syntax

YEARCUTOFF= *nnnn* | *nnnnn*

Syntax Description

nnnn* | *nnnnn

specifies the first year of the 100-year span.

Range: 1582–19900

Default: 1920

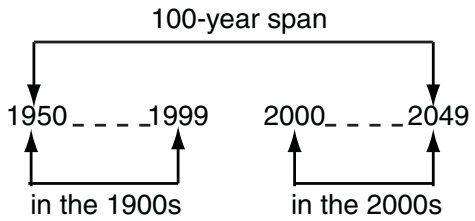
Details

The YEARCUTOFF= value is the default that is used by various date and datetime informats and functions.

If the default value of *nnnn* (1920) is in effect, the 100-year span begins with 1920 and ends with 2019. Therefore, any informat or function that uses a two-digit year value that ranges from 20 to 99 assumes a prefix of 19. For example, the value 92 refers to the year 1992.

The value that you specify in YEARCUTOFF= can result in a range of years that span two centuries. For example, if you specify YEARCUTOFF=1950, any two-digit value between 50 and 99 inclusive refers to the first half of the 100-year span, which is in the 1900s. Any two-digit value between 00 and 49, inclusive, refers to the second half of the 100-year span, which is in the 2000s. The following figure illustrates the relationship between the 100-year span and the two centuries if YEARCUTOFF=1950.

Figure 7.1 A 100–Year Span with Values in Two Centuries



Note: YEARCUTOFF= has no effect on existing SAS dates or dates that are read from input data that include a four-digit year, except years with leading zeros. For example, 0076 with yearcutoff=1990 indicates 2076. △

Operating Environment Information: The syntax that is shown here applies to the OPTIONS statement. On the command line or in a configuration file, the syntax is specific to your operating environment. For more information, see the SAS documentation for your operating environment. △

See Also

"Year 2000" in *SAS Language Reference: Concepts*.

SAS System Options Documented in Other SAS Publications

In addition to system options documented in *SAS Language Reference: Dictionary*, system options are also documented in the following publications:

Encryption in SAS

System Option	Description
NETENCRYPT=	Specifies whether client/server data transfers are encrypted.
NETENCRYPTALGORITHM=	Specifies one or more algorithms to be used for encrypted client/server data transfers.
NETENCRYPTKEYLEN=	Specifies the key length to use for encrypted client/server data transfers.
SSLCALISTLOC=	Specifies the location of digital certificates for trusted certification authorities (CA).

System Option	Description
SSLCERTISS=	Specifies the name of the issuer of the digital certificate that SSL should use.
SSLCERTLOC=	Specifies the location of the digital certificate that is used for authentication.
SSLCERTSERIAL=	Specifies the serial number of the digital certificate that SSL should use.
SSLCERTSUBJ=	Specifies the subject name of the digital certificate that SSL should use.
SSLCLIENTAUTH=	Specifies whether a server should perform client authentication.
SSLCRLCHECK=	Specifies whether a Certificate Revocation List (CRL) is checked when a digital certificate is validated.
SSLCRLLOC=	Specifies the location of a Certificate Revocation List (CRL).
SSLPVTKEYLOC=	Specifies the location of the private key that corresponds to the digital certificate.
SSLPVTKEYPASS=	Specifies the password that SSL requires for decrypting the private key.

Grid Computing in SAS

System Option	Description
CONNECTMETACONNECTION	Specifies whether a SAS/CONNECT server is authorized to access a SAS Metadata Server at server sign-on.
IPADDRESS	Specifies whether the grid node sends its IP address to the client session during sign-on to the grid.
SSPI	Enables a SAS session that runs on a grid node to access the SAS Metadata Server using credentials that are supplied by Windows SSPI (Security Provider Interface).

For more information, see *Grid Computing in SAS 9.2* on http://support.sas.com.

SAS Interface to Application Response Measurement (ARM): Reference

System Option	Description
ARMAGENT=	Specifies another vendor's ARM agent, which is an executable module that contains a vendor's implementation of the ARM API.
ARMLOC=	Specifies the location of the ARM log.
ARMSUBSYS=	Specifies whether to enable or disable the ARM subsystems that determine the internal SAS processing transactions to be logged.

SAS Companion for Windows

The system options listed here are documented only in *SAS Companion for Windows*. Other system options in *SAS Companion for Windows* contain information specific to the Windows operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter system options are not listed here.

System Option	Description
ACCESSIBILITY	Enables the accessibility features on the Customize Tools dialog box.
ALTLOG	Specifies a destination for a copy of the SAS log.
ALTPRINT	Specifies the destination for the copies of the output files from SAS procedures.
AUTHSERVER	Specifies the authentication domain server to search for secure server logins.
AUTOEXEC	Specifies the SAS autoexec file.
AWSCONTROL	Specifies whether the main SAS window includes a title bar, a system/control menu, and minimize/maximize buttons.
AWSDEF	Specifies the location and dimensions of the main SAS window when SAS initializes.
AWSMENU	Specifies whether to display the menu bar in the main SAS window.
AWSMENUMERGE	Specifies whether to embed menu items that are specific to Windows in the main menus.
AWSTITLE	Replaces the default text in the main SAS title bar.
COMDEF	Specifies the location where the SAS Command window is displayed.
CONFIG	Specifies the configuration file that is used when initializing or overriding the values of SAS system options.
ECHO	Specifies a message to be echoed to the SAS log while initializing SAS.
EMAILDLG	Specifies whether to use the native e-mail dialog box provided by your e-mail application or the e-mail dialog box provided by SAS.
EMAILSYS	Specifies the e-mail protocol to use for sending electronic mail.

System Option	Description
ENHANCEDEDITOR	Specifies whether to enable the Enhanced Editor during SAS invocation.
FILTERLIST	Specifies an alternative set of file filter specifications to use for the Open and Save As dialog boxes.
FONT	Specifies a font to use for SAS windows.
FONTALIAS	Assigns a Windows font to one of the SAS fonts.
FULLSTIMER	Specifies whether to write all available system performance statistics to the SAS log.
HELPINDEX	Specifies one or more index files for the SAS Help and Documentation.
HELPLOC	Specifies the location of Help files that are used to view SAS Help and Documentation using Microsoft HTML Help.
HELPREGISTER	Registers help files to access from the main SAS window Help menu.
HELPTOC	Specifies the table of contents files for the SAS Help and Documentation.
HOSTPPRINT	Specifies that the Windows Print Manager is to be used for printing.
ICON	Minimizes the SAS window.
JREOPTIONS	Identifies Java Runtime Environment (JRE) options for SAS.
LOADMEMSIZE	Specifies a suggested amount of memory needed for executable programs loaded by SAS.
LOG	Specifies a destination for a copy of the SAS log when running in batch mode.
MAXMEMQUERY	Specifies the limit on the maximum amount of memory that is allocated for procedures.
MEMBLKSZ	Specifies the memory block size for memory-based libraries for Windows operating environments.
MEMCACHE	Specifies to use the memory-based libraries as a SAS file cache.
MEMLIB	Specifies to process the Work library as a memory-based library.
MEMMAXSZ	Specifies the maximum amount of memory to allocate for using memory-based libraries in Windows operating environments.
MEMSIZE	Specifies the limit on the amount of virtual memory that can be used during a SAS session.
MSG	Specifies the library that contains the SAS error messages.
MSGCASE	Specifies whether notes, warnings, and error messages that are generated by SAS are displayed in uppercase characters.
NUMKEYS	Controls the number of available function keys.
NUMMOUSEKEYS	Specifies the number of mouse buttons SAS displays in the KEYS window.
PATH	Specifies one or more search paths for SAS executable files.
PFKEY	Specifies which set of function keys to designate as the primary set of function keys.
PRINT	Specifies a destination for SAS output when running in batch mode.

System Option	Description
PRNGETLIST	Specifies if printers attached to the system are recognized.
PRTABORTDLGS	Specifies when to display the Print Abort dialog box.
PRTPERSISTDEFAULT	Specifies to use the same destination printer from SAS session to SAS session.
PRTSETFORMS	Specifies whether to include the Use Forms check box in the Print Setup dialog box.
REALMEMSIZE	Specifies the amount of virtual memory SAS can expect to allocate.
REGISTER	Adds an application to the Tools menu in the main SAS window.
RESOURCESLOC	Specifies a directory location of the files that contain SAS resources.
RTRACE	Produces a list of resources that are read or loaded during a SAS session.
RTRACELOC	Specifies the pathname of the file to which the list of resources that are read or loaded during a SAS session is written.
SASCONTROL	Specifies whether the SAS application windows include system/ control menus and minimize/maximize buttons.
SASINITIALFOLDER	Changes the working folder and the default folders for the Open and Save As dialog boxes to the specified folder after SAS initialization is complete.
SCROLLBARFLASH	Specifies whether to allow the mouse or keyboard to focus on a scroll bar.
SET	Defines a SAS environment variable.
SGIO	Activates the Scatter/Gather I/O feature.
SLEEPWINDOW	Enables or disables the SLEEP window.
SORTANOM	Specifies certain options for the SyncSort utility.
SORTCUT	Specifies the number of observations above which SyncSort is used instead of the SAS sort program.
SORTCUTP	Specifies the number of bytes above which SyncSort is used instead of the SAS sort program.
SORTDEV	Specifies the pathname used for temporary files created by the SyncSort utility.
SORTPARM	Specifies parameters for the SyncSort utility.
SORTPGM	Specifies the sort utility that is used in the SORT procedure.
SPLASH	Specifies whether to display the splash screen (logo screen) when SAS starts.
SPLASHLOC	Specifies the location of the splash screen bitmap that appears when SAS starts.
STIMEFMT	Specifies the format to use for displaying the time on STIMER output.
STIMER	Writes a subset of system performance statistics to the SAS log.
SYSGUIFONT	Specifies a font to use for the button text and the descriptive text.
SYSPRINT	Specifies a destination printer for printing SAS output.
SYSIN	Specifies a batch mode source file.

System Option	Description
TOOLDEF	Specifies the Toolbox display location.
UPRINTMENUSWITCH	Enables the universal print commands in the File menu.
USERICON	Specifies the pathname of the resource file associated with your user-defined icon.
VERBOSE	Controls whether SAS writes the settings of all the system options specified in the configuration file to either the terminal or the batch log.
WEBUI	Specifies to enable Web enhancements.
WINDOWSMENU	Specifies to include or suppress the Window menu in windows that display menus.
XCMD	Specifies that the X command is valid in the current SAS session.
XMIN	Specifies to open the application specified in the X command in a minimized state or in the default active state.
XSYNC	Controls whether an X command or statement executes synchronously or asynchronously.
XWAIT	Specifies whether you have to type EXIT at the DOS prompt before the DOS shell closes.

SAS Companion for OpenVMS on HP Integrity Servers

The system options listed here are documented only in *SAS Companion for OpenVMS on HP Integrity Servers*. Other system options in *SAS Companion for OpenVMS on HP Integrity Servers* contain information specific to the OpenVMS operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter system options are not listed here.

System Option	Description
ALTMULT	Specifies the number of pages that are preallocated to a file.
ALTLOG	Specifies a destination for a copy of the SAS log.
ALTPRINT	Specifies the destination for the copies of the output files from SAS procedures.
APPLETLOC	Specifies the location of Java applets.
AUTOEXEC	Specifies the SAS autoexec file.
CACHENUM	Specifies the number of caches used per SAS file.
CACHESIZE	Specifies the size of cache that is used for each open SAS file.
CC	Tells SAS what type of carriage control to use when it writes to external files.
CONFIG	Specifies the configuration file that is used when initializing or overriding the values of SAS system options.
DEQMULT	Specifies the number of pages to extend a file.

System Option	Description
DETACH	Specifies that the asynchronous host command uses a detached process.
DUMP	Specifies when to create a process dump file.
EDITCMD	Specifies the host editor to be used with the HOSTEDIT command.
EMAILSYS	Specifies the e-mail protocol to use for sending electronic mail.
EXPANDLNM	Specifies whether concealed logical names are expanded when libref paths are displayed to the user.
FILECC	Specifies how to treat data in column 1 of a print file.
FULLSTIMER	Specifies whether to write all available system performance statistics to the SAS log.
GSFCC	Tells SAS what type of carriage control to use for writing to graphics stream files.
HELPHOST	Specifies the name of the local computer where the remote browsing system is to be displayed.
HELPINDEX	Specifies one or more index files for the SAS Help and Documentation.
HELPLOC	Specifies the location of the text and index files for the facility that is used to view SAS Help and Documentation.
HELPTOC	Specifies the table of contents files for the SAS Help and Documentation.
JREOPTIONS	Identifies the Java Runtime Environment (JRE) options for SAS.
LOADLIST	Specifies whether to print to the specified file the information about images that SAS has loaded into memory.
LOG	Specifies a destination for a copy of the SAS log when running in batch mode.
LOGMULTREAD	Specifies the session log file to be opened for shared read access.
MEMSIZE	Specifies the limit on the total amount of memory that can be used by a SAS session.
MSG	Specifies the library that contains SAS error messages.
MSGCASE	Specifies whether notes, warnings, and error messages that are generated by SAS are displayed in uppercase characters.
OPLIST	Specifies whether the settings of the SAS system options are written to the SAS log.
PRINT	Specifies a destination for SAS output when running in batch mode.
REALMEMSIZE	Specifies the amount of real memory SAS can expect to allocate.
SORTPGM	Specifies whether SAS sorts using use the SAS sort utility or the host sort utility.
SORTWORK	Defines locations for host sort work files.
SPAWN	Specifies that SAS is invoked in a SPAWN/NOWAIT subprocess.
STIMEFMT	Specifies the format that is used to display time on STIMER output.

System Option	Description
STIMER	Specifies whether to write a subset of system performance statistics to the SAS log.
SYSIN	Specifies the default location of SAS source programs.
SYSPRINT	Specifies the destination for printed output.
TERMIO	Specifies whether terminal I/O is blocking or non-blocking.
USER	Specifies the default permanent SAS library.
VERBOSE	Specifies whether SAS writes the system options that are set when SAS starts to the VMS computer in the SAS windowing environment or, in batch, to the batch log.
WORKCACHE	Specifies the size of the I/O data cache allocated for a file in the WORK library.
XCMD	Specifies whether the X command is valid in the SAS session.
XCMDWIN	Specifies whether to create a DECTERM window for X command output when in the SAS windowing environment.
XKEYPAD	Specifies that subprocesses use the keypad settings that were in effect before you invoked SAS.
XLOG	Specifies whether to display the output from the X command in the SAS log file.
XLOGICAL	Specifies that process-level logical names are passed to the subprocess that is spawned by an X statement or X command.
XOUTPUT	Specifies whether to display the output from the X command.
XRESOURCES	Specifies a character string of X resource options or the application instance name for the SAS interface to Motif.
XSYMBOL	Specifies that global symbols are passed to the subprocess that is spawned by an X statement or X command.
XTIMEOUT	Specifies how long a subprocess that has been spawned by an X statement or X command remains inactive before being deleted.

SAS Companion for UNIX Environments

The system options listed here are documented only in *SAS Companion for UNIX Environments*. Other system options in *SAS Companion for UNIX Environments* contain information specific to the UNIX operating environment, where the main documentation is in *SAS Language Reference: Dictionary*. These latter system options are not listed here.

System Option	Description
ALTLOG	Specifies a destination for a copy of the SAS log.
ALTPRINT	Specifies the destination for the copies of the output files from SAS procedures.
AUTOEXEC	Specifies the SAS autoexec file.

System Option	Description
CONFIG	Specifies the configuration file that is used when initializing or overriding the values of SAS system options.
ECHO	Specifies a message to be echoed to the computer.
EDITCMD	Specifies the host editor to be used with the HOSTEDIT command.
EMAILSYS	Specifies the e-mail protocol to use for sending electronic mail.
FILELOCKS	Specifies whether external file locking is turned on or off and what action should be taken if a file cannot be locked.
FILELOCKWAITMAX	Sets an upper limit on the time SAS will wait for a locked file.
FULLSTIMER	Specifies whether to write all available system performance statistics and the datetime stamp to the SAS log.
HELPINDEX	Specifies one or more index files for the SAS Help and Documentation.
HELPLOC	Specifies the location of the text and index files for the facility that is used to view SAS Help and Documentation.
HELPTOC	Specifies the location of the table of contents files for the SAS Help and Documentation.
JREOPTIONS	Identifies the Java Runtime Environment (JRE) options for SAS.
LOG	Specifies a destination for a copy of the SAS log when running in batch mode.
LPTYPE	Specifies which UNIX command and options settings will be used to route files to the printer.
MAXMEMQUERY	Specifies the maximum amount of memory that is allocated per request for certain procedures.
MEMSIZE	Specifies the limit on the total amount of virtual memory that can be used by a SAS session.
MSG	Specifies the library that contains the SAS error messages.
MSGCASE	Specifies whether notes, warnings, and error messages that are generated by SAS are displayed in uppercase characters.
OPTLIST	Specifies whether the settings of the SAS system options are written to the SAS log.
PATH	Specifies one or more search paths for SAS executable files.
PRINT	Specifies a destination for SAS output when running in batch mode.
PRINTCMD	Specifies the print command SAS is to use.
REALMEMSIZE	Specifies the amount of real (physical) memory SAS can expect to allocate.
RTRACE	Produces a list of resources that are read or loaded during a SAS session.
RTRACELOC	Specifies the pathname of the file to which the list of resources that are read or loaded during a SAS session is written.
SASSCRIPT	Specifies one or more storage locations of SAS/CONNECT script files.
SET	Defines an environment variable.
SORTANOM	Specifies certain options for the host sort utility.

System Option	Description
SORTCUT	Specifies the number of observations that SAS sorts. If the number of observation in the data set is greater than the specified number, the host sort program sorts the remaining observations.
ojbedSORTCUTP	Specifies the number of bytes that SAS sorts. If the number of bytes in the data set is greater than the specified number, the host sort program sorts the remaining data set.
SORTDEV	Specifies the pathname used for temporary files created by the host sort utility.
SORTNAME	Specifies the name of the host sort utility.
SORTPARM	Specifies parameters for the host sort utility.
SORTPGM	Specifies whether SAS sorts using the SAS sort utility or the host sort utility.
STDIO	Specifies whether SAS should use stdin, stdout, and stderr.
STIMEFMT	Specifies the format that is used to display the time on FULLSTIMER and STIMER output.
STIMER	Specifies whether to write a subset of system performance statistics to the SAS log.
SYSIN	Specifies the default location of SAS source code when running in batch mode.
SYSPRINT	Specifies the destination for printed output.
VERBOSE	Specifies whether SAS writes the system option settings to the SAS log.
WORKPERMS	Sets the permissions of the SAS Work library when it is initially created.
XCMD	Specifies whether the X command is valid in the SAS session.

SAS Companion for z/OS

The system options listed here are documented only in *SAS Companion for z/OS*. Other system options in *SAS Companion for z/OS* contain information specific to the z/OS operating environment, where the main documentation is in *SAS Language Reference: Dictionary* . These latter system options are not listed here.

System Option	Description
ALTLOG=	Specifies a destination for a copy of the SAS log.
ALTPRINT=	Specifies the destination for the copies of the output files from SAS procedures.
APPEND=	Appends the specified value to the existing value of the specified system option.
AUTOEXEC=	Specifies the SAS autoexec file.
BLKALLOC	Causes SAS to set LRECL and BLKSIZE values for a SAS library when it is allocated rather than when it is first accessed.

System Option	Description
BLKSIZE=	Specifies the default block size for SAS libraries.
BLSKIZE(device-type)=	Specifies the default block size for SAS libraries by *device-type*.
CAPSOUT	Specifies that all output is to be converted to uppercase.
CHARTYPE=	Specifies a character set or screen size to use for a device.
CLIST	Specifies that SAS obtains its input from a CLIST.
CONFIG=	Specifies the configuration file that is used when initializing or overriding the values of SAS system options.
DLDISPCHG	Controls changes in allocation disposition for an existing library data set.
DLDSNTYPE	Specifies the default value of the DSNTYPE LIBNAME option.
DLEXCPCOUNT	Reports number of EXCPs to direct access bound SAS libraries.
DLHFSDIRCREATE	Creates an HFS directory for a SAS library that is specified with LIBNAME if the library does not exist.
DLMSGLEVEL=	Specifies the level of messages to generate for SAS libraries.
DLSEQDSNTYPE	Specifies the default value of the DSNTYPE LIBNAME option for sequential-format disk files.
DLTRUNCHK	Enables checking for SAS library truncation.
DLRESV	Requests exclusive use of shared disk volumes when accessing partitioned data sets on shared disk volumes.
DYNALLOC	Controls whether SAS or the host sort utility allocates sort work data sets.
ECHO=	Specifies a message to be echoed to the SAS log while initializing SAS.
EMAILSYS=	Specifies the e-mail protocol to use for sending electronic mail.
FILEAUTHDEFER	Controls whether SAS performs file authorization checking for z/OS data sets or defers authorization checking to z/OS system services such as OPEN.
FILEBLKSIZE(device-type)=	Specifies the default maximum block size for external files.
FILECC	Specifies whether to treat data in column 1 of a printer file as carriage-control data when reading the file.
FILEDEST=	Specifies the default printer destination.
FILEDEV=	Specifies the device name used for allocating new physical files.
FILEDIRBLK=	Specifies the number of default directory blocks to allocate for new partitioned data sets.
FILEEXT=	Specifies how to handle file extensions when accessing members of partitioned data sets.
FILEFORMS=	Specifies the default SYSOUT form for a print file.
FILELBI	Controls the use of the z/OS Large Block Interface support for BSAM and QSAM files, as well as files on tapes that have standard labels.

System Option	Description
FILELOCKS=	Specifies the default SAS system file locking that is to be used for external files (both USS and native MVS). Also specifies the operating system file locking to be used for USS files (both SAS files and external files).
FILEMOUNT	Specifies whether an off-line volume is to be mounted.
FILEMSGS	Controls whether you receive expanded dynamic allocation error messages when you are assigning a physical file.
FILENULL	Specifies whether zero-length records are written to external files.
FILEPROMPT	Controls whether you are prompted if you reference a data set that does not exist.
FILEREUSE	Specifies whether to reuse an existing allocation for a file that is being allocated to a temporary ddname.
FILESEQDSNTYPE	Specifies the default value that is assigned to DSNTYPE when it is not specified with a filename statement, a DD statement, or a TSO ALLOC command.
FILESPPRI=	Specifies the default primary space allocation for new physical files.
FILESPEC=	Specifies the default secondary space allocation for new physical files.
FILESTAT	Specifies whether ISPF statistics are written.
FILESYSOUT=	Specifies the default SYSOUT CLASS for a printer file.
FILESYSTEM=	Specifies the default file system used when the filename is ambiguous.
FILEUNIT=	Specifies the default unit of allocation for new physical files.
FILEVOL=	Specifies which VOLSER to use for new physical files.
FILSZ	Specifies that the host sort utility supports the FILSZ parameter.
FSBCOLOR	Specifies whether you can set background colors in SAS windows on vector graphics devices.
FSBORDER=	Specifies what type of symbols are to be used in borders.
FSDEVICE=	Specifies the full-screen device driver for your terminal.
FSMODE=	Specifies the full-screen data stream type.
FULLSTATS	Specifies whether to write all available system performance statistics to the SAS log.
GHFONT=	Specifies the default graphics hardware font.
HELPCASE	Controls how text is displayed in the help browser.
HELPHOST	Specifies the name of the computer where the remote help browser is running.
HELPLOC=	Specifies the location of the text and index files for the facility that is used to view SAS Help and Documentation.
HSLXTNTS=	Specifies the size of each physical hyperspace that is created for a SAS library.
HSMAXPGS=	Specifies the maximum number of hyperspace pages allowed in a SAS session.

System Option	Description
HSMAXSPC=	Specifies the maximum number of hyperspaces allowed in a SAS session.
HSSAVE	Controls how often the DIV data set pages are updated when a DIV data set backs a hyperspace library.
HSWORK	Tells SAS to place the WORK library in a hyperspace.
INSERT	Inserts the specified value at the beginning of the specified system option.
ISPCAPS	Specifies whether to convert to uppercase printable ISPF parameters that are used in CALL ISPEXEC and CALL ISPLINK.
ISPCHARF	Specifies whether the values of SAS character variables are converted using their automatically specified informats or formats each time they are used as ISPF variables.
ISPCSR=	Tells SAS to set an ISPF variable to the name of a variable whose value is found to be invalid.
ISPEXECV=	Specifies the name of an ISPF variable that passes its value to an ISPF service.
ISPMISS=	Specifies the value assigned to SAS character variables defined to ISPF when the associated ISPF variable has a length of zero.
ISPMSG=	Tells SAS to set an ISPF variable to a message ID when a variable is found to be invalid.
ISPNOTES	Specifies whether ISPF error messages are to be written to the SAS log.
ISPNZTRC	Specifies whether nonzero ISPF service return codes are to be written to the SAS log.
ISPPPT	Specifies whether ISPF parameter value pointers and lengths are to be written to the SAS log.
ISPTRACE	Specifies whether the parameter lists and service return codes are to be written to the SAS log.
ISPVDEFA	Specifies whether all current SAS variables are to be identified to ISPF via the SAS VDEFINE user exit.
ISPVDLT	Specifies whether VDELETE is executed before each SAS variable is identified to ISPF via VDEFINE.
ISPVDTRC	Specifies whether to trace every VDEFINE for SAS variables.
ISPVIMSG=	Specifies the ISPF message ID that is to be set by the SAS VDEFINE user exit when the informat for a variable returns a nonzero return code.
ISPVRMSG=	Specifies the ISPF message ID that is to be set by the SAS VDEFINE user exit when a variable has a null value.
ISPVTMSG=	Specifies the ISPF message ID that is to be displayed by the SAS VDEFINE user exit when the ISPVTRAP option is in effect.
ISPVTNAM=	Restricts the information that is displayed by the ISPVTRAP option to the specified variable only.
ISPVTPNL=	Specifies the name of the ISPF panel that is to be displayed by the SAS VDEFINE user exit when the ISPVTRAP option is in effect.

System Option	Description
ISPVTRAP	Specifies whether the SAS VDEFINE user exit is to write information to the SAS log (for debugging purposes) each time it is entered.
ISPVTVARS=	Specifies the prefix for the ISPF variables to be set by the SAS VDEFINE user exit when the ISPVTRAP option is in effect.
JREOPTIONS=	Identifies the Java Runtime Environment (JRE) options for SAS.
LOG=	Specifies a destination for a copy of the SAS log when running in batch mode.
MEMLEAVE=	Specifies the amount of memory in the user's region that is reserved exclusively for the use of the operating environment.
MEMRPT	Specifies whether memory usage statistics are to be written to the SAS log for each step.
MEMSIZE=	Specifies the limit on the total amount of memory that can be used by a SAS session.
MINSTG	Tells SAS whether to minimize its use of storage.
MSG=	Specifies the library that contains the SAS error messages.
MSGCASE	Specifies whether notes, warnings, and error messages that are generated by SAS are displayed in uppercase characters.
MSGSIZE=	Specifies the size of the message cache.
OPLIST	Specifies whether the settings of the SAS system options are written to the SAS log.
PFKEY=	Specifies which set of function keys to designate as the primary set of function keys.
PGMPARM=	Specifies the parameter that is passed to the external program specified by the SYSINP= option.
PRINT=	Specifies a destination for SAS output when running in batch mode.
PROCLEAVE=	Specifies how much memory to leave unallocated for SAS procedures to use to complete critical functions during out-of-memory conditions.
REALMEMSIZE=	Specifies the amount of real memory SAS can expect to allocate.
REXXLOC=	Specifies the ddname of the REXX library to be searched when the REXXMAC option is in effect.
REXXMAC	Enables or disables the REXX interface.
SASLIB=	Specifies the ddname for an alternate load library.
SASSCRIPT	Specifies one or more storage locations of SAS/CONNECT script files.
SEQENGINE=	Specifies the default engine for sequential SAS libraries.
SET=	Defines an environment variable.
SORT=	Specifies the minimum size of all allocated sort work data sets.
SORTALTMSGF	Enables sorting with alternate message flags.
SORTBLKMODE	Enables block mode sorting.
SORTBUFMOD	Enables modification of the sort utility output buffer.

System Option	Description
SORTCUTP=	Specifies the number of bytes that SAS sorts. If the number of observations in the data set is greater that the specified number, the host sort program sorts the remaining observations.
SORTDEV=	Specifies the unit device name if SAS dynamically allocates the sort work file.
SORTDEVWARN	Enables device type warnings.
SORTEQOP	Specifies whether the host sort utility supports the EQUALS option.
SORTLIB=	Specifies the name of the sort library.
SORTLIST	Enables passing of the LIST parameter to the host sort utility.
SORTMSG	Controls the class of messages to be written by the host sort utility.
SORTMSG=	Specifies the ddname to be dynamically allocated for the message print file of the host sort utility.
SORTNAME=	Specifies the name of the host sort utility.
SORTOPTS	Specifies whether the host sort utility supports the OPTIONS statement.
SORTPARM=	Specifies parameters for the host sort utility.
SORTPGM=	Specifies whether SAS sorts using the SAS sort utility or the host sort utility.
SORTSHRB	Specifies whether the host sort interface can modify data in buffers.
SORTSUMF	Specifies whether the host sort utility supports the SUM FIELDS=NONE control statement.
SORTUADCON	Specifies whether the host sort utility supports passing a user address constant to the E15/E35 exits.
SORTUNIT=	Specifies the unit of allocation for sort work files.
SORTWKDD=	Specifies the prefix of sort work data sets.
SORTWKNO=	Specifies how many sort work data sets to allocate.
SORT31PL	Controls what type of parameter list is used to invoke the host sort utility.
STAE	Enables or disables a system abend exit.
STATS	Specifies whether statistics are to be written to the SAS log.
STAX	Specifies whether to enable attention handling.
STIMER	Specifies whether to write a subset of system performance statistics to the SAS log.
SVC11SCREEN	Specifies whether to enable SVC 11 screening to obtain host date and time information.
SYNCHIO	Specifies whether synchronous I/O is enabled.
SYSIN=	Specifies the location of the primary SAS input data stream.
SYSINP=	Specifies the name of an external program that provides SAS input statements.
SYSLEAVE=	Specifies how much memory to leave unallocated to ensure that SAS software tasks are able to terminate successfully.

System Option	Description
SYSPREF=	Specifies a prefix for partially qualified physical filenames.
SYSPRINT=	Specifies the handling of output that is directed to the default print file.
S99NOMIG	Tells SAS whether to recall a migrated data set.
TAPECLOSE=	Specifies the default CLOSE setting for a SAS library that is on tape.
USER=	Specifies the location of the default SAS library.
V6GUIMODE	Specifies whether SAS uses Version 6 SCL selection list windows.
VERBOSE	Specifies whether SAS writes the system option settings to the SAS log or to the batch log.
WTOUSERDESC=	Specifies a WTO DATA step function descriptor code.
WTOUSERMCSF=	Specifies WTO DATA step function MCS flags.
WTOUSERROUT=	Specifies a WTO DATA step function routing code.
XCMD	Specifies whether the X command is valid in the SAS session.

SAS Data Quality Server: Reference

System Option	Description
DQLOCALE=	Specifies an ordered list of locales.
DQOPTIONS	Specifies SAS session parameters for data quality programs.
DQSETUPLOC=	Specifies the location of the SAS Data Quality Server setup file.

SAS Intelligence Platform: Application Server Administration Guide

System Option	Description
OBJECTSERVER	Specifies whether SAS is to run as an Integrated Object Model (IOM) server.
OBJECTSERVERPARMS	Specifies startup parameters for the SAS object servers.
SECPACKAGE	Identifies the security package that the IOM server uses to authenticate incoming client connections.
SECPACKAGELIST	Specifies the security authorization packages used by the server.
SSPI	Identifies support for the Security Provider Interface for SSO connections to IOM servers.

For more information, see the SAS Intelligence Platform documentation on http://support.sas.com.

SAS Language Interfaces to Metadata

System Option	Description
METAAUTORESOURCES=	Identifies the metadata resources that are assigned when SAS starts.
METACONNECT=	Identifies the named connection from the metadata user profiles to use as the default values for logging in to the SAS Metadata Server.
METAENCRYPTALG=	Specifies the type of encryption to use when communicating with a SAS Metadata Server.
METAENCRYPTLEVEL=	Specifies what is to be encrypted when communicating with a SAS Metadata Server.
METAPASS=	Specifies the default password for the SAS Metadata Server.
METAPORT=	Specifies the TCP port for the SAS Metadata Server.
METAPROFILE=	Identifies the file that contains the SAS Metadata Server user profiles.
METAPROTOCOL=	Specifies the network protocol for communicating with the SAS Metadata Server.
METAREPOSITORY=	Specifies the default SAS Metadata Repository to use with the SAS Metadata Server.
METASERVER=	Specifies the address of the SAS Metadata Server.
METASPN=	Specifies the service principal name (SPN) for the SAS Metadata Server.
METAUSER=	Specifies the default user ID for logging on to the SAS Metadata Server.

SAS Logging: Configuration and Programming Reference

System Option	Description
LOGAPPLNAME	Specifies a SAS session name for SAS logging.
LOGCONFIGLOC	Specifies the name of the configuration file that is used to initialize SAS logging.

SAS Macro Language: Reference

System Option	Description
CMDMAC	Controls command-style macro invocation.
IMPLMAC	Controls statement-style macro invocation.

System Option	Description
MACRO	Controls whether the SAS macro language is available.
MAUTOLCDISPLAY	Specifies whether to display the source location of the autocall macros in the log when the autocall macro is invoked.
MAUTOSOURCE	Specifies whether the autocall feature is available.
MCOMPILENOTE	Issues a NOTE to the SAS log containing the size and number of instructions upon the completion of the compilation of a macro.
MCOMPILE	Specifies whether to allow new definitions of macros.
MERROR	Specifies whether the macro processor issues a warning message when a macro reference cannot be resolved.
MEXECNOTE	Specifies whether to display macro execution information in the SAS log at macro invocation.
MEXECSIZE	Specifies the maximum macro size that can be executed in memory.
MFILE	Specifies whether MPRINT output is routed to an external file.
MINDELIMITER=	Specifies the character to be used as the delimiter for the macro IN operator.
MINOPERATOR	Specifies whether the macro processor recognizes and evaluates the IN (#) logical operator.
MLOGIC	Specifies whether the macro processor traces its execution for debugging.
MLOGICNEST	Specifies whether to display the macro nesting information in the MLOGIC output in the SAS log.
MPRINT	Specifies whether SAS statements generated by macro execution are traced for debugging.
MPRINTNEST	Specifies whether to display the macro nesting information in the MPRINT output in the SAS log.
MRECALL	Specifies whether autocall libraries are searched for a member that was not found during an earlier search.
MREPLACE	Specifies whether to enable existing macros to be redefined.
MSTORED	Specifies whether the macro facility searches a specific catalog for a stored compiled macro.
MSYMTABMAX	Specifies the maximum amount of memory available to the macro variable symbol tables.
MVARSIZE	Specifies the maximum size for macro variable values that are stored in memory.
SASAUTOS	Specifies the location of one or more autocall libraries.
SASMSTORE=	Identifies the libref of a SAS library with a catalog that contains, or will contain, stored compiled SAS macros.
SERROR	Specifies whether the macro processor issues a warning message when a macro variable reference does not match a macro variable.

System Option	Description
SYMBOLGEN	Specifies whether the results of resolving macro variable references are written to the SAS log for debugging.
SYSPARM	Specifies a character string that can be passed to SAS programs.

SAS National Language Support (NLS): Reference Guide

System Option	Description
BOMFILE	Specifies whether to write the Byte Order Mark (BOM) prefix on Unicode encoded external files.
DATESTYLE	Identifies the sequence of month, date, and year when the ANYDTDTM, ANYDTDTE, or ANYDTTME informats encounter input where the year, month, and day determination is ambiguous.
DBCS	Recognizes double-byte character sets.
DBCSLANG	Specifies a double-byte character set (DBCS) language.
DBCSTYPE	Specifies the encoding method to use for a double-byte character set (DBCS).
DFLANG	Specifies the language for international date informats and formats.
ENCODING	Specifies the default character-set encoding for the SAS session.
FSDBTYPE	Specifies a full-screen double-byte character set (DBCS) encoding method.
FSIMM	Specifies input method modules (IMMs) for a full-screen double-byte character set (DBCS).
FSIMMOPT	Specifies options for input method modules (IMMs) that are used with a full-screen double-byte character set (DBCS).
LOCALE	Specifies a set of attributes in a SAS session that reflect the language, local conventions, and culture for a geographical region.
LOCALELANGCHG	Determines whether the language of the ODS output text can be changed.
NLSCOMPATMODE	Provides national language compatibility with a previous release of SAS.
RSASIOTRANSERROR	Displays a transcoding error when illegal data is read from a remote application.
SORTSEQ	Specifies a language-specific collating sequence for the SORT procedure to use in the current SAS session.
TRANTAB	Specifies the translation tables that are used by various parts of SAS.

SAS Scalable Performance Data Engine: Reference

System Option	Description
COMPRESS=	Specifies to compress the SPD Engine data sets on disk as they are being created.
MAXSEGRATIO=	When evaluating a WHERE expression that contains indexed variables, controls what percentage of index segments to identify as candidate segments before processing the WHERE expression.
MINPARTSIZE=	Specifies a minimum partition size to use for creating SPE Engine data sets.
SPDEINDEXSORTSIZE=	Specifies the size of memory space that the sorting utility can use when sorting values for creating an index.
SPDEMAXTHREADS=	Specifies the upper limit on the number of threads that the SPD Engine can spawn for I/O processing.
SPDESORTSIZE=	Specifies the size of memory space needed for sorting operations used by the SPD Engine.
SPDEUTILLOC=	Specifies one or more file system locations in which the SPD Engine can temporarily store utility files.
SPDEWHEVAL=	Specifies the process used to determine which observations meet the conditions of a WHERE expression.

SAS VSAM Processing for Z/OS

System Option	Description
VSAMLOAD	Enables you to load a VSAM data set.
VSAMREAD	Enables the user to read a VSAM data set.
VSAMRLS	Enables record-level sharing for a VSAM data set.
VSAMUPDATE	Enables you to update a VSAM data set.

SAS/ACCESS for Relational Databases: Reference

System Option	Description
DBIDIRECTEXEC=	Controls SQL optimization for SAS/ACCESS engines.
DBSRVTP=	Specifies whether SAS/ACCESS engines put a hold (or block) on the originating client while making performance-critical calls to the database. This option applies when SAS is invoked as a server responding to multiple clients .
DBSLICEPARM=	Controls the scope of DBMS threaded reads and the number of threads.
SASTRACE=	Generates trace information from a DBMS engine.
SASTRACELOC=	Prints SASTRACE information to a specified location.
SQLMAPPUTTO=	Specifies whether the PUT function in the SQL procedure is processed by SAS or by the SAS_PUT( ) function inside the Teradata database.
VALIDVARNAME=	Controls the type of SAS variable names that can be used or created during a SAS session.

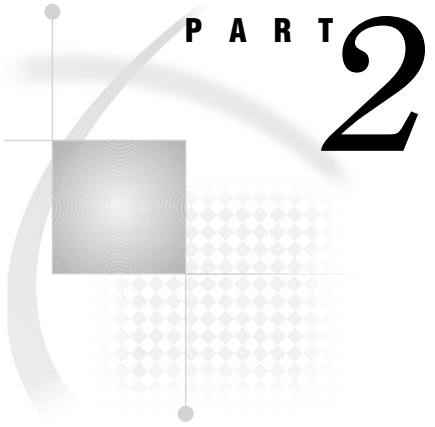
SAS/CONNECT User's Guide

System Option	Description
AUTOSIGNON	Automatically signs on to the server when the client issues a remote submit request for server processing.
COMAMID=	Identifies the communication access method for connecting a client and a server across a network.
CONNECTPERSIST	Specifies whether a connection between a client and a server persists (continues) after the RSUBMIT has completed.
CONNECTREMOTE=	Identifies the server session that a SAS/CONNECT client connects to.
CONNECTSTATUS	Specifies the default setting for the display of the Transfer Status window.
CONNECTWAIT	Specifies whether remote submits are executed synchronously or asynchronously.
DMR	Specifies to invoke a server session.
SASCMD=	Specifies the command that starts a server session on a multi-processor (SMP) machine.
SASFRSCR	Is a read-only option that contains the fileref that is generated by the SASSCRIPT= option.
SASSCRIPT=	Specifies one or more storage locations for SAS/CONNECT script files.
SIGNONWAIT	Specifies whether a SAS/CONNECT SIGNON should be executed asynchronously or synchronously.

System Option	Description
SYSRPUTSYNC	Sets %SYSRPUT macro variable in the client session when the %SYSRPUT statements are executed rather than when a synchronization point in encountered.
TBUFSIZE=	Specifies the size of the buffer that is used by the SAS application layer for transferring data between a client and a server across a network.
TCPPORTFIRST=	Specifies the first value in a range of TCP/IP ports for a client to use to connect to a server.
TCPPORTLAST=	Specifies the last value in a range of TCP/IP ports for a client to use to connect to a server.

SAS/SHARE User's Guide

System Option	Description
COMAMID=	Identifies the communications access method to connect a SAS/SHARE client a server SAS session.
COMAUX1=	Specifies the first alternate communications access method.
SHARESESSIONCNTL=	Specifies the condition under which subsequent sessions can be created on a SAS/SHARE server.
TBUFSIZE=	Specifies the value of the default buffer size that the server uses for transferring data.



PART 2

Dictionary of Component Object Language Elements

CHAPTER

8

Component Objects

DATA Step Component Objects

SAS provides these five predefined component objects for use in a DATA step:

hash and hash iterator objects	enable you to quickly and efficiently store, search, and retrieve data based on lookup keys. For more information about the hash object, see "Using the Hash Object" in *SAS Language Reference: Concepts*. For more information about the hash iterator, see "Using the Hash Iterator Object" in *SAS Language Reference: Concepts*.
java object	provides a mechanism similar to the Java Native Interface (JNI) for instantiating Java classes and accessing fields and methods on the resultant objects. For more information about the java object, see "Using the Java Object" in *SAS Language Reference: Concepts*.
logger and appender objects	enable you to record logging events and write these events to the appropriate destination. For complete information about the logger and appender objects, see "Logger and Appender Object Language Reference" in the *SAS Logging: Configuration and Programming Reference*.

The DATA Step Component Interface

The DATA step component object interface enables you to create and manipulate predefined component objects in a DATA step.

To declare and create a component object, you use either the DECLARE statement by itself or the DECLARE statement and _NEW_ operator together.

Component objects are data elements that consist of attributes, methods, and operators. *Attributes* are the properties that specify the information that is associated with an object. *Methods* define the operations that an object can perform. For component objects, *operators* provide special functionality.

You use the DATA step object dot notation to access the component object's attributes and methods.

Note: The DATA step component object statements, attributes, methods, and operators are limited to those defined for these objects. You cannot use the SAS Component Language functionality with these predefined DATA step objects. △

Dot Notation and DATA Step Component Objects

Definition

Dot notation provides a shortcut for invoking methods and for setting and querying attribute values. Using dot notation makes your SAS programs easier to read.

To use dot notation with a DATA step component object, you must declare and instantiate the component object by using either the DECLARE statement by itself or the DECLARE statement and the _NEW_ operator together. For more information about creating the hash, hash iterator, and java DATA step component objects, see "Using DATA Step Component Objects" in *SAS Language Reference: Concepts*. For more information about creating the logger and appender objects, see "Logger and Appender Object Language Reference" in *SAS Logging: Configuration and Programming Reference*.

Syntax

The syntax for dot notation is as follows:

```
object.attribute
```

or

```
object.method(<argument_tag-1: value-1<, ...argument_tag-n: value-n>>);
```

Where

object
specifies the variable name for the DATA step component object

attribute
specifies an object attribute to assign or query.

When you set an attribute for an object, the code takes this form:

```
object.attribute = value;
```

When you query an object attribute, the code takes this form:

```
value = object.attribute;
```

method
specifies the name of the method to invoke.

argument_tag
identifies the arguments that are passed to the method. Enclose the argument tag in parentheses. The parentheses are required whether the method contains argument tags.

All DATA step component object methods take this form:

```
return_code=object.method(<argument_tag-1:value-1<, ...argument_tag-n value-n>>);
```

The return code indicates method success or failure. A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, an appropriate error message will be printed to the log.

value
specifies the argument value.

Rules When Using Component Objects

- You can assign objects in the same manner as you assign DATA step variables. However, the object types must match. The first set of code is valid, but the second generates an error.

```
declare hash h();
declare hash t();
t=h;
```

```
declare hash t();
declare javaobj j();
j=t;
```

- You cannot declare arrays of objects. The following code would generate an error:

```
declare hash h1();
declare hash h2();
array h h1--h2;
```

- You can store a component object in a hash object as data but not as keys.

```
data _null_;
   declare hash h1();
   declare hash h2();

   length key1 key2 $20;

   h1.defineKey('key1');
   h1.defineData('key1', 'h2');
   h1.defineDone();

   key1 = 'abc';
   h2 = _new_ hash();
   h2.defineKey('key2');
   h2.defineDone();

   key2 = 'xyz';
   h2.add();
   h1.add();

   key1 = 'def';
   h2 = _new_ hash();
   h2.defineKey('key2');
   h2.defineDone();

   key1 = 'abc';
```

```
   rc = h1.find();
   h2.output(dataset: 'work.h2');
run;

proc print data=work.h2;
run;
```

The data set WORK.H2 is displayed.

Output 8.1

Obs	key2
1	xyz

- You cannot use component objects with comparison operators other than the equal sign (=). If H1 and H2 are hash objects, the following code will generate an error:

```
if h1>h2 then
```

- After you declare and instantiate a component object, you cannot assign a scalar value to it. If J is a java object, the following code will generate an error:

```
j=5;
```

- You have to be careful to not delete object references that might still be in use or that have already been deleted by reference. In the following code, the second DELETE statement will generate an error because the original H1 object has already been deleted through the reference to H2. The original H2 can no longer be referenced directly.

```
declare hash h1();
declare hash h2();
declare hash t();
t=h2;
h2=h1;
h2.delete();
t.delete();
```

- You cannot use component objects in argument tag syntax. In the following example, using the H2 hash object in the ADD methods will generate an error.

```
declare hash h2();
declare hash h();
h.add(key: h2);
h.add(key: 99, data: h2);
```

CHAPTER 9

Hash and Hash Iterator Object Language Elements

ADD Method

Adds the specified data that is associated with the given key to the hash object.

Applies to: Hash object

Syntax

rc=*object*.**ADD**(<KEY: *keyvalue-1*,..., KEY: *keyvalue-n*, DATA: *datavalue-1*, ..., DATA: *datavalue-n*>);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

KEY: *keyvalue*
specifies the key value whose type must match the corresponding key variable that is specified in a DEFINEKEY method call.

The number of "KEY: *keyvalue*" pairs depends on the number of key variables that you define by using the DEFINEKEY method.

DATA: *datavalue*
specifies the data value whose type must match the corresponding data variable that is specified in a DEFINEDATA method call.

The number of "DATA: *datavalue*" pairs depends on the number of data variables that you define by using the DEFINEDATA method.

Details

You can use the ADD method in one of two ways to store data in a hash object.

You can define the key and data item, and then use the ADD method as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   /* Declare hash object and key and data variable names */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();
   end;

   /* Define constant key and data values */
   k = 'Joyce';
   d = 'Ulysses';
   /* Add key and data values to hash object */
   rc = h.add();
run;
```

Alternatively, you can use a shortcut and specify the key and data directly in the ADD method call as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   /* Define hash object and key and data variable names */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;

   /* Define constant key and data values and add to hash object */
   rc = h.add(key: 'Joyce', data: 'Ulysses');
run;
```

If you add a key that is already in the hash object, then the ADD method will return a non-zero value to indicate that the key is already in the hash object. Use the REPLACE method to replace the data that is associated with the specified key with new data.

If you do not specify the data variables with the DEFINEDATA method, the data variables are automatically assumed to be same as the keys.

If you use the KEY: and DATA: argument tags to specify the key and data directly, you must use both argument tags.

The ADD method does not set the value of the data variable to the value of the data item. It only sets the value in the hash object.

See Also

Statements:

"DEFINEDATA Method" on page 2092
"DEFINEKEY Method" on page 2096
"REF Method" on page 2124

"Storing and Retrieving Data" in *SAS Language Reference: Concepts*

CHECK Method

Checks whether the specified key is stored in the hash object.

Applies to: Hash object

Syntax

rc=object.**CHECK**(<KEY: *keyvalue-1*,..., KEY: *keyvalue-n*>);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

KEY: *keyvalue*
specifies the key value whose type must match the corresponding key variable that is specified in a DEFINEKEY method call.

The number of "KEY: *keyvalue*" pairs depends on the number of key variables that you define by using the DEFINEKEY method.

Details

You can use the CHECK method in one of two ways to find data in a hash object.

You can specify the key, and then use the CHECK method as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   /* Declare hash object and key and data variable names */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
```

```
      rc = h.defineData('d');
      rc = h.defineDone();

  /* avoid uninitialized variable notes */
      call missing(k, d);
   end;

   /* Define constant key and data values and add to hash object */
   rc = h.add(key: 'Joyce', data: 'Ulysses');

   /* Verify that JOYCE key is in hash object */
   k = 'Joyce';
   rc = h.check();
   if (rc = 0) then
      put 'Key is in the hash object.';
run;
```

Alternatively, you can use a shortcut and specify the key directly in the CHECK method call as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   /* Declare hash object and key and data variable names */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();

  /* avoid uninitialized variable notes */
      call missing(k, d);
   end;

   /* Define constant key and data values and add to hash object */
   rc = h.add(key: 'Joyce', data: 'Ulysses');

   /* Verify that JOYCE key is in hash object */
   rc = h.check(key: 'Joyce');
   if (rc  =0) then
      put 'Key is in the hash object.';
run;
```

Comparisons

The CHECK method only returns a value that indicates whether the key is in the hash object. The data variable that is associated with the key is not updated. The FIND method also returns a value that indicates whether the key is in the hash object. However, if the key is in the hash object, then the FIND method also sets the data variable to the value of the data item so that it is available for use after the method call.

See Also

Methods:

"FIND Method" on page 2100

"DEFINEKEY Method" on page 2096

CLEAR Method

Removes all items from the hash object without deleting the hash object instance.

Applies to: Hash object

Syntax

rc=*object*.**CLEAR**();

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

Details

The CLEAR method enables you to remove items from and reuse an existing hash object without having to delete the object and create a new one. If you want to remove the hash object instance completely, use the DELETE method.

Note: The CLEAR method does not change the value of the DATA step variables. It only clears the values in the hash object. △

Examples

The following example declares a hash object, gets the number of items in the hash object, and then clears the hash object without deleting it.

```
data mydata;
 do i = 1 to 10000;
   output;
 end;
run;

data _null_;
 length i 8;

/* Declares the hash object named MYHASH using the data set MyData. */
dcl hash myhash(dataset: 'mydata');
 myhash.definekey('i');
 myhash.definedone();
 call missing (i);
```

```
/* Uses the NUM_ITEMS attribute, which returns the number of items in
    the hash object. */
 n = myhash.num_items;
 put n=;

/* Uses the CLEAR method to delete all items within MYHASH. */
 rc = myhash.clear();

/* Writes the number of items in the log. */
 n = myhash.num_items;
 put n=;

 run;
```

The first PUT statement writes the number of items in the hash table MYHASH before it is cleared.

```
n=10000
```

The second PUT statement writes the number of items in the hash table MYHASH after it is cleared.

```
n=0
```

See Also

Methods:
"DELETE Method" on page 2097

DECLARE Statement, Hash and Hash Iterator Objects

Declares a hash or hash iterator object; creates an instance of and initializes data for a hash or hash iterator object.

Valid in: DATA step
See: "DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

DEFINEDATA Method

Defines data, associated with the specified data variables, to be stored in the hash object.

Applies to: Hash object

Syntax

rc=object.**DEFINEDATA**('*datavarname-1*'<,...'*datavarname-n*'>);

rc=object.**DEFINEDATA**(ALL: 'YES' | "YES");

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

'datavarname'
specifies the name of the data variable.

The data variable name can also be enclosed in double quotation marks.

ALL: 'YES' | "YES"
specifies all the data variables as data when the data set is loaded in the object constructor.

If the *dataset* argument tag is used in the DECLARE statement or _NEW_ operator to automatically load a data set, then you can define all the data set variables as data by using the ALL: 'YES' option.

Note: If you use the shortcut notation for the ADD or REPLACE method (for example, **`h.add(key:99, data:'apple', data:'orange')`**) and use the ALL:'YES' option on the DEFINEDATA method, then you must specify the data in the same order as it exists in the data set. △

Note: The hash object does not assign values to key variables (for example, **`h.find(key:'abc')`**), and the SAS compiler cannot detect the key and data variable assignments that are performed by the hash object and the hash iterator. Therefore, if no assignment to a key or data variable appears in the program, then SAS will issue a note stating that the variable is uninitialized. To avoid receiving these notes, you can perform one of the following actions:

- Set the NONOTES system option.
- Provide an initial assignment statement (typically to a missing value) for each key and data variable.
- Use the CALL MISSING routine with all the key and data variables as parameters. Here is an example:

```
length d $20;
length k $20;

if _N_ = 1 then do;
   declare hash h();
   rc = h.defineKey('k');
   rc = h.defineData('d');
   rc = h.defineDone();
   call missing(k, d);
end;
```

△

Details

The hash object works by storing and retrieving data based on lookup keys. The keys and data are DATA step variables, which you use to initialize the hash object by using dot notation method calls. You define a key by passing the key variable name to the DEFINEKEY method. You define data by passing the data variable name to the DEFINEDATA method. When you have defined all key and data variables, you must call the DEFINEDONE method to complete initialization of the hash object. Keys and data consist of any number of character or numeric DATA step variables.

For detailed information about how to use the DEFINEDATA method, see "Defining Keys and Data" in *SAS Language Reference: Concepts*.

Examples

The following example creates a hash object and defines the key and data variables:

```
data _null_;
   length d $20;
   length k $20;
   /* Declare the hash object and key and data variables */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;
run;
```

See Also

Methods:

"DEFINEDONE Method" on page 2095

"DEFINEKEY Method" on page 2096

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Defining Keys and Data" in *SAS Language Reference: Concepts*

DEFINEDONE Method

Indicates that all key and data definitions are complete.

Applies to: Hash object

Syntax

rc = *object*.**DEFINEDONE**( );

rc = *object*.**DEFINEDONE**(MEMRC: 'y');

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

memrc:'y'
enables recovery from memory failure when loading a data set into a hash object.

If a call fails because of insufficient memory to load a data set, a nonzero return code is returned. The hash object frees the principal memory in the underlying array. The only allowable operation after this type of failure is deletion via the DELETE method.

Details

When the DEFINEDONE method is called and the *dataset* argument tag is used with the constructor, the data set is loaded into the hash object.

The hash object works by storing and retrieving data based on lookup keys. The keys and data are DATA step variables, which you use to initialize the hash object by using dot notation method calls. You define a key by passing the key variable name to the DEFINEKEY method. You define data by passing the data variable name to the DEFINEDATA method. When you have defined all key and data variables, you must call the DEFINEDONE method to complete initialization of the hash object. Keys and data consist of any number of character or numeric DATA step variables.

For detailed information about how to use the DEFINEDONE method, see "Defining Keys and Data" in *SAS Language Reference: Concepts*.

See Also

Methods:

"DEFINEDATA Method" on page 2092

"DEFINEKEY Method" on page 2096

"Defining Keys and Data" in *SAS Language Reference: Concepts*.

DEFINEKEY Method

Defines key variables for the hash object.

Applies to: Hash object

Syntax

rc=object.**DEFINEKEY**(*'keyvarname-1'<…, 'keyvarname-n'>*);

rc=object.**DEFINEKEY**(ALL: 'YES' | "YES");

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

`object`
specifies the name of the hash object.

`'keyvarname'`
specifies the name of the key variable.

The key variable name can also be enclosed in double quotation marks.

ALL: 'YES' | "YES"
specifies all the data variables as keys when the data set is loaded in the object constructor.

If you use the *dataset* argument tag in the DECLARE statement or _NEW_ operator to automatically load a data set, then you can define all the key variables by using the ALL: 'YES' option.

Note: If you use the shortcut notation for the ADD, CHECK, FIND, REMOVE, or REPLACE methods (for example, **`h.add(key:99, data:'apple', data:'orange')`**) and the ALL:'YES' option on the DEFINEKEY method, then you must specify the keys and data in the same order as they exist in the data set. △

Details

The hash object works by storing and retrieving data based on lookup keys. The keys and data are DATA step variables, which you use to initialize the hash object by using dot notation method calls. You define a key by passing the key variable name to the DEFINEKEY method. You define data by passing the data variable name to the DEFINEDATA method. When you have defined all key and data variables, you must call the DEFINEDONE method to complete initialization of the hash object. Keys and data consist of any number of character or numeric DATA step variables.

For detailed information about how to use the DEFINEKEY method, see "Defining Keys and Data" in *SAS Language Reference: Concepts*.

Note: The hash object does not assign values to key variables (for example, **h.find(key:'abc')**), and the SAS compiler cannot detect the key and data variable assignments done by the hash object and the hash iterator. Therefore, if no assignment to a key or data variable appears in the program, SAS will issue a note stating that the variable is uninitialized. To avoid receiving these notes, you can perform one of the following actions:

- Set the NONOTES system option.
- Provide an initial assignment statement (typically to a missing value) for each key and data variable.
- Use the CALL MISSING routine with all the key and data variables as parameters. Here is an example:

```
length d $20;
length k $20;

if _N_ = 1 then do;
   declare hash h();
   rc = h.defineKey('k');
   rc = h.defineData('d');
   rc = h.defineDone();
   call missing(k, d);
end;
```

△

See Also

Methods:

"DEFINEDATA Method" on page 2092

"DEFINEDONE Method" on page 2095

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Defining Keys and Data" in *SAS Language Reference: Concepts*.

DELETE Method

Deletes the hash or hash iterator object.

Applies to: Hash object

Hash interator object

Syntax

rc=object.**DELETE**( );

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is printed to the log.

object
specifies the name of the hash or hash iterator object.

Details

DATA step component objects are deleted automatically at the end of the DATA step. If you want to reuse the object reference variable in another hash or hash iterator object constructor, you should delete the hash or hash iterator object by using the DELETE method.

If you attempt to use a hash or hash iterator object after you delete it, you will receive an error in the log.

If you want to delete all the items from within a hash object and save the hash object to use again, use the "CLEAR Method" on page 2091.

EQUALS Method

Determines whether two hash objects are equal.

Applies to: Hash object

Syntax

rc=*object*.**EQUALS**(HASH: '*object*', RESULT: *variable name*);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of a hash object.

HASH:'*object*'
specifies the name of the second hash object that is compared to the first hash object.

RESULT: *variable name*
specifies the name of a numeric variable name to hold the result. If the hash objects are equal, the result variable is 1. Otherwise, the result variable is zero.

Details

The following example compares H1 to H2 hash objects:

```
length eq k 8;
declare hash h1();
h1.defineKey('k');
h1.defineDone();

declare hash h2();
h2.defineKey('k');
h2.defineDone();

rc = h1.equals(hash: 'h2', result: eq);
if eq then
  put 'hash objects equal';
else
  put 'hash objects not equal';
```

The two hash objects are defined as equal when all of the following conditions occur:

- Both hash objects are the same size—that is, the HASHEXP sizes are equal.
- Both hash objects have the same number of items—that is, H1.NUM_ITEMS = H2.NUM_ITEMS.
- Both hash objects have the same key and data structure.
- In an unordered iteration over H1 and H2 hash objects, each successive record from H1 has the same key and data fields as the corresponding record in H2—that is, each record is in the same position in each hash object and each such record is identical to the corresponding record in the other hash object.

Examples

In the following example, the first return call to EQUALS returns a nonzero value and the second return call returns a zero value.

```
data x;
 length k eq 8;
 declare hash h1();
 h1.defineKey('k');
 h1.defineDone();

 declare hash h2();
 h2.defineKey('k');
 h2.defineDone();

 k = 99;
 h1.add();
 h2.add();

 rc = h1.equals(hash: 'h2', result: eq);
 put eq=;

 k = 100;
 h2.replace();
```

```
rc = h1.equals(hash: 'h2', result: eq);
put eq=;

run;
```

FIND Method

Determines whether the specified key is stored in the hash object.

Applies to: Hash object

Syntax

rc=object.**FIND**(<KEY: *keyvalue–1*,..., KEY: *keyvalue-n*>);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

KEY: *keyvalue*
specifies the key value whose type must match the corresponding key variable that is specified in a DEFINEKEY method call.

The number of "KEY: **`keyvalue`**" pairs depends on the number of key variables that you define by using the DEFINEKEY method.

Details

You can use the FIND method in one of two ways to find data in a hash object.

You can specify the key, and then use the FIND method as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   /* Declare hash object and key and data variables */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;
```

```
   /* Define constant key and data values */
   rc = h.add(key: 'Joyce', data: 'Ulysses');

   /* Find the key JOYCE */
   k = 'Joyce';
   rc = h.find();
   if (rc = 0) then
      put 'Key is in the hash object.';
run;
```

Alternatively, you can use a shortcut and specify the key directly in the FIND method call as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   /* Declare hash object and key and data variables */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;

    /* Define constant key and data values */
   rc = h.add(key: 'Joyce', data: 'Ulysses');

  /* Find the key JOYCE */
   rc = h.find(key: 'Joyce');
   if (rc = 0) then
      put 'Key is in the hash object.';
run;
```

If the hash object has multiple data items for each key, use the "FIND_NEXT Method" on page 2102 and the "FIND_PREV Method" on page 2104 in conjunction with the FIND method to traverse a multiple data item list.

Comparisons

The FIND method returns a value that indicates whether the key is in the hash object. If the key is in the hash object, then the FIND method also sets the data variable to the value of the data item so that it is available for use after the method call. The CHECK method only returns a value that indicates whether the key is in the hash object. The data variable is not updated.

Examples

The following example creates a hash object. Two data values are added. The FIND method is used to find a key in the hash object. The data value is returned to the data set variable that is associated with the key.

```
data _null_;
   length k $8;
```

```
   length d $12;
   /* Declare hash object and key and data variable names */
   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      /* avoid uninitialized variable notes */
      call missing(k, d);
      rc = h.defineDone();
   end;
   /* Define constant key and data values and add to hash object */
   rc = h.add(key: 'Joyce', data: 'Ulysses');
   rc = h.add(key: 'Homer', data: 'Odyssey');
   /* Verify that key JOYCE is in hash object and */
   /* return its data value to the data set variable D */
   rc = h.find(key: 'Joyce');
   put d=;
run;
```

d=Ulysses is written to the SAS log.

See Also

Methods:

"CHECK Method" on page 2089
"DEFINEKEY Method" on page 2096
"FIND_NEXT Method" on page 2102
"FIND_PREV Method" on page 2104
"REF Method" on page 2124

"Storing and Retrieving Data" in *SAS Language Reference: Concepts*

FIND_NEXT Method

Sets the current list item to the next item in the current key's multiple item list and sets the data for the corresponding data variables.

Applies to: Hash object

Syntax

rc=object.**FIND_NEXT**( );

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, an appropriate error message is printed to the log.

object
specifies the name of the hash object.

Details

The FIND method determines whether the key exists in the hash object. The HAS_NEXT method determines whether the key has multiple data items associated with it. When you have determined that the key has another data item, that data item can be retrieved by using the FIND_NEXT method, which sets the data variable to the value of the data item so that it is available for use after the method call. Once you are in the data item list, you can use the HAS_NEXT and FIND_NEXT methods to traverse the list.

Examples

This example uses the FIND_NEXT method to iterate through a data set where several keys have multiple data items. If a key has more than one data item, subsequent items are marked **dup**.

```
data dup;
   length key data 8;
   input key data;
   datalines;
   1 10
   2 11
   1 15
   3 20
   2 16
   2 9
   3 100
   5 5
   1 5
   4 6
   5 99
   ;

data _null_;
   dcl hash h(dataset:'dup', multidata: 'y');
   h.definekey('key');
   h.definedata('key', 'data');
   h.definedone();
   /* avoid uninitialized variable notes */
   call missing (key, data);

   do key = 1 to 5;
      rc = h.find();
      if (rc = 0) then do;
         put key= data=;
         rc = h.find_next();
         do while(rc = 0);
            put 'dup ' key= data;
            rc = h.find_next();
```

```
        end;
      end;
    end;
run;
```

The following lines are written to the SAS log.

Output 9.1 Keys with Multiple Data Items

```
key=1 data=10
dup key=1 5
dup key=1 15
key=2 data=11
dup key=2 9
dup key=2 16
key=3 data=20
dup key=3 100
key=4 data=6
key=5 data=5
dup key=5 99
```

See Also

Methods:

"FIND Method" on page 2100
"FIND_PREV Method" on page 2104
"HAS_NEXT Method" on page 2107

"Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

FIND_PREV Method

Sets the current list item to the previous item in the current key's multiple item list and sets the data for the corresponding data variables.

Applies to: Hash object

Syntax

rc=*object*.**FIND_PREV**( );

Arguments

rc

specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, an appropriate error message is printed to the log.

object
specifies the name of the hash object.

Details

The FIND method determines whether the key exists in the hash object. The HAS_PREV method determines whether the key has multiple data items associated with it. When you have determined that the key has a previous data item, that data item can be retrieved by using the FIND_PREV method, which sets the data variable to the value of the data item so that it is available for use after the method call. Once you are in the data item list, you can use the HAS_PREV and FIND_PREV methods in addition to the HAS_NEXT and FIND_NEXT methods to traverse the list. See "HAS_NEXT Method" on page 2107 for an example.

See Also

Methods:
"FIND Method" on page 2100
"FIND_NEXT Method" on page 2102
"Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

FIRST Method

Returns the first value in the underlying hash object.

Applies to: Hash iterator object

Syntax

rc=*object*.**FIRST**( );

Arguments

rc
specifies whether the method succeeded or failed.
A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, an appropriate error message will be printed to the log.

object
specifies the name of the hash iterator object.

Details

The FIRST method returns the first data item in the hash object. If you use the *ordered: 'yes'* or *ordered: 'ascending'* argument tag in the DECLARE statement or _NEW_ operator when you instantiate the hash object, then the data item that is returned is the one with the 'least' key (smallest numeric value or first alphabetic character),

because the data items are sorted in ascending key-value order in the hash object. Repeated calls to the NEXT method will iteratively traverse the hash object and return the data items in ascending key order. Conversely, if you use the *ordered: 'descending'* argument tag in the DECLARE statement or _NEW_ operator when you instantiate the hash object, then the data item that is returned is the one with the 'highest' key (largest numeric value or last alphabetic character), because the data items are sorted in descending key-value order in the hash object. Repeated calls to the NEXT method will iteratively traverse the hash object and return the data items in descending key order.

Use the LAST method to return the last data item in the hash object.

Note: The FIRST method sets the data variable to the value of the data item so that it is available for use after the method call. △

Examples

The following example creates a data set that contains sales data. You want to list products in order of sales. The data is loaded into a hash object and the FIRST and NEXT methods are used to retrieve the data.

```
data work.sales;
   input prod $1-6 qty $9-14;
   datalines;
banana  398487
apple   384223
orange  329559
;

data _null_;
   /* Declare hash object and read SALES data set as ordered */
   if _N_ = 1 then do;
      length prod $10;
      length qty $6;
      declare hash h(dataset: 'work.sales', ordered: 'yes');
      declare hiter iter('h');
      /* Define key and data variables */
      h.defineKey('qty');
      h.defineData('prod');
      h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(qty, prod);
   end;

   /* Iterate through the hash object and output data values */
   rc = iter.first();
   do while (rc = 0);
      put prod=;
      rc = iter.next();
   end;
run;
```

The following lines are written to the SAS log:

```
prod=orange
prod=banana
prod=apple
```

See Also

Method:

"LAST Method" on page 2111

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Using the Hash Iterator Object" in *SAS Language Reference: Concepts*

HAS_NEXT Method

Determines whether there is a next item in the current key's multiple data item list.

Applies to: Hash object

Syntax

rc=*object*.**HAS_NEXT**(RESULT: R);

Arguments

rc

specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object

specifies the name of the hash object.

RESULT:R

specifies the numeric variable **R**, which receives a zero value if there is not another data item in the data item list or a nonzero value if there is another data item in the data item list.

Details

If a key has multiple data items, you can use the HAS_NEXT method to determine whether there is a next item in the current key's multiple data item list. If there is another item, the method will return a nonzero value in the numeric variable **R**. Otherwise, it will return a zero.

The FIND method determines whether the key exists in the hash object. The HAS_NEXT method determines whether the key has multiple data items associated with it. When you have determined that the key has another data item, that data item can be retrieved by using the FIND_NEXT method, which sets the data variable to the value of the data item so that it is available for use after the method call. Once you are

in the data item list, you can use the HAS_PREV and FIND_PREV methods in addition to the HAS_NEXT and FIND_NEXT methods to traverse the list.

Examples

This example creates a hash object where several keys have multiple data items. It uses the HAS_NEXT method to find all the data items.

```
data testdup;
   length key data 8;
   input key data;
   datalines;
   1 100
   2 11
   1 15
   3 20
   2 16
   2 9
   3 100
   5 5
   1 5
   4 6
   5 99
;

data _null_;
   length r 8;
   dcl hash h(dataset:'testdup', multidata: 'y');
   h.definekey('key');
   h.definedata('key', 'data');
   h.definedone();
   call missing (key, data);

   do key = 1 to 5;
      rc = h.find();
      if (rc = 0) then do;
         put key= data=;
         h.has_next(result: r);
         do while(r ne 0);
             rc = h.find_next();
             put 'dup ' key= data;
             h.has_next(result: r);
         end;
      end;
   end;
run;
```

The following lines are written to the SAS log.

Output 9.2 Output of Keys with Multiple Data Items

```
key=1 data=100
dup key=1 5
dup key=1 15
key=2 data=11
dup key=2 9
dup key=2 16
key=3 data=20
dup key=3 100
key=4 data=6
key=5 data=5
dup key=5 99
```

See Also

Methods:

"FIND Method" on page 2100
"FIND_NEXT Method" on page 2102
"FIND_PREV Method" on page 2104
"HAS_PREV Method" on page 2109

"Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

HAS_PREV Method

Determines whether there is a previous item in the current key's multiple data item list.

Applies to: Hash object

Syntax

rc=object.**HAS_PREV**(RESULT: R);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

RESULT:R
specifies the numeric variable **R**, which receives a zero value if there is not another data item in the data item list or a nonzero value if there is another data item in the data item list.

Details

If a key has multiple data items, you can use the HAS_PREV method to determine whether there is a previous item in the current key's multiple data item list. If there is a previous item, the method will return a nonzero value in the numeric variable **R**. Otherwise, it will return a zero.

The FIND method determines whether the key exists in the hash object. The HAS_NEXT method determines whether the key has multiple data items associated with it. When you have determined that the key has a previous data item, that data item can be retrieved by using the FIND_PREV method, which sets the data variable to the value of the data item so that it is available for use after the method call. Once you are in the data item list, you can use the HAS_PREV and FIND_PREV methods in addition to the HAS_NEXT and FIND_NEXT methods to traverse the list. See "HAS_NEXT Method" on page 2107 for an example.

See Also

Methods:

"FIND Method" on page 2100
"FIND_NEXT Method" on page 2102
"FIND_PREV Method" on page 2104
"HAS_NEXT Method" on page 2107

"Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

ITEM_SIZE Attribute

Returns the size (in bytes) of an item in a hash object.

Applies to: Hash object

Syntax

variable_name=*object*.**ITEM_SIZE**;

Arguments

variable_name
specifies name of the variable that contains the size of the item in the hash object.

object
specifies the name of the hash object.

Details

The ITEM_SIZE attribute returns the size (in bytes) of an item, which includes the key and data variables and some additional internal information. You can set an estimate of how much memory the hash object is using with the ITEM_SIZE and NUM_ITEMS attributes. The ITEM_SIZE attribute does not reflect the initial overhead that the hash

object requires, nor does it take into account any necessary internal alignments. Therefore, the use of ITEM_SIZE does not provide exact memory usage, but it does return a good approximation.

Examples

The following example uses ITEM_SIZE to return the size of the item in MYHASH:

```
data work.stock;
   input prod $1-10 qty 12-14;
   datalines;
broccoli 345
corn 389
potato 993
onion 730
;

data _null_;
   if _N_ = 1 then do;
      length prod $10;
   /* Declare hash object and read STOCK data set as ordered */
      declare hash myhash(dataset: "work.stock");
      /* Define key and data variables */
      myhash.defineKey('prod');
      myhash.defineData('qty');
      myhash.defineDone();
   end;
   /* Add a key and data value to the hash object */
   prod = 'celery';
   qty = 183;
   rc = myhash.add();

   /* Use ITEM_SIZE to return the size of the item in hash object */
   itemsize = myhash.item_size;
   put itemsize=;
run;
```

The following lines are written to the log:

```
itemsize=40
```

LAST Method

Returns the last value in the underlying hash object.

Applies to: Hash iterator object

Syntax

rc=object.**LAST**( );

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash iterator object.

Details

The LAST method returns the last data item in the hash object. If you use the *ordered: 'yes'* or *ordered: 'ascending'* argument tag in the DECLARE statement or _NEW_ operator when you instantiate the hash object, then the data item that is returned is the one with the 'highest' key (largest numeric value or last alphabetic character), because the data items are sorted in ascending key-value order in the hash object. Conversely, if you use the *ordered: 'descending'* argument tag in the DECLARE statement or _NEW_ operator when you instantiate the hash object, then the data item that is returned is the one with the 'least' key (smallest numeric value or first alphabetic character), because the data items are sorted in descending key-value order in the hash object.

Use the FIRST method to return the first data item in the hash object.

Note: The LAST method sets the data variable to the value of the data item so that it is available for use after the method call. △

See Also

Methods:

"FIRST Method" on page 2105

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Using the Hash Iterator Object" in *SAS Language Reference: Concepts*

NEW Operator, Hash or Hash Iterator Object

Creates an instance of a hash or hash iterator object.

Applies to: Hash object
Hash iterator object

Syntax

object-reference = **_NEW_** *object*(<*argument_tag-1*: *value-1*<, *...argument_tag-n*: *value-n*>>);

Arguments

object-reference
specifies the object reference name for the hash or hash iterator object.

object
specifies the component object. It can be one of the following:

hash — indicates a hash object. The hash object provides a mechanism for quick data storage and retrieval. The hash object stores and retrieves data based on lookup keys. For more information about the hash object, see "Using the Hash Object" in *SAS Language Reference: Concepts*.

hiter — indicates a hash iterator object. The hash iterator object enables you to retrieve the hash object's data in forward or reverse key order. For more information about the hash iterator object, see "Using the Hash Iterator Object" in *SAS Language Reference: Concepts*.

argument-tag
specifies the information that is used to create an instance of the hash object.

Valid hash object argument tags are

dataset: '*dataset_name* <(*datasetoption*)>'
Names a SAS data set to load into the hash object.

The name of the SAS data set can be a literal or character variable. The data set name must be enclosed in single or double quotation marks. Macro variables must be enclosed in double quotation marks.

You can use SAS data set options when declaring a hash object in the DATASET argument tag. Data set options specify actions that apply only to the SAS data set with which they appear. They enable you to perform the following operations:

- renaming variables
- selecting a subset of observations based on observation number for processing
- selecting observations using the WHERE option
- dropping or keeping variables from a data set loaded into a hash object, or for an output data set specified in an OUTPUT method call
- specifying a password for a data set.

The following syntax is used:

```
dcl hash h;
h = _new_ hash (dataset: 'x (where = (i > 10))');
```

For a list of SAS data set options, see "Data Set Options by Category" on page 12.

Note: If the data set contains duplicate keys, the default is to keep the first instance in the hash object; subsequent instances will be ignored. To store the last instance in the hash object or have an error message written in the SAS log if there is a duplicate key, use the DUPLICATE argument tag. △

duplicate: '*option*'
determines whether to ignore duplicate keys when loading a data set into the hash object. The default is to store the first key and ignore all subsequent duplicates. Option can be one of the following values:

'replace' | 'r'
stores the last duplicate key record.

'error' | 'e'
reports an error to the log if a duplicate key is found.

The following example using the REPLACE option stores **brown** for the key 620 and **blue** for the key 531 . If you use the default, **green** would be stored for 620 and **yellow** would be stored for 531.

```
data table;
  input key data $;
  datalines;
  531 yellow
  620 green
  531 blue
  908 orange
  620 brown
  143 purple
 run;

data _null_;
length key 8 data $ 8;
if (_n_ = 1) then do;
    declare hash myhash;
    myhash = _new_ hash (dataset: "table", duplicate: "r");
    rc = myhash.definekey('key');
    rc = myhash.definedata('data');
    myhash.definedone();
 end;

rc = myhash.output(dataset:"otable");
run;
```

hashexp: *n*
The hash object's internal table size, where the size of the hash table is 2^n.

The value of HASHEXP is used as a power-of-two exponent to create the hash table size. For example, a value of 4 for HASHEXP equates to a hash table size of 2^4, or 16. The maximum value for HASHEXP is 20.

The hash table size is not equal to the number of items that can be stored. Imagine the hash table as an array of 'buckets.' A hash table size of 16 would have 16 'buckets.' Each bucket can hold an infinite number of items. The efficiency of the hash table lies in the ability of the hashing function to map items to and retrieve items from the buckets.

You should set the hash table size relative to the amount of data in the hash object in order to maximize the efficiency of the hash object lookup routines. Try different HASHEXP values until you get the best result. For example, if the hash object contains one million items, a hash table size of 16 (HASHEXP = 4) would work, but not very efficiently. A hash table size of 512 or 1024 (HASHEXP = 9 or 10) would result in the best performance.

Default: 8, which equates to a hash table size of 2^8 or 256

ordered: '*option*'
Specifies whether or how the data is returned in key-value order if you use the hash object with a hash iterator object or if you use the hash object OUTPUT method.

option can be one of the following values:

'ascending' | 'a' Data is returned in ascending key-value order. Specifying '**ascending**' is the same as specifying '**yes**'.

'descending' | 'd' Data is returned in descending key-value order.

'YES' | 'Y' Data is returned in ascending key-value order. Specifying '**yes**' is the same as specifying '**ascending**'.

'NO' | 'N' Data is returned in some undefined order.

Default: NO

The argument value can also be enclosed in double quotation marks.

multidata: '*option*'
specifies whether multiple data items are allowed for each key.
option can be one of the following values:

'YES' | 'Y' Multiple data items are allowed for each key.

'NO' | 'N' Only one data item is allowed for each key.

Default: NO

See Also: "Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

The argument value can also be enclosed in double quotation marks.

suminc: '*variable-name*'
maintains a summary count of hash object keys. The SUMINC argument tag is given a DATA step variable, which holds the sum increment, that is, how much to add to the key summary for each reference to the key. The SUMINC value treats a missing value as zero, like the SUM function. For example, a key summary changes using the current value of the DATA step variable.

```
dcl hash myhash(suminc: 'count');
```

For more information, see "Maintaining Key Summaries" in *SAS Language Reference: Concepts*.

See Also: "Initializing Hash Object Data Using a Constructor" and "Declaring and Instantiating a Hash Iterator Object" in *SAS Language Reference: Concepts*.

Details

To use a DATA step component object in your SAS program, you must declare and create (instantiate) the object. The DATA step component interface provides a mechanism for accessing the predefined component objects from within the DATA step.

If you use the _NEW_ operator to instantiate the component object, you must first use the DECLARE statement to declare the component object. For example, in the following lines of code, the DECLARE statement tells SAS that the object reference H is a hash object. The _NEW_ operator creates the hash object and assigns it to the object reference H.

```
declare hash h();
h = _new_ hash( );
```

Note: You can use the DECLARE statement to declare and instantiate a hash or hash iterator object in one step. △

A constructor is a method that is used to instantiate a component object and to initialize the component object data. For example, in the following lines of code, the _NEW_ operator instantiates a hash object and assigns it to the object reference H. In addition, the data set WORK.KENNEL is loaded into the hash object.

```
declare hash h();
h = _new_ hash(datset: "work.kennel");
```

For more information about the predefined DATA step component objects and constructors, see “Using DATA Step Component Objects” in *SAS Language Reference: Concepts*.

Comparisons

You can use the DECLARE statement and the _NEW_ operator, or the DECLARE statement alone to declare and instantiate an instance of a hash or hash iterator object.

Examples

This example uses the _NEW_ operator to instantiate and initialize data for a hash object and instantiate a hash iterator object.

The hash object is filled with data, and the iterator is used to retrieve the data in key order.

```
data kennel;
   input name $1-10 kenno $14-15;
   datalines;
Charlie      15
Tanner       07
Jake         04
Murphy       01
Pepe         09
Jacques      11
Princess Z   12
;
run;

data _null_;
   if _N_ = 1 then do;
      length kenno $2;
      length name $10;
      /* Declare the hash object */
      declare hash h();
      /* Instantiate and initialize the hash object */
      h = _new_ hash(dataset:"work.kennel", ordered: 'yes');
      /* Declare the hash iterator object */
      declare hiter iter;
      /* Instantiate the hash iterator object */
      iter = _new_ hiter('h');
      /* Define key and data variables */
      h.defineKey('kenno');
      h.defineData('name', 'kenno');
      h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(kenno, name);
   end;

   /* Find the first key in the ordered hash object and output to the log */
   rc = iter.first();
   do while (rc = 0);
      put kenno '   ' name;
      rc = iter.next();
```

```
   end;
run;
```

The following lines are written to the SAS log:

Output 9.3 Output of Data Written in Key Order

```
NOTE: There were 7 observations read from the data set WORK.KENNEL.
01    Murphy
04    Jake
07    Tanner
09    Pepe
11    Jacques
12    Princess Z
15    Charlie
```

See Also

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Using DATA Step Component Objects" in *SAS Language Reference: Concepts*

NEXT Method

Returns the next value in the underlying hash object.

Applies to: Hash iterator object

Syntax

rc=*object*.**NEXT**( );

Arguments

rc

specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object

specifies the name of the hash iterator object.

Details

Use the NEXT method iteratively to traverse the hash object and return the data items in key order.

The FIRST method returns the first data item in the hash object.

You can use the PREV method to return the previous data item in the hash object.

Note: The NEXT method sets the data variable to the value of the data item so that it is available for use after the method call. △

Note: If you call the NEXT method without calling the FIRST method, then the NEXT method will still start at the first item in the hash object. △

See Also

Methods:

"FIRST Method" on page 2105

"PREV Method" on page 2123

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Using the Hash Iterator Object" in *SAS Language Reference: Concepts*

NUM_ITEMS Attribute

Returns the number of items in the hash object.

Applies to: Hash object

Syntax

variable_name=*object*.**NUM_ITEMS**;

Arguments

variable_name
specifies the name of the variable that contains the number of items in the hash object.

object
specifies the name of the hash object.

Examples

This example creates a data set and loads the data set into a hash object. An item is added to the hash object and the total number of items in the resulting hash object is returned by the NUM_ITEMS attribute.

```
data work.stock;
   input item $1-10 qty $12-14;
   datalines;
broccoli 345
```

```
corn 389
potato 993
onion 730
;

data _null_;
   if _N_ = 1 then do;
      length item $10;
      length qty 8;
      length totalitems 8;
      /* Declare hash object and read STOCK data set as ordered */
      declare hash myhash(dataset: "work.stock");
      /* Define key and data variables */
      myhash.defineKey('item');
      myhash.defineData('qty');
      myhash.defineDone();
   end;
   /* Add a key and data value to the hash object */
   item = 'celery';
   qty = 183;
   rc = myhash.add();
   if (rc ne 0) then
      put 'Add failed';
   /* Use NUM_ITEMS to return updated number of items in hash object */
   totalitems = myhash.num_items;
   put totalitems=;
run;
```

totalitems=5 is written to the SAS log.

OUTPUT Method

Creates one or more data sets each of which contain the data in the hash object.

Applies to: Hash object

Syntax

rc=*object*.**OUTPUT**(DATASET: '*dataset-1* <(*datasetoption*)>' <…, DATASET: '*dataset-n*'>('*datasetoption*<(*datasetoption*)>');

Arguments

rc

specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

DATASET: '*dataset*'
specifies the name of the output data set.

The name of the SAS data set can be a character literal or character variable. The data set name can also be enclosed in double quotation marks. When specifying the name of the output data set, you can use SAS data set options in the DATASET argument tag. Macro variables must be enclosed in double quotation marks.

datasetoption
specifies a data set option.

For complete information about how to specify data set options, see "Syntax" on page 10.

Details

Hash object keys are not automatically stored as part of the output data set. The keys must be defined as data items by using the DEFINEDATA method to be included in the output data set.

If you use the *ordered: 'yes'* or *ordered: 'ascending'* argument tag in the DECLARE statement or _NEW_ operator when you instantiate the hash object, then the data items are written to the data set in ascending key-value order. If you use the *ordered: 'descending'* argument tag in the DECLARE statement or _NEW_ operator when you instantiate the hash object, then the data items are written to the data set in descending key-value order. If you do not use the *ordered* argument tag, the order is undefined.

When specifying the name of the output data set, you can use SAS data set options in the DATASET argument tag. Data set options specify actions that apply only to the SAS data set with which they appear. They let you perform the following operations:

- renaming variables
- selecting a subset of observations based on the observation number for processing
- selecting observations using the WHERE option
- dropping or keeping variables from a data set loaded into a hash object, or for an output data set that is specified in an OUTPUT method call
- specifying a password for a data set.

The following example uses the WHERE data set option to select specific data for the output data set named OUT:

```
data x;
 do i = 1 to 20;
   output;
 end;
run;

/* Using the WHERE option. */
data _null_;
  length i 8;
  dcl hash h();
  h.definekey(all: 'y');
  h.definedone();
  h.output(dataset: 'out (where =( i < 8))');
run;
```

The following example uses the RENAME data set option to rename the variable J to K for the output data set named OUT:

```
data x;
 do i = 1 to 20;
   output;
 end;
run;

/* Using the RENAME option. */
data _null_;
  length i j 8;
  dcl hash h();
  h.definekey(all: 'y');
  h.definedone();
  h.output(dataset: 'out (rename =(j=k))');
run;
```

For a list of data set options, see "Data Set Options by Category" on page 12.

Note: When you use the OUTPUT method to create a data set, the hash object is not part of the output data set. In the following example, the H2 hash object will be omitted from the output data set.

```
data _null_;
   length k 8;
   length d $10;
   declare hash h2();
   declare hash h(ordered: 'y');
   h.defineKey('k');
   h.defineData('k', 'd', 'h2');
   h.defineDone();
   k = 99;
   d = 'abc';
   h.add();
   k = 199;
   d = 'def';
   h.add();
   h.output(dataset:'work.x');
run;
```

△

Examples

Using the data set ASTRO that contains astronomical data, the following code creates a hash object with the Messier (OBJ) objects sorted in ascending order by their right-ascension (RA) values and uses the OUTPUT method to save the data to a data set.

```
data astro;
input obj $1-4 ra $6-12 dec $14-19;
datalines;
 M31 00 42.7 +41 16
 M71 19 53.8 +18 47
 M51 13 29.9 +47 12
 M98 12 13.8 +14 54
 M13 16 41.7 +36 28
 M39 21 32.2 +48 26
 M81 09 55.6 +69 04
```

```
M100 12 22.9 +15 49
 M41 06 46.0 -20 44
 M44 08 40.1 +19 59
 M10 16 57.1 -04 06
 M57 18 53.6 +33 02
  M3 13 42.2 +28 23
 M22 18 36.4 -23 54
 M23 17 56.8 -19 01
 M49 12 29.8 +08 00
 M68 12 39.5 -26 45
 M17 18 20.8 -16 11
 M14 17 37.6 -03 15
 M29 20 23.9 +38 32
 M34 02 42.0 +42 47
 M82 09 55.8 +69 41
 M59 12 42.0 +11 39
 M74 01 36.7 +15 47
 M25 18 31.6 -19 15
;
run;

data _null_;
   if _N_ = 1 then do;
      length obj $10;
      length ra $10;
      length dec $10;
      /* Read ASTRO data set as ordered */
   declare hash h(hashexp: 4, dataset:"work.astro", ordered: 'yes');
      /* Define variables RA and OBJ as key and data for hash object */
      h.defineKey('ra');
      h.defineData('ra', 'obj');
      h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(ra, obj);
end;
/* Create output data set from hash object */
rc = h.output(dataset: 'work.out');
run;

proc print data=work.out;
   var ra obj;
   title 'Messier Objects Sorted by Right-Ascension Values';
run;
```

Output 9.4 Messier Objects Sorted by Right-Ascension Values

Messier Objects Sorted by Right-Ascension Values 1

Obs	ra	obj
1	00 42.7	M31
2	01 36.7	M74
3	02 42.0	M34
4	06 46.0	M41
5	08 40.1	M44
6	09 55.6	M81
7	09 55.8	M82
8	12 13.8	M98
9	12 22.9	M100
10	12 29.8	M49
11	12 39.5	M68
12	12 42.0	M59
13	13 29.9	M51
14	13 42.2	M3
15	16 41.7	M13
16	16 57.1	M10
17	17 37.6	M14
18	17 56.8	M23
19	18 20.8	M17
20	18 31.6	M25
21	18 36.4	M22
22	18 53.6	M57
23	19 53.8	M71
24	20 23.9	M29
25	21 32.2	M39

See Also

Methods:

"DEFINEDATA Method" on page 2092

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Saving Hash Object Data in a Data Set" in *SAS Language Reference: Concepts*

PREV Method

Returns the previous value in the underlying hash object.

Applies to: Hash iterator object

Syntax

rc=object.**PREV**( );

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash iterator object.

Details

Use the PREV method iteratively to traverse the hash object and return the data items in reverse key order.

The FIRST method returns the first data item in the hash object. The LAST method returns the last data item in the hash object.

You can use the NEXT method to return the next data item in the hash object.

Note: The PREV method sets the data variable to the value of the data item so that it is available for use after the method call. △

See Also

Methods:

"FIRST Method" on page 2105

"LAST Method" on page 2111

"NEXT Method" on page 2117

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Using the Hash Iterator Object" in *SAS Language Reference: Concepts*

REF Method

Consolidates the FIND and ADD methods into a single method call.

Applies to: Hash object

Syntax

rc=*object*.**REF**(<KEY: *keyvalue-1*,..., KEY: *keyvalue-n*, DATA: *datavalue-1*, ..., DATA: *datavalue-n*>);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

KEY: ***keyvalue***
specifies the key value whose type must match the corresponding key variable that is specified in a DEFINEKEY method call.

The number of "KEY: *keyvalue*" pairs depends on the number of key variables that you define by using the DEFINEKEY method.

DATA: ***datavalue***
specifies the data value whose type must match the corresponding data variable that is specified in a DEFINEDATA method call.

The number of "DATA: *datavalue*" pairs depends on the number of data variables that you define by using the DEFINEDATA method.

Details

You can consolidate FIND and ADD methods into a single REF method. You can change the following code:

```
rc = h.find();
  if (rc ne = 0) then
    rc = h.add();
```

to

```
rc = h.ref();
```

The REF method is useful for counting the number of occurrences of each key in a hash object. The REF method initializes the key summary for each key on the first ADD, and then changes the ADD for each subsequent FIND.

Note: The REF method sets the data variable to the value of the data item so that it is available for use after the method call. △

For more information about key summaries, see *SAS Language Reference: Concepts*.

Examples

The following example uses the REF method for key summaries:

```
data keys;
 input key;
datalines;
1
2
1
3
5
2
```

```
3
2
4
1
5
1
;

data count;
 length count key 8;
 keep key count;

 if _n_ = 1 then do;
   declare hash myhash(suminc: "count", ordered: "y");
   declare hiter iter("myhash");
   myhash.defineKey('key');
   myhash.defineDone();
   count = 1;
  end;

 do while (not done);
   set keys end=done;
   rc = myhash.ref();
 end;

 rc = iter.first();
 do while(rc = 0);
   rc = myhash.sum(sum: count);
   output;
   rc = iter.next();
 end;

 stop;
run;
```

The following lines are written to the SAS log.

Output 9.5 Output of DATA Using the REF Method

Obs	count	key
1	4	1
2	3	2
3	2	3
4	1	4
5	2	5

See Also

Methods:

"ADD Method" on page 2087
"FIND Method" on page 2100
"CHECK Method" on page 2089

REMOVE Method

Removes the data that is associated with the specified key from the hash object.

Applies to: Hash object

Syntax

rc=*object*.**REMOVE**(<KEY: *keyvalue-1*,..., KEY: *keyvalue-n*>);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

KEY: *keyvalue*
specifies the key value whose type must match the corresponding key variable that is specified in a DEFINEKEY method call

The number of "KEY: *keyvalue*" pairs depends on the number of key variables that you define by using the DEFINEKEY method.

Restriction: If an associated hash iterator is pointing to the *keyvalue*, then the REMOVE method will not remove the key or data from the hash object. An error message is issued.

Details

The REMOVE method deletes both the key and the data from the hash object.

You can use the REMOVE method in one of two ways to remove the key and data in a hash object.

You can specify the key, and then use the REMOVE method as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;

   rc = h.add(key: 'Joyce', data: 'Ulysses');
```

```
   /* Specify the key */
   k = 'Joyce';
   /* Use the REMOVE method to remove the key and data */
   rc = h.remove();
   if (rc = 0) then
      put 'Key and data removed from the hash object.';
run;
```

Alternatively, you can use a shortcut and specify the key directly in the REMOVE method call as shown in the following code:

```
data _null_;
   length k $8;
   length d $12;

   if _N_ = 1 then do;
      declare hash h();
      rc = h.defineKey('k');
      rc = h.defineData('d');
      rc = h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(k, d);
   end;

   rc = h.add(key: 'Joyce', data: 'Ulysses');
   rc = h.add(key: 'Homer', data: 'Iliad');

   /* Specify the key in the REMOVE method parameter */
   rc = h.remove(key: 'Homer');
   if (rc  =0) then
      put 'Key and data removed from the hash object.';
run;
```

Note: The REMOVE method does not modify the value of data variables. It only removes the value in the hash object. △

Note: If you specify **multidata:'y'** in the hash object constructor, the REMOVE method will remove all data items for the specified key. △

Examples

This example illustrates how to remove a key in the hash table.

```
/* Generate test data */
data x;
   do k = 65 to 70;
      d = byte (k);
      output;
   end;
run;

data _null_;
   length k 8 d $1;
   /* define the hash table and iterator */
   declare hash H (dataset:'x', ordered:'a');
   H.defineKey  ('k');
```

```
   H.defineData ('k', 'd');
   H.defineDone ();
   call missing (k,d);
   declare hiter HI ('H');
   /* Use this logic to remove a key in the hash table
   when an iterator is pointing to that key */
   do while (hi.next() = 0);
      if flag then rc=h.remove(key:key);
         if d = 'C' then do;
            key=k;
            flag=1;
         end;
      end;

rc = h.output(dataset: 'work.out');
stop;
run;

proc print;
run;
```

The following output shows that the key and data for the third object (key=67, data=C) is deleted.

Output 9.6 Key and Data Removed from Output

```
                              The SAS System                              1

Obs      k    d

 1      65    A
 2      66    B
 3      68    D
 4      69    E
 5      70    F
```

See Also

Methods:

"ADD Method" on page 2087

"DEFINEKEY Method" on page 2096

"REMOVEDUP Method" on page 2130

"Replacing and Removing Data" in *SAS Language Reference: Concepts*

REMOVEDUP Method

Removes the data that is associated with the specified key's current data item from the hash object.

Applies to: Hash object

Syntax

rc=*object*.**REMOVEDUP**(<KEY: *keyvalue-1*,..., KEY: *keyvalue-n*>);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

KEY: *keyvalue*
specifies the key value whose type must match the corresponding key variable that is specified in a DEFINEKEY method call.

The number of "KEY: *keyvalue*" pairs depends on the number of key variables that you define by using the DEFINEKEY method.

Restriction: If an associated hash iterator is pointing to the **`keyvalue`**, then the REMOVEDUP method will not remove the key or data from the hash object. An error message is issued.

Details

The REMOVEDUP method deletes both the key and the data from the hash object.

You can use the REMOVEDUP method in one of two ways to remove the key and data in a hash object. You can specify the key, and then use the REMOVEDUP method. Alternatively, you can use a shortcut and specify the key directly in the REMOVEDUP method call.

Note: The REMOVEDUP method does not modify the value of data variables. It only removes the value in the hash object. △

Note: If only one data item is in the key's data item list, the key and data will be removed from the hash object. △

Comparisons

The REMOVEDUP method removes the data that is associated with the specified key's current data item from the hash object. The REMOVE method removes the data that is associated with the specified key from the hash object.

Examples

This example creates a hash object where several keys have multiple data items. The last data item in the key is removed.

```
data testdup;
   length key data 8;
   input key data;
   datalines;
   1 10
   2 11
   1 15
   3 20
   2 16
   2 9
   3 100
   5 5
   1 5
   4 6
   5 99
;

data _null_;
   length r 8;
   dcl hash h(dataset:'testdup', multidata: 'y', ordered: 'y');
   h.definekey('key');
   h.definedata('key', 'data');
   h.definedone();
   call missing (key, data);

   do key = 1 to 5;
      rc = h.find();
      if (rc = 0) then do;
         h.has_next(result: r);
         if (r ne 0) then do;
            h.find_next();
            h.removedup();
         end;
      end;
   end;

   dcl hiter i('h');
   rc = i.first();
   do while (rc = 0);
      put key= data=;
      rc = i.next();
   end;
run;
```

The following lines are written to the SAS log.

Output 9.7 Last Data Item Removed from the Key

```
key=1 data=10
key=1 data=15
key=2 data=11
key=2 data=16
key=3 data=20
key=4 data=6
key=5 data=5
```

See Also

Methods:

"REMOVE Method" on page 2127

"Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

REPLACE Method

Replaces the data that is associated with the specified key with new data.

Applies to: Hash object

Syntax

rc=*object*.**REPLACE**(<KEY: *keyvalue-1*,..., KEY: *keyvalue-n*, DATA: *datavalue-1*,..., DATA: *datavalue-n*>);

Arguments

rc

specifies whether the method succeeded or failed.

A return code of zero indicates success; a non-zero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object

specifies the name of the hash object.

KEY: ***keyvalue***

specifies the key value whose type must match the corresponding key variable that is specified in a DEFINEKEY method call.

The number of "KEY: **`keyvalue`**" pairs depends on the number of key variables that you define by using the DEFINEKEY method.

DATA: ***datavalue***

specifies the data value whose type must match the corresponding data variable that is specified in a DEFINEDATA method call.

The number of "DATA: **`datavalue`**" pairs depends on the number of data variables that you define by using the DEFINEDATA method.

Details

You can use the REPLACE method in one of two ways to replace data in a hash object.

You can define the key and data item, and then use the REPLACE method as shown in the following code. In this example the data for the key 'Rottwlr' is changed from '1st' to '2nd'.

```
data work.show;
   input brd $1-10 plc $12-14;
datalines;
Terrier    2nd
LabRetr    3rd
Rottwlr    1st
Collie     bis
ChinsCrstd 2nd
Newfnlnd   3rd
;


data _null_;
   length brd $12;
   length plc $8;

   if _N_ = 1 then do;
      declare hash h(dataset: 'work.show');
      rc = h.defineKey('brd');
      rc = h.defineData('plc');
      rc = h.defineDone();
   end;

   /* Specify the key and new data value */
   brd = 'Rottwlr';
   plc = '2nd';
   /* Call the REPLACE method to replace the data value */
   rc = h.replace();
run;
```

Alternatively, you can use a shortcut and specify the key and data directly in the REPLACE method call as shown in the following code:

```
data work.show;
   input brd $1-10 plc $12-14;
datalines;
Terrier    2nd
LabRetr    3rd
Rottwlr    1st
Collie     bis
ChinsCrstd 2nd
Newfnlnd   3rd
;

data _null_;
   length brd $12;
   length plc $8;

   if _N_ = 1 then do;
      declare hash h(dataset: 'work.show');
```

```
      rc = h.defineKey('brd');
      rc = h.defineData('plc');
      rc = h.defineDone();
      /* avoid uninitialized variable notes */
      call missing(brd, plc);
   end;

   /* Specify the key and new data value in the REPLACE method */
   rc = h.replace(key: 'Rottwlr', data: '2nd');
run;
```

Note: If you call the REPLACE method and the key is not found, then the key and data are added to the hash object. △

Note: The REPLACE method does not replace the value of the data variable with the value of the data item. It only replaces the value in the hash object. △

Comparisons

The REPLACE method replaces the data that is associated with the specified key with new data. The REPLACEDUP method replaces the data that is associated with the current key's current data item with new data.

See Also

Methods:

"DEFINEDATA Method" on page 2092
"DEFINEKEY Method" on page 2096
"REPLACEDUP Method" on page 2134

"Replacing and Removing Data" in *SAS Language Reference: Concepts*

REPLACEDUP Method

Replaces the data that is associated with the current key's current data item with new data.

Applies to: Hash object

Syntax

rc=*object*.**REPLACEDUP**(<DATA: *datavalue-1*,..., DATA: *datavalue-n*>);

Arguments

rc

specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash object.

DATA: ***datavalue***
specifies the data value whose type must match the corresponding data variable that is specified in a DEFINEDATA method call.

The number of "DATA: **`datavalue`**" pairs depends on the number of data variables that you define by using the DEFINEDATA method for the current key.

Details

You can use the REPLACEDUP method in one of two ways to replace data in a hash object.

You can define the data item, and then use the REPLACEDUP method. Alternatively, you can use a shortcut and specify the data directly in the REPLACEDUP method call.

Note: If you call the REPLACEDUP method and the key is not found, then the key and data are added to the hash object. △

Note: The REPLACEDUP method does not replace the value of the data variable with the value of the data item. It only replaces the value in the hash object. △

Comparisons

The REPLACEDUP method replaces the data that is associated with the current key's current data item with new data. The REPLACE method replaces the data that is associated with the specified key with new data.

Examples

This example creates a hash object where several keys have multiple data items. When a duplicate data item is found, 300 is added to the value of the data item.

```
data testdup;
   length key data 8;
   input key data;
   datalines;
   1 10
   2 11
   1 15
   3 20
   2 16
   2 9
   3 100
   5 5
   1 5
   4 6
   5 99
;

data _null_;
   length r 8;
   dcl hash h(dataset:'testdup', multidata: 'y', ordered: 'y');
   h.definekey('key');
   h.definedata('key', 'data');
   h.definedone();
   call missing (key, data);
```

```
   do key = 1 to 5;
      rc = h.find();
      if (rc = 0) then do;
         put key= data=;
         h.has_next(result: r);
         do while(r ne 0);
            rc = h.find_next();
            put 'dup ' key= data;
            data = data + 300;
            rc = h.replacedup();
            h.has_next(result: r);
         end;
      end;
   end;

   put 'iterating...';

   dcl hiter i('h');

   rc = i.first();
   do while (rc = 0);
      put key= data=;
      rc = i.next();
    end;
run;
```

The following lines are written to the SAS log.

Output 9.8 Ouput Showing Alteration of Duplicate Data Items

```
key=1 data=10
dup key=1 5
dup key=1 15
key=2 data=11
dup key=2 9
dup key=2 16
key=3 data=20
dup key=3 100
key=4 data=6
key=5 data=5
dup key=5 99
iterating...
key=1 data=10
key=1 data=305
key=1 data=315
key=2 data=11
key=2 data=309
key=2 data=316
key=3 data=20
key=3 data=400
key=4 data=6
key=5 data=5
key=5 data=399
```

See Also

Methods:

"REPLACE Method" on page 2132

"Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

SETCUR Method

Specifies a starting key item for iteration.

Applies to: Hash iterator object

Syntax

rc=*object*.**SETCUR**(KEY: *'keyvalue-1'*<,...,KEY: *'keyvalue-n'*>);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object
specifies the name of the hash iterator object.

KEY: *'keyvalue'*
specifies a key value as the starting key for the iteration.

Details

The hash iterator enables you to start iteration on any item in the hash object. The SETCUR method sets the starting key for iteration. You use the KEY option to specify the starting item.

Examples

The following example creates a data set that contains astronomical data. You want to start iteration at RA= 18 31.6 instead of the first or last items. The data is loaded into a hash object and the SETCUR method is used to start the iteration. Because the *ordered* argument tag was set to YES, note that the output is sorted in ascending order.

```
data work.astro;
input obj $1-4 ra $6-12 dec $14-19;
datalines;
 M31 00 42.7 +41 16
 M71 19 53.8 +18 47
 M51 13 29.9 +47 12
```

```
 M98 12 13.8 +14 54
 M13 16 41.7 +36 28
 M39 21 32.2 +48 26
 M81 09 55.6 +69 04
M100 12 22.9 +15 49
 M41 06 46.0 -20 44
 M44 08 40.1 +19 59
 M10 16 57.1 -04 06
 M57 18 53.6 +33 02
  M3 13 42.2 +28 23
 M22 18 36.4 -23 54
 M23 17 56.8 -19 01
 M49 12 29.8 +08 00
 M68 12 39.5 -26 45
 M17 18 20.8 -16 11
 M14 17 37.6 -03 15
 M29 20 23.9 +38 32
 M34 02 42.0 +42 47
 M82 09 55.8 +69 41
 M59 12 42.0 +11 39
 M74 01 36.7 +15 47
 M25 18 31.6 -19 15
;
```

The following code sets the starting key for iteration to **'18 31.6'**:

```
data _null_;
length obj $10;
length ra $10;
length dec $10;
declare hash myhash(hashexp: 4, dataset:"work.astro", ordered:"yes");

declare hiter iter('myhash');
myhash.defineKey('ra');
myhash.defineData('obj', 'ra');

myhash.defineDone();
call missing (ra, obj, dec);

rc = iter.setcur(key: '18 31.6');
  do while (rc = 0);
    put obj= ra=;
    rc = iter.next();
  end;
run;
```

The following lines are written to the SAS log.

Output 9.9 Output Showing Starting Key of 18.31.6

```
obj=M25 ra=18 31.6
 obj=M22 ra=18 36.4
 obj=M57 ra=18 53.6
 obj=M71 ra=19 53.8
 obj=M29 ra=20 23.9
 obj=M39 ra=21 32.2
```

You can use the FIRST method or the LAST method to start iteration on the first item or the last item, respectively.

See Also

Methods:

"FIRST Method" on page 2105

"LAST Method" on page 2111

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

"Using the Hash Iterator Object" in *SAS Language Reference: Concepts*

SUM Method

Retrieves the summary value for a given key from the hash table and stores the value in a DATA step variable.

Applies to: Hash object

Syntax

rc=object.**SUM**(SUM: *variable-name*);

Arguments

rc

specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, then an appropriate error message is written to the log.

object

specifies the name of the hash object.

SUM: ***variable-name***
specifies a DATA step variable that stores the current summary value of a given key.

Details

You use the SUM method to retrieve key summaries from the hash object. For more information, see "Maintaining Key Summaries" in *SAS Language Reference: Concepts*.

Comparisons

The SUM method retrieves the summary value for a given key when only one data item exists per key. The SUMDUP method retrieves the summary value for the current data item of the current key when more than one data item exists for a key.

Examples

The following example uses the SUM method to retrieve the key summary for each given key, K=99 and K=100.

```
k = 99;
count = 1;
h.add();
/* key=99 summary is now 1 */

k = 100;
h.add();
/* key=100 summary is now 1 */

k = 99;
h.find();
/* key=99 summary is now 2 */

count = 2;
h.find();
/* key=99 summary is now 4 */

k = 100;
h.find();
/* key=100 summary is now 3 */
h.sum(sum: total);
put 'total for key 100 = ' total;

k = 99;
h.sum(sum:total);
put 'total for key 99 = ' total;

run;
```

The first PUT statement prints the summary for k=100:

```
total for key 100 = 3
```

The second PUT statement prints the summary for k=99:

```
total for key 99 = 4
```

See Also

Methods:

"ADD Method" on page 2087
"FIND Method" on page 2100
"CHECK Method" on page 2089
"REF Method" on page 2124
"SUMDUP Method" on page 2141

Operators:

"_NEW_ Operator, Hash or Hash Iterator Object" on page 2112

Statements:

"DECLARE Statement, Hash and Hash Iterator Objects" on page 1476

SUMDUP Method

Retrieves the summary value for the current data item of the current key and stores the value in a DATA step variable.

Applies to: Hash object

Syntax

rc=*object*.**SUMDUP**(SUM: *variable-name*);

Arguments

rc
specifies whether the method succeeded or failed.

A return code of zero indicates success; a nonzero value indicates failure. If you do not supply a return code variable for the method call and the method fails, an appropriate error message is printed to the log.

object
specifies the name of the hash object.

SUM: ***variable-name***
specifies a DATA step variable that stores the summary value for the current data item of the current key.

Details

You use the SUMDUP method to retrieve key summaries from the hash object when a key has multiple data items. For more information, see "Maintaining Key Summaries" in *SAS Language Reference: Concepts*.

Comparisons

The SUMDUP method retrieves the summary value for the current data item of the current key when more than one data item exists for a key. The SUM method retrieves the summary value for a given key when only one data item exists per key.

Example

The following example uses the SUMDUP method to retrieve the summary value for the current data item. It also illustrates that it is possible to loop backward through the list by using the HAS_PREV and FIND_PREV methods. The FIND_PREV method works similarly to the FIND_NEXT method with respect to the current list item except that it moves backward through the multiple item list.

```
data dup;
   length key data 8;
   input key data;
   cards;
   1 10
   2 11
   1 15
   3 20
   2 16
   2 9
   3 100
   5 5
   1 5
   4 6
   5 99
;

data _null_;
   length r i sum 8;
   i = 0;
   dcl hash h(dataset:'dup', multidata: 'y', suminc: 'i');
   h.definekey('key');
   h.definedata('key', 'data');
   h.definedone();
   call missing (key, data);

   i = 1;
  do key = 1 to 5;
     rc = h.find();
     if (rc = 0) then do;
        h.has_next(result: r);
        do while(r ne 0);
           rc = h.find_next();
           rc = h.find_prev();
           rc = h.find_next();
           h.has_next(result: r);
        end;
     end;
  end;

 i = 0;
```

```
do key = 1 to 5;
  rc = h.find();
  if (rc = 0) then do;
      h.sum(sum: sum);
      put key= data= sum=;
      h.has_next(result: r);
      do while(r ne 0);
         rc = h.find_next();
         h.sumdup(sum: sum);
         put 'dup ' key= data= sum=;
         h.has_next(result: r);
      end;
    end;
end;
run;
```

The following lines are written to the SAS log.

Output 9.10 Key Summary

```
key=1 data=10 sum=2
dup key=1 data=5 sum=3
dup key=1 data=15 sum=2
key=2 data=11 sum=2
dup key=2 data=9 sum=3
dup key=2 data=16 sum=2
key=3 data=20 sum=2
dup key=3 data=100 sum=2
key=4 data=6 sum=1
key=5 data=5 sum=2
dup key=5 data=99 sum=2
```

To see how this works, consider the key 1,which has three data values: 10, 5, and 15 (which are stored in that order).

```
key=1 data=10 sum=2
dup key=1 data=5 sum=3
dup key=1 data=15 sum=2
```

When traveling through the data list in the loop, the key summary for 10 is set to 1 on the initial FIND method call. The first FIND_NEXT method call sets the key summary for 5 to 1. The next FIND_PREV method call moves back to the data value 10 and increments its key summary to 2. Finally, the last call to the FIND_NEXT method increments the key summary for 5 to 2. The next iteration through the loop sets the key summary for 15 to 1 and the key summary for 5 to 3 (because 5 is stored before 15 in the list). Finally, the key summary for 15 is incremented to 2. This processing results in the output for key 1 as shown in Output 5.10.

Note that you do not call the HAS_PREV method before calling the FIND_PREV method in this example because you already know that there is a previous entry in the list. Otherwise, you would not have gotten into the loop.

This example illustrates that there is no guaranteed order for multiple data items for a given key because they all have the same key. SAS cannot sort on the key. The order in the list (10, 5, 15) does not match the order that the items were added.

Also shown here is the necessity of having special methods for some duplicate operations (in this case, the SUMDUP method works similarly to the SUM method by retrieving the key summary for the current list item).

See Also

Methods:

"SUM Method" on page 2139

"Non-Unique Key and Data Pairs" in *SAS Language Reference: Concepts*

CHAPTER 10

Java Object Language Elements

Java Object Methods by Category

There are five categories of java object methods:

Table 10.1 Java Object Methods by Category

Category	Description
Deletion	enables you to delete a java object.
Exception	enables you to gather information about and clear an exception.
Field reference	enables you to return or set the value of static and non-static instance fields of the java object.
Method reference	enables you to access static and non-static Java methods.
Output	enables you to send the Java output to its destination immediately.

The following table provides brief descriptions of the java object methods. For more detailed descriptions, see the dictionary entry for each method.

Table 10.2 Categories and Descriptions of Java Object Language Elements

Category	Java Object Language Elements	Description
Deletion	"DELETE Method" on page 2151	Deletes the Java object.
Exception	"EXCEPTIONCHECK Method" on page 2152	Determines whether an exception occurred during a method call.
	"EXCEPTIONCLEAR Method" on page 2153	Clears any exception that is currently being thrown.
	"EXCEPTIONDESCRIBE Method" on page 2156	Turns the exception debug logging on or off and prints exception information.
Field reference	"GET*type*FIELD Method" on page 2159	Returns the value of a non-static field for a Java object.
	"GETSTATIC*type*FIELD Method" on page 2161	Returns the value of a static field for a Java object.
	"SET*type*FIELD Method" on page 2165	Modifies the value of a non-static field for a Java object.
	"SETSTATIC*type*FIELD Method" on page 2167	Modifies the value of a static field for a Java object.
Method reference	"CALL*type*METHOD Method" on page 2146	Invokes an instance method on a Java object from a non-static Java method.
	"CALLSTATIC*type*METHOD Method" on page 2149	Invokes an instance method on a Java object from a static Java method.
Output	"FLUSHJAVAOUTPUT Method" on page 2157	Specifies that the Java output is sent to its destination.

Dictionary

CALL*type*METHOD Method

Invokes an instance method on a java object from a non-static Java method.

Category: Method reference

Syntax

object.**CALL***type***METHOD** ("*method-name*", <*method-argument-1 …*, *method-argument-n*>, <*return value*>);

Arguments

object
specifies the name of the java object.

type
specifies the result type for the non-static Java method. The type can be one of the following values:

BOOLEAN
specifies that the result type is BOOLEAN.

BYTE
specifies that the result type is BYTE.

CHAR
specifies that the result type is CHAR.

DOUBLE
specifies that the result type is DOUBLE.

FLOAT
specifies that the result type is FLOAT.

INT
specifies that the result type is INT.

LONG
specifies that the result type is LONG.

SHORT
specifies that the result type is SHORT.

STRING
specifies that the result type is STRING.

VOID
specifies that the result type is VOID.

See Also: "Type Issues" in *SAS Language Reference: Concepts*

method-name
specifies the name of the non-static Java method.

Requirement: The method name must be enclosed in either single or double quotation marks.

method-argument
specifies the parameters to pass to the method.

return-value
specifies the return value if the method returns one.

Details

Once you instantiate a java object, you can access any non-static Java method through method calls on the java object by using the CALL*type*METHOD method.

Note: *type* represents a Java data type. For more information about how Java data types relate to SAS data types, see "Type Issues" in *SAS Language Reference: Concepts*. △

Comparisons

Use the CALL*type*METHOD method for non-static Java methods. If the Java method is static, use the CALLSTATIC*type*METHOD method.

Example

The following example creates a simple class that contains three non-static fields. The java object **j** is instantiated, the field values are set and then retrieved by using the CALL*type*FIELD method .

```
/* Java code */
import java.util.*;
import java.lang.*;
public class ttest
{
   public int i;
   public double d;
   public string s;

   public int im()
   {
      return i;
   }

   public String sm()
   {
       return s;
   }

   public double dm()
   {
      return d;
   }
}


/*  DATA step code */
data _null_;
   dcl javaobj j("ttest");
   length val 8;
   length str $20;
    j.setIntField("i", 100);
   j.setDoubleField("d", 3.14159);
   j.setStringField("s", "abc");

   j.callIntMethod("im", val);
   put val=;
   j.callDoubleMethod("dm", val);
   put val=;
   j.callStringMethod("sm", str);
   put str=;
run;
```

The following lines are written to the SAS log:

```
val=100
val=3.14159
str=abc
```

See Also

Method:

"CALLSTATIC*type*METHOD Method" on page 2149

CALLSTATIC*type*METHOD Method

Invokes an instance method on a java object from a static Java method.

Category: Method reference

Syntax

object.**CALLSTATIC***type***METHOD** ("*method-name*", <*method-argument-1 …*, *method-argument-n*>, <*return value*>);

Arguments

object
specifies the name of the java object.

type
specifies the result type for the static Java method. The type can be one of the following values:

BOOLEAN
specifies that the result type is BOOLEAN.

BYTE
specifies that the result type is BYTE.

CHAR
specifies that the result type is CHAR.

DOUBLE
specifies that the result type is DOUBLE.

FLOAT
specifies that the result type is FLOAT.

INT
specifies that the result type is INT.

LONG
specifies that the result type is LONG.

SHORT
specifies that the result type is SHORT.

STRING
specifies that the result type is STRING.

VOID
specifies that the result type is VOID.

See Also: "Type Issues" in *SAS Language Reference: Concepts*

method-name
specifies the name of the static Java method.

Requirement: The method name must be enclosed in either single or double quotation marks.

method-argument
specifies the parameters to pass to the method.

return-value
specifies the return value if the method returns one.

Details

Once you instantiate a java object, you can access any static Java method through method calls on the java object by using the CALLSTATIC*type*METHOD method.

Note: *type* represents a Java data type. For more information about how Java data types relate to SAS data types, see "Type Issues" in *SAS Language Reference: Concepts*. △

Comparisons

Use the CALLSTATIC*type*METHOD method for static Java methods. If the Java method is not static, use the CALL*type*METHOD method.

Example

The following example creates a simple class that contains three static fields. The java object **j** is instantiated, the field values are set and then retrieved by using the CALLSTATIC*type*FIELD method.

```
/* Java code */
import java.util.*;
import java.lang.*;
public class ttestc
{
   public static double d;
   public static double dm()
   {
      return d;
   }
```

```
/* DATA step code */
data x;
   declare javaobj j("ttestc");
   length d 8;

   j.SetStaticDoubleField("d", 3.14159);
   j.callStaticDoubleMethod("dm", d);
   put d=;
run;
```

The following line is written to the SAS log:

```
d=3.14159
```

See Also

Method:

"CALL*type*METHOD Method" on page 2146

DECLARE Statement, Java Object

Declares a java object; creates an instance of and initializes data for a java object.

Valid in: DATA step

See: "DECLARE Statement, Java Object" on page 1483 in *SAS Language Reference: Dictionary*

DELETE Method

Deletes the java object.

Category: Deletion

Syntax

object.**DELETE**( );

Arguments

object
specifies the name of the java object.

Details

DATA step component objects are deleted automatically at the end of the DATA step. If you want to reuse the object reference variable in another java object constructor, you should delete the java object by using the DELETE method.

If you attempt to use a java object after you delete it, you will receive an error in the log.

EXCEPTIONCHECK Method

Determines whether an exception occurred during a method call.

Category: Exception

Syntax

object.**EXCEPTIONCHECK**(*status*);

Arguments

object
specifies the name of the java object.

status
specifies the exception status that is returned.

Tip: The status value that is returned by Java is of type DOUBLE which corresponds to a SAS numeric data value.

Details

Java exceptions are handled through the EXCEPTIONCHECK, EXCEPTIONCLEAR, and EXCEPTIONDESCRIBE methods.

The EXCEPTIONCHECK method is used to determine whether an exception occurred during a method call. Ideally, the EXCEPTIONCHECK method should be called after every call to a Java method that can throw an exception.

Example

In the following example, the Java class contains a method that throws an exception. The method is called in the DATA step and a check is made for the exception.

```
/* Java code */
public class a
   {
      public void m() throws NullPointerException
      {
         throw new NullPointerException();
      }
   }
```

```
/* DATA step code */
data _null_;
   length e 8;
   dcl javaobj j('a');

   rc = j.callvoidmethod('m');

   /* Check for exception.  Value is returned in variable 'e' */
   rc = j.exceptioncheck(e);
   if (e) then
     put 'exception';
   else
     put 'no exception';

   run;
```

The following line is written to the SAS log.

```
exception
```

See Also

Method:

"EXCEPTIONCLEAR Method" on page 2153

"EXCEPTIONDESCRIBE Method" on page 2156

EXCEPTIONCLEAR Method

Clears any exception that is currently being thrown.

Category: Exception

Syntax

object.**EXCEPTIONCLEAR**( );

Arguments

object
specifies the name of the java object.

Details

Java exceptions are handled through the EXCEPTIONCHECK, EXCEPTIONCLEAR, and EXCEPTIONDESCRIBE methods.

If you call a method that throws an exception, it is strongly recommended that you check for an exception after the call. If an exception was thrown, you should perform some appropriate action and then clear the exception by using the EXCEPTIONCLEAR method.

If no exception is currently being thrown, this method has no effect.

Example

Example 1: Checking and Clearing an Exception In the following example, the Java class contains a method that throws an exception. The method is called in the DATA step and the exception is cleared.

```
/* Java code */
public class a
   {
      public void m() throws NullPointerException
      {
         throw new NullPointerException();
      }
   }


/* DATA step code */
data _null_;
   length e 8;
   dcl javaobj j('a');

   rc = j.callvoidmethod('m');

   /* Check for exception.  Value is returned in variable 'e' */
   rc = j.exceptioncheck(e);
   if (e) then
     put 'exception';
   else
     put 'no exception';

   /* Clear the exception and check it again */
   rc = j.exceptionclear( );
   rc = j.exceptioncheck(e);
   if (e) then
      put 'exception';
   else
      put 'no exception';
```

```
run;
```

The following lines are written to the SAS log.

```
exception
no exception
```

Example 2: Chekcing for an Exception When Reading an External File Iin this example, the Java IO classes are used to read an external file from the DATA step. The Java code creates a wrapper class for **DataInputStream** which enables you to pass a **FileInputStream** to the constructor. The wrapper is necessary because the constructor actually takes an **InputStream**, the parent of **FileInputStream**, and the current method lookup is not robust enough to do the superclass lookup.

```
/* Java code */
public class myDataInputStream extends java.io.DataInputStream
{
   myDataInputStream(java.io.FileInputStream fi)
   {
      super(fi);
   }
}
```

After you create the wrapper class, you can use it to create a **DataInputStream** for an external file and read the file until the end–of–file is reached. The EXCEPTIONCHECK method is used to determine when the **readInt** method throws an **EOFException**, which enables you to end the input loop.

```
/* DATA step code */
data _null_;
   length d e 8;
   dcl javaobj f("java/io/File", "c:\temp\binint.txt");
   dcl javaobj fi("java/io/FileInputStream", f);
   dcl javaobj di("myDataInputStream", fi);

   do while(1);
      di.callIntMethod("readInt", d);
      di.ExceptionCheck(e);
      if (e) then
         leave;
      else
         put d=;
   end;
run;
```

See Also

Method:

"EXCEPTIONCHECK Method" on page 2152
"EXCEPTIONDESCRIBE Method" on page 2156

EXCEPTIONDESCRIBE Method

Turns the exception debug logging off or on and prints exception information.

Category: Exception

Syntax

object.**EXCEPTIONDESCRIBE**(*status*);

Arguments

object
specifies the name of the java object.

status
specifies whether exception debug logging is on or off. *status* can be one of the following values:

0
specifies that debug logging is off.

1
specifies that debug logging is on.

Default: 0 (off)

Tip: The status value that is returned by Java is of type DOUBLE which corresponds to a SAS numeric data value.

Details

The EXCEPTIONDESCRIBE method is used to turn exception debug logging on or off. If exception debug logging is on, exception information is printed to the JVM standard output.

Note: By default, JVM standard output is redirected to the SAS log. △

Example

In the following example, exception information is printed to the standard output.

```
/* Java code */
public class a
   {
      public void m() throws NullPointerException
      {
         throw new NullPointerException();
      }
   }
```

```
/* DATA step code */
data _null_;
   length e 8;
   dcl javaobj j('a');

   j.exceptiondescribe(1);
   rc = j.callvoidmethod('m');
run;
```

The following lines are written to the SAS log:

```
java.lang.NullPointerException
     at a.m(a.java:5)
```

See Also

Method:

"EXCEPTIONCHECK Method" on page 2152
"EXCEPTIONCLEAR Method" on page 2153

FLUSHJAVAOUTPUT Method

Specifies that the Java output is sent to its destination.

Category: Output

Syntax

object.**FLUSHJAVAOUTPUT**( );

Arguments

object
specifies the name of the java object.

Details

Java output that is directed to the SAS log is flushed when the DATA step terminates. If you use the FLUSHJAVAOUTPUT method, the Java output will appear after any output that was issued while the DATA step was running.

Example

In the following example, the "In Java class" lines are written after the DATA step is complete.

```
:

/* Java code */
 public class p
{
  void p()
     {
        System.out.println("In Java class");
     }
}


/* DATA step code */
data _null_;
   dcl javaobj j('p');
   do i = 1 to 3;
        j.callVoidMethod('p');
        put 'In DATA Step';
   end;
run;
```

The following lines are written to the SAS log.

```
In DATA Step
In DATA Step
In DATA Step
In Java class
In Java class
In Java class
```

If you use the FLUSHJAVAOUTPUT method, the Java output is written to the SAS log in the order of execution.

```
/* DATA step code */
data _null_;
   dcl javaobj j('p');
   do i = 1 to 3;
        j.callVoidMethod('p');
        j.flushJavaOutput();
        put 'In DATA Step';
   end;
run;
```

The following lines are written to the SAS log.

```
In Java class
In DATA Step
In Java class
In DATA Step
In Java class
In DATA Step
```

See Also

"Java Standard Output" in *SAS Language Reference: Concepts*

GET*type*FIELD Method

Returns the value of a non-static field for a java object.

Category: Field reference

Syntax

object.**GET***type***FIELD**("*field–name*", *value*);

Arguments

object
specifies the name of a java object.

type
specifies the type for the Java field. The type can be one of the following values:

BOOLEAN
specifies that the field type is BOOLEAN.

BYTE
specifies that the field type is BYTE.

CHAR
specifies that the field type is CHAR.

DOUBLE
specifies that the field type is DOUBLE.

FLOAT
specifies that the field type is FLOAT.

INT
specifies that the field type is INT.

LONG
specifies that the field type is LONG.

SHORT
specifies that the field type is SHORT.

STRING
specifies that the field type is STRING.

See Also: "Type Issues" in *SAS Language Reference: Concepts*

field-name
specifies the Java field name.

Requirement: The field name must be enclosed in either single or double quotation marks.

value
specifies the name of variable that receives the returned field value.

Details

Once you instantiate a java object, you can access and modify its public fields through method calls on the java object. The GET*type*FIELD method enables you to access non-static fields.

Note: type represents a Java data type. For more information about how Java data types relate to SAS data types, see "Type Issues" in *SAS Language Reference: Concepts*. △

Comparisons

The GET*type*FIELD method returns the value of a non-static field for a java object. To return the value of a static field, use the GETSTATIC*type*FIELD method.

Example

The following example creates a simple class that contains three non-static fields. The java object **j** is instantiated, the field values are modified and then retrieved by using the GET*type*FIELD method.

```
/* Java code */
import java.util.*;
import java.lang.*;
public class ttest
{
   public int i;
   public double d;
   public string s;
}
}
```

```
/*  DATA step code */
data _null_;
   dcl javaobj j("ttest");
   length val 8;
   length str $20;
   j.setIntField("i", 100);
   j.setDoubleField("d", 3.14159);
   j.setStringField("s", "abc");

   j.getIntField("i", val);
   put val=;
   j.getDoubleField("d", val);
   put val=;
   j.getStringField("s", str);
   put str=;
run;
```

The following lines are written to the SAS log:

```
val=100
val=3.14159
str=abc
```

See Also

Method:

"GETSTATIC*type*FIELD Method" on page 2161
"SET*type*FIELD Method" on page 2165

GETSTATIC*type*FIELD Method

Returns the value of a static field for a java object.

Category: Field reference

Syntax

object.**GETSTATIC***type***FIELD**("*field–name*", *value*);

Arguments

object
specifies the name of a java object.

type
specifies the type for the Java field. The type can be one of the following values:

BOOLEAN
specifies that the field type is BOOLEAN.

BYTE
specifies that the field type is BYTE.

CHAR
specifies that the field type is CHAR.

DOUBLE
specifies that the field type is DOUBLE.

FLOAT
specifies that the field type is FLOAT.

INT
specifies that the field type is INT.

LONG
specifies that the field type is LONG.

SHORT
specifies that the field type is SHORT.

STRING
specifies that the field type is STRING.

See Also: "Type Issues" in *SAS Language Reference: Concepts*

field-name
specifies the Java field name.

Requirement: The field name must be enclosed in either single or double quotation marks.

value
specifies the name of variable that receives the returned field value.

Details

Once you instantiate a java object, you can access and modify its public fields through method calls on the java object. The GETSTATIC*type*FIELD method enables you to access static fields.

Note: type represents a Java data type. For more information about how Java data types relate to SAS data types, see "Type Issues" in *SAS Language Reference: Concepts*. △

Comparisons

The GETSTATIC*type*FIELD method returns the value of a static field for a java object. To return the value of a non-static field, use the GET*type*FIELD method.

Example

The following example creates a simple class that contains three static fields. The java object **j** is instantiated, the field values are set and then retrieved by using the GETSTATIC*type*FIELD method.

```
/* Java code */
import java.util.*;
import java.lang.*;
public class ttestc
{
   public static double d;
   public static double dm()
   {
      return d;
   }
}
```

```
/* DATA step code */
data x;
   declare javaobj j("ttestc");
   length d 8;
```

```
   j.callSetStaticDoubleField("d", 3.14159);
   j.callStaticDoubleMethod("dm", d);
   put d=;
run;
```

The following line is written to the SAS log:

```
d=3.14159
```

See Also

Method:

"GET*type*FIELD Method" on page 2159
"SETSTATIC*type*FIELD Method" on page 2167

NEW Operator, Java Object

Creates an instance of a java object.

Valid in: DATA step

Syntax

object-reference = **_NEW_ JAVAOBJ** ("*java-class*", <*argument–1* , … *argument–n*>);

Arguments

object-reference
specifies the object reference name for the java object.

java-class
specifies the name of the Java class to be instantiated.

Requirement: The Java class name must be enclosed in either double or single quotation marks.

argument
specifies the information that is used to create an instance of the java object. Valid values for *argument* depend on the java object.

See Also: "Details" on page 2163

Details

To use a DATA step component object in your SAS program, you must declare and create (instantiate) the object. The DATA step component interface provides a mechanism for accessing the predefined component objects from within the DATA step.

If you use the _NEW_ operator to instantiate the java object, you must first use the DECLARE statement to declare the java object. For example, in the following lines of code, the DECLARE statement tells SAS that the object reference J is a java object. The _NEW_ operator creates the java object and assigns it to the object reference J.

```
declare javaobj j;
j = _new_ javaobj("somejavaclass" );
```

Note: You can use the DECLARE statement to declare and instantiate a java object in one step. △

A constructor is a method that is used to instantiate a component object and to initialize the component object data. For example, in the following lines of code, the _NEW_ operator instantiates a java object and assigns it to the object reference J. Note that the only required argument to a java object constructor is the name of the Java class to be instantiated. All other arguments are constructor arguments to the Java class itself. In addition, the Java classname, **testjavaclass**, he constructor and the values **100** and **.8** are constructor arguments.

```
declare javaobj j;
j = _new_ javaobj("testjavaclass", 100, .8);
```

For more information about the predefined DATA step component objects and constructors, see "Using DATA Step Component Objects" in *SAS Language Reference: Concepts*.

Comparisons

You can use the DECLARE statement and the _NEW_ operator, or the DECLARE statement alone to declare and instantiate an instance of a java object.

Examples

In the following example, a Java class is created for a hash table. The _NEW_ operator is used to create and instantiate an instance of this class by specifying the capacity and load factor. In this example, a wrapper class, **mhash**, is necessary because the DATA step's only numeric type is equivalent to the Java type DOUBLE.

```
/* Java code */
import java.util.*;

public class mhash extends Hashtable;
{
   mhash (double size, double load)
      {
         super ((int)size, (float)load);
      }
}
```

```
/* DATA step code */
data _null_;
   declare javaobj h;
   h = _new_ javaobj("mhash", 100, .8);
run;
```

See Also

Statements:

"DECLARE Statement, Java Object" on page 1483

"Using DATA Step Component Objects" in *SAS Language Reference: Concepts*

SET*type*FIELD Method

Modifies the value of a non-static field for a java object.

Category: Field reference

Syntax

object.**SET***type***FIELD**("*field–name*", *value*);

Arguments

object
specifies the name of a java object.

type
specifies the type for the Java field. The type can be one of the following values:

BOOLEAN
specifies that the field type is BOOLEAN.

BYTE
specifies that the field type is BYTE.

CHAR
specifies that the field type is CHAR.

DOUBLE
specifies that the field type is DOUBLE.

FLOAT
specifies that the field type is FLOAT.

INT
specifies that the field type is INT.

LONG
specifies that the field type is LONG.

SHORT
specifies that the field type is SHORT.

STRING
specifies that the field type is STRING.

See Also: "Type Issues" in *SAS Language Reference: Concepts*

field-name
specifies the Java field name.

Requirement: The field name must be enclosed in either single or double quotation marks.

value
specifies the value.

Details

Once you instantiate a java object, you can access and modify its public fields through method calls on the java object. The SET*type*FIELD method enables you to modify non-static fields.

Note: *type* represents a Java data type. For more information about how Java data types relate to SAS data types, see "Type Issues" in *SAS Language Reference: Concepts*. △

Comparisons

The SET*type*FIELD method modifies the value of a non-static field for a java object. To modify the value of a static field, use the SETSTATIC*type*FIELD method.

Example

The following example creates a simple class that contains three non-static fields. The java object **j** is instantiated, the field values are set by using the SET*type*FIELD method and then retrieved.

```
/* Java code */
import java.util.*;
import java.lang.*;
public class ttest
{
   public int i;
   public double d;
   public string s;
}
}
```

```
/*  DATA step code */
data _null_;
   dcl javaobj j("ttest");
   length val 8;
   length str $20;
   j.setIntField("i", 100);
   j.setDoubleField("d", 3.14159);
   j.setStringField("s", "abc");

   j.getIntField("i", val);
   put val=;
   j.getDoubleField("d", val);
```

```
   put val=;
   j.getStringField("s", str);
   put str=;
run;
```

The following lines are written to the SAS log:

```
val=100
val=3.14159
str=abc
```

See Also

Method:

"GET*type*FIELD Method" on page 2159
"SETSTATIC*type*FIELD Method" on page 2167

SETSTATIC*type*FIELD Method

Modifies the value of a static field for a java object.

Category: Field reference

Syntax

object.**SETSTATIC***type***FIELD**("*field–name*", *value*);

Arguments

object
specifies the name of a java object.

type
specifies the type for the Java field. The type can be one of the following values:

BOOLEAN
specifies that the field type is BOOLEAN.

BYTE
specifies that the field type is BYTE.

CHAR
specifies that the field type is CHAR.

DOUBLE
specifies that the field type is DOUBLE.

FLOAT
specifies that the field type is FLOAT.

INT
specifies that the field type is INT.

LONG
specifies that the field type is LONG.

SHORT
specifies that the field type is SHORT.

STRING
specifies that the field type is STRING.

See Also: "Type Issues" in *SAS Language Reference: Concepts*

field-name
specifies the Java field name.

Requirement: The field name must be enclosed in either single or double quotation marks.

value
specifies the value.

Details

Once you instantiate a java object, you can access and modify its public fields through method calls on the java object. The SETSTATIC*type*FIELD method enables you to modify static fields.

Note: type represents a Java data type. For more information about how Java data types relate to SAS data types, see "Type Issues" in *SAS Language Reference: Concepts*. △

Comparisons

The SETSTATIC*type*FIELD method modifies the value of a static field for a java object. To modify the value of a non–static field, use the SET*type*FIELD method.

Example

The following example creates a simple class that contains three static fields. The java object **j** is instantiated, the field values are set by using the SETSTATIC*type*FIELD method and then retrieved.

```
/* Java code */
import java.util.*;
import java.lang.*;
public class ttestc
{
   public static double d;
   public static double dm()
   {
      return d;
   }
}
```

```
/* DATA step code */
data x;
   declare javaobj j("ttestc");
   length d 8;

   j.callSetStaticDoubleField("d", 3.14159);
   j.callStaticDoubleMethod("dm", d);
   put d=;
run;
```

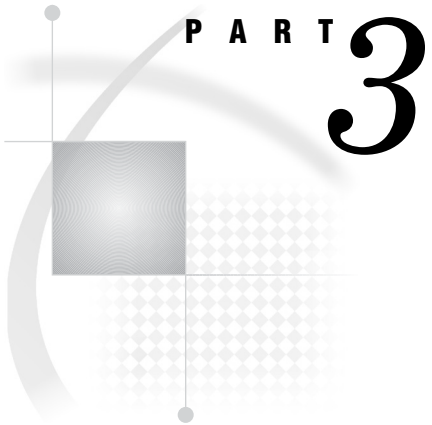
The following line is written to the SAS log:

```
d=3.14159
```

See Also

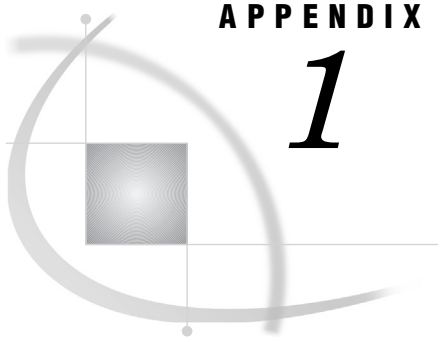
Method:

"GETSTATIC*type*FIELD Method" on page 2161
"SET*type*FIELD Method" on page 2165



PART 3

Appendixes



APPENDIX

1

DATA Step Debugger

Introduction

Definition: What Is Debugging?

Debugging is the process of removing logic errors from a program. Unlike syntax errors, logic errors do not stop a program from running. Instead, they cause the program to produce unexpected results. For example, if you create a DATA step that keeps track of inventory, and your program shows that you are out of stock but your warehouse is full, you have a logic error in your program.

To debug a DATA step, you could do the following:

- copy a few lines of the step into another DATA step, execute it, and print the results of those statements
- insert PUT statements at selected places in the DATA step, submit the step, and examine the values that are displayed in the SAS log.
- use the DATA step debugger.

While the SAS log can help you identify data errors, the DATA step debugger offers you an easier, interactive way to identify logic errors, and sometimes data errors, in DATA steps.

Definition: The DATA Step Debugger

The DATA step debugger is part of Base SAS software and consists of windows and a group of commands. By issuing commands, you can execute DATA step statements one by one and pause to display the resulting variable values in a window. By observing the results that are displayed, you can determine where the logic error lies. Because the debugger is interactive, you can repeat the process of issuing commands and observing the results as many times as needed in a single debugging session. To invoke the debugger, add the DEBUG option to the DATA statement and execute the program.

The DATA step debugger enables you to perform the following tasks:

- execute statements one by one or in groups
- bypass execution of one or more statements
- suspend execution at selected statements, either in each iteration of DATA step statements or on a condition you specify, and resume execution on command
- monitor the values of selected variables and suspend execution at the point a value changes
- display the values of variables and assign new values to them
- display the attributes of variables
- receive help for individual debugger commands
- assign debugger commands to function keys
- use the macro facility to generate customized debugger commands.

Basic Usage

How a Debugger Session Works

When you submit a DATA step with the DEBUG option, SAS compiles the step, displays the debugger windows, and pauses until you enter a debugger command to begin execution. For example, if you begin execution with the GO command, SAS executes each statement in the DATA step. To suspend execution at a particular line in the DATA step, use the BREAK command to set breakpoints at statements you select. Then issue the GO command. The GO command starts or resumes execution until the breakpoint is reached.

To execute the DATA step one statement at a time or a few statements at a time, use the STEP command. By default, the STEP command is mapped to the ENTER key.

In a debugging session, statements in a DATA step can iterate as many times as they would outside the debugging session. When the last iteration has finished, a message appears in the DEBUGGER LOG window.

You cannot restart DATA step execution in a debugging session after the DATA step finishes executing. You must resubmit the DATA step in your SAS session. However, you can examine the final values of variables after execution has ended.

You can debug only one DATA step at a time. You can use the debugger only with a DATA step, and not with a PROC step.

Using the Windows

The DATA step debugger contains two primary windows, the DEBUGGER LOG and the DEBUGGER SOURCE windows. The windows appear when you execute a DATA step with the DEBUG option.

The DEBUGGER LOG window records the debugger commands you issue and their results. The last line is the debugger command line, where you issue debugger commands. The debugger command line is marked with a greater than (>) prompt.

The DEBUGGER SOURCE window contains the SAS statements that comprise the DATA step you are debugging. The window enables you to view your position in the DATA step as you debug your program. In the window, the SAS statements have the same line numbers as they do in the SAS log.

You can enter windowing environment commands on the window command lines. You can also execute commands by using function keys.

Entering Commands

Enter DATA step debugger commands on the debugger command line. For a list of commands and their descriptions, refer to “Debugger Commands by Category” on page 2190. Follow these rules when you enter a command:

- A command can occupy only one line (except for a DO group).
- A DO group can extend over more than one line.
- To enter multiple commands, separate the commands with semicolons:

```
examine _all_; set letter='bill'; examine letter
```

Working with Expressions

All SAS operators that are described in "SAS Operators in Expressions" in *SAS Language Reference: Concepts*, are valid in debugger expressions. Debugger expressions cannot contain functions.

A debugger expression must fit on one line. You cannot continue an expression on another line.

Assigning Commands to Function Keys

To assign debugger commands to function keys, open the Keys window. Position your cursor in the Definitions column of the function key you want to assign, and begin the command with the term DSD. To assign more than one command to a function key, enclose the commands (separated by semicolons) in quotation marks. Be sure to save your changes. These examples show commands assigned to function keys:

□ `dsd step3`

□ `dsd 'examine cost saleprice; go 120;'`

Advanced Usage: Using the Macro Facility with the Debugger

You can use the SAS macro facility with the debugger to invoke macros from the DEBUGGER LOG command line. You can also define macros and use macro program statements, such as %LET, on the debugger command line.

Using Macros as Debugging Tools

Macros are useful for storing a series of debugger commands. Executing the macro at the DEBUGGER LOG command line then generates the entire series of debugger commands. You can also use macros with parameters to build different series of debugger commands based on various conditions.

Creating Customized Debugging Commands with Macros

You can create a customized debugging command by defining a macro on the DEBUGGER LOG command line. Then invoke the macro from the command line. For example, to examine the variable COST, to execute five statements, and then to examine the variable DURATION, define the following macro (in this case the macro is called EC). Note that the example uses the alias for the EXAMINE command.

```
%macro ec; ex cost; step 5; ex duration; %mend ec;
```

To issue the commands, invoke macro EC from the DEBUGGER LOG command line:

```
%ec
```

The DEBUGGER LOG displays the value of COST, executes the next five statements, and then displays the value of DURATION.

Note: Defining a macro on the DEBUGGER LOG command line allows you to use the macro only during the current debugging session, because the macro is not permanently stored. To create a permanently stored macro, use the Program Editor. △

Debugging a DATA Step Generated by a Macro

You can use a macro to generate a DATA step, but debugging a DATA step that is generated by a macro can be difficult. The SAS log displays a copy of the macro, but not the DATA step that the macro generated. If you use the DEBUG option at this point, the text that the macro generates appears as a continuous stream to the debugger. As a result, there are no line breaks where execution can pause.

To debug a DATA step that is generated by a macro, use the following steps:

1. Use the MPRINT and MFILE system options when you execute your program.
2. Assign the fileref MPRINT to an existing external file. MFILE routes the program output to the external file. Note that if you rerun your program, current output appends to the previous output in your file.
3. Invoke the macro from a SAS session.
4. In the Program Editor window, issue the INCLUDE command or use the File menu to open your external file.
5. Add the DEBUG option to the DATA statement and begin a debugging session.
6. When you locate the logic error, correct the portion of the macro that generated that statement or statements.

Examples

Example 1: Debugging a Simple DATA Step

This example shows how to debug a DATA step when output is missing.

Discovering a Problem

This program creates information about a travel tour group. The data files contain two types of records. One type contains the tour code, and the other type contains customer information. The program creates a report listing tour number, name, age, and sex for each customer.

```
/* first execution */
data tours (drop=type);
   input @1 type $ @;
   if type='H' then do;
      input @3 Tour $20.;
      return;
      end;
   else if type='P' then do;
      input @3 Name $10. Age 2. +1 Sex $1.;
      output;
      end;
   datalines;
H Tour 101
P Mary E    21 F
P George S  45 M
P Susan K    3 F
```

```
H Tour 102
P Adelle S  79 M
P Walter P  55 M
P Fran I    63 F
;

proc print data=tours;
   title 'Tour List';
run;
```

```
                          Tour List                                1

            Obs     Tour      Name       Age     Sex

             1              Mary E        21      F
             2              George S      45      M
             3              Susan K        3      F
             4              Adelle S      79      M
             5              Walter P      55      M
             6              Fran I        63      F
```

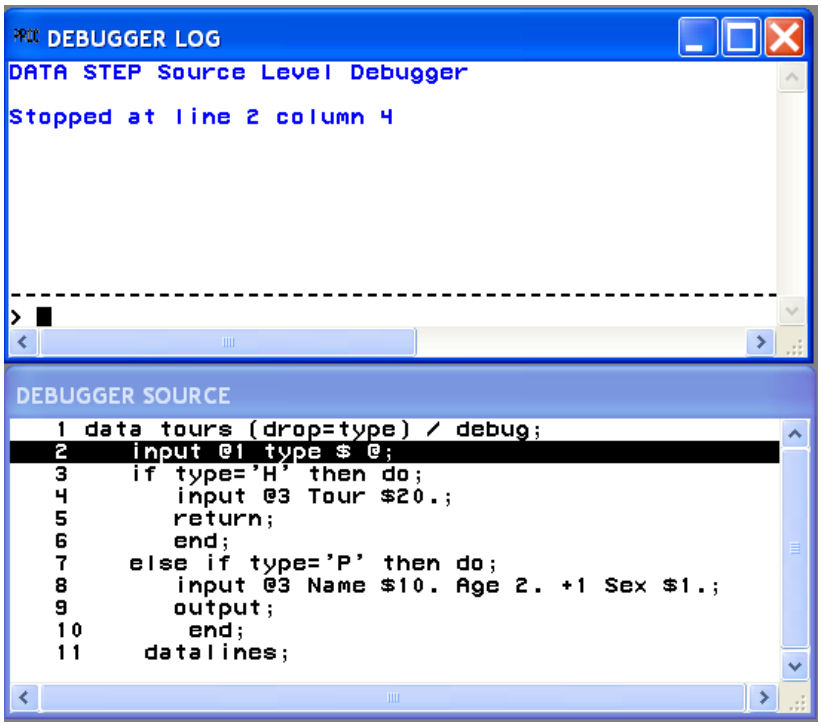
The program executes without error, but the output is unexpected. The output does not contain values for the variable Tour. Viewing the SAS log will not help you debug the program because the data are valid and no errors appear in the log. To help identify the logic error, run the DATA step again using the DATA step debugger.

Using the DEBUG Option

To invoke the DATA step debugger, add the DEBUG option to the DATA statement and resubmit the DATA step:

```
data tours (drop=type) / debug;
```

The following display shows the resulting two debugger windows.



The upper window is the DEBUGGER LOG window. Issue debugger commands in this window by typing commands on the debugger command line (the bottom line, marked by a >). The debugger displays the command and results in the upper part of the window.

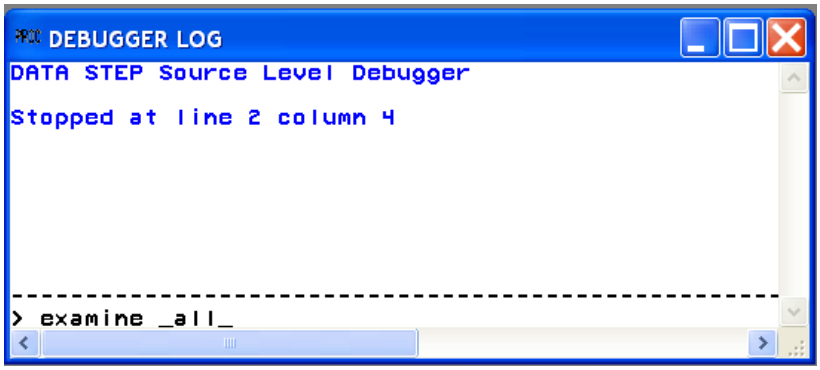
The lower window is the DEBUGGER SOURCE window. It displays the DATA step submitted with the DEBUG option. Each line in the DATA step is numbered with the same line number used in the SAS log. One line appears in reverse video (or other highlighting, depending on your monitor). DATA step execution pauses *just before* the execution of the highlighted statement.

At the beginning of your debugging session, the first executable line after the DATA statement is highlighted. This means that SAS has compiled the step and will begin to execute the step at the top of the DATA step loop.

Examining Data Values after the First Iteration

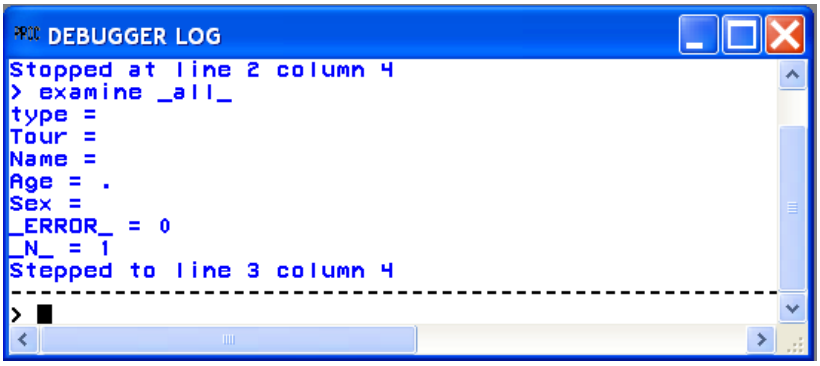
To debug a DATA step, create a hypothesis about the logic error and test it by examining the values of variables at various points in the program. For example, issue the EXAMINE command from the debugger command line to display the values of all variables in the program data vector before execution begins:

```
examine _all_
```



Note: Most debugger commands have abbreviations, and you can assign commands to function keys. The examples in this section, however, show the full command name to help you find the commands in "Debugger Commands by Category" on page 2190. △

When you press ENTER, the following display appears:



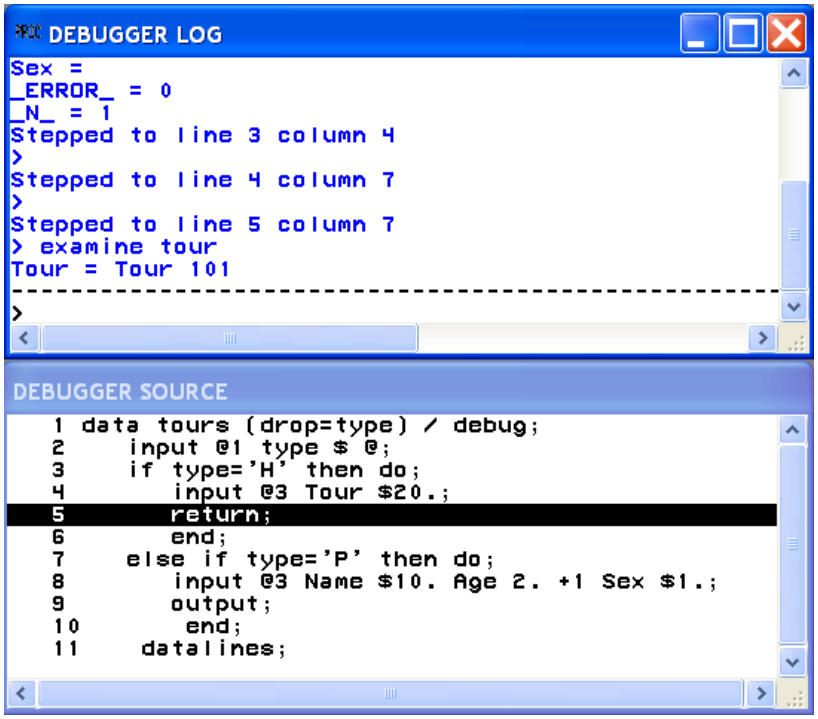
The values of all variables appear in the DEBUGGER LOG window. SAS has compiled, but not yet executed, the INPUT statement.

Use the STEP command to execute the DATA step statements one at a time. By default, the STEP command is assigned to the ENTER key. Press ENTER repeatedly to step through the first iteration of the DATA step, and stop when the RETURN statement in the program is highlighted in the DEBUGGER SOURCE window.

Because Tour information was missing in the program output, enter the EXAMINE command to view the value of the variable Tour for the first iteration of the DATA step.

```
examine tour
```

The following display shows the results:



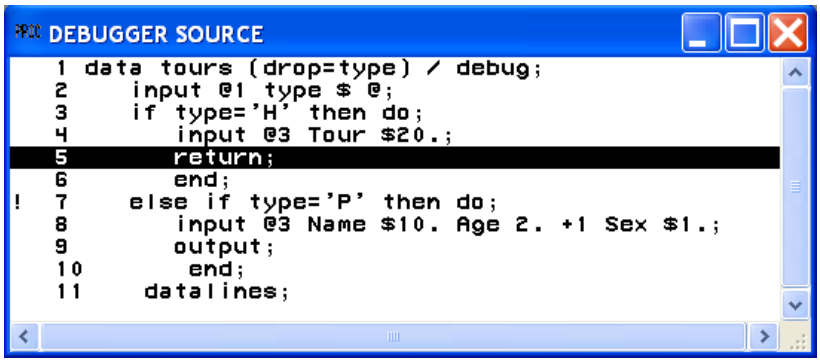
The variable Tour contains the value Tour 101, showing you that Tour was read. The first iteration of the DATA step worked as intended. Press ENTER to reach the top of the DATA step.

Examining Data Values after the Second Iteration

You can use the BREAK command (also known as *setting a breakpoint*) to suspend DATA step execution at a particular line you designate. In this example, suspend execution before executing the ELSE statement by setting a breakpoint at line 7.

```
break 7
```

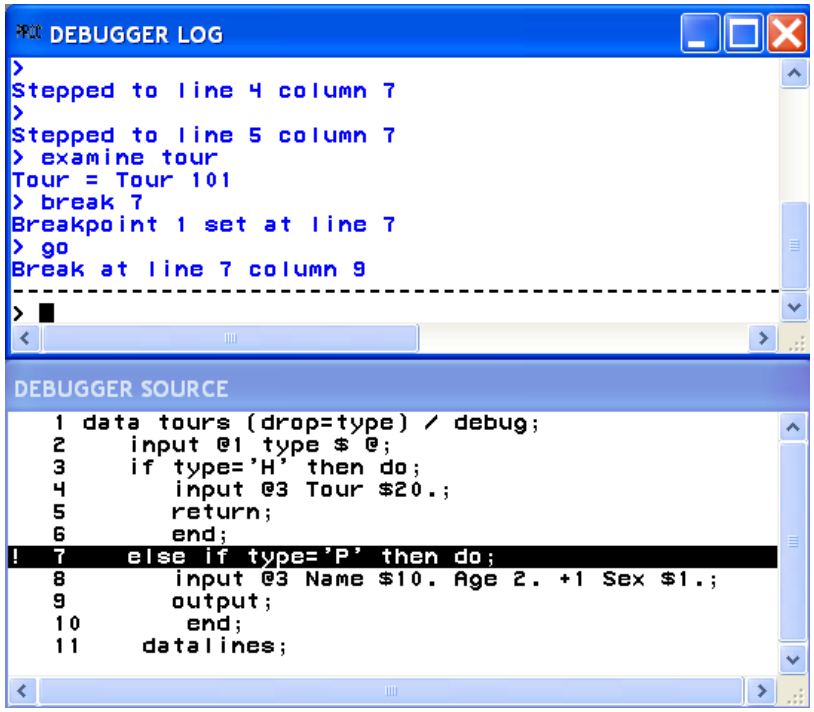
When you press ENTER, an exclamation point appears at line 7 in the DEBUGGER SOURCE window to mark the breakpoint:



Execute the GO command to continue DATA step execution until it reaches the breakpoint (in this case, line 7):

```
go
```

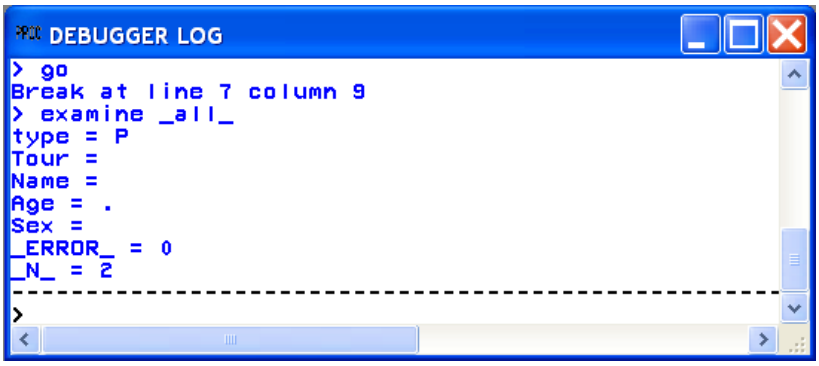
The following display shows the result:



SAS suspended execution *just before* the ELSE statement in line 7. Examine the values of all the variables to see their status at this point.

```
examine _all_
```

The following display shows the values:



You expect to see a value for Tour, but it does not appear. The program data vector gets reset to missing values at the beginning of each iteration and therefore does not retain the value of Tour. To solve the logic problem, you need to include a RETAIN statement in the SAS program.

Ending the Debugger

To end the debugging session, issue the QUIT command on the debugger command line:

```
quit
```

The debugging windows disappear, and the original SAS session resumes.

Correcting the DATA Step

Correct the original program by adding the RETAIN statement. Delete the DEBUG option from the DATA step, and resubmit the program:

```
   /* corrected version */
data tours (drop=type);
   retain Tour;
   input @1 type $ @;
   if type='H' then do;
      input @3 Tour $20.;
      return;
      end;
   else if type='P' then do;
      input @3 Name $10. Age 2. +1 Sex $1.;
      output;
      end;
datalines;
H Tour 101
P Mary E    21 F
P George S  45 M
P Susan K    3 F
H Tour 102
P Adelle S  79 M
P Walter P  55 M
P Fran I    63 F
;
```

```
run;

proc print;
   title 'Tour List';
run;
```

The values for Tour now appear in the output:

```
                          Tour List                                1
          Obs      Tour         Name       Age     Sex

            1     Tour 101    Mary E        21      F
            2     Tour 101    George S      45      M
            3     Tour 101    Susan K        3      F
            4     Tour 102    Adelle S      79      M
            5     Tour 102    Walter P      55      M
            6     Tour 102    Fran I        63      F
```

Example 2: Working with Formats

This example shows how to debug a program when you use format statements to format dates. The following program creates a report that lists travel tour dates for specific countries.

```
options yearcutoff=1920;

data tours;
  length Country $ 10;
  input Country $10. Start : mmddyy. End : mmddyy.;
  Duration=end-start;
datalines;
Italy       033000 041300
Brazil      021900 022800
Japan       052200 061500
Venezuela   110300 11800
Australia   122100 011501
;

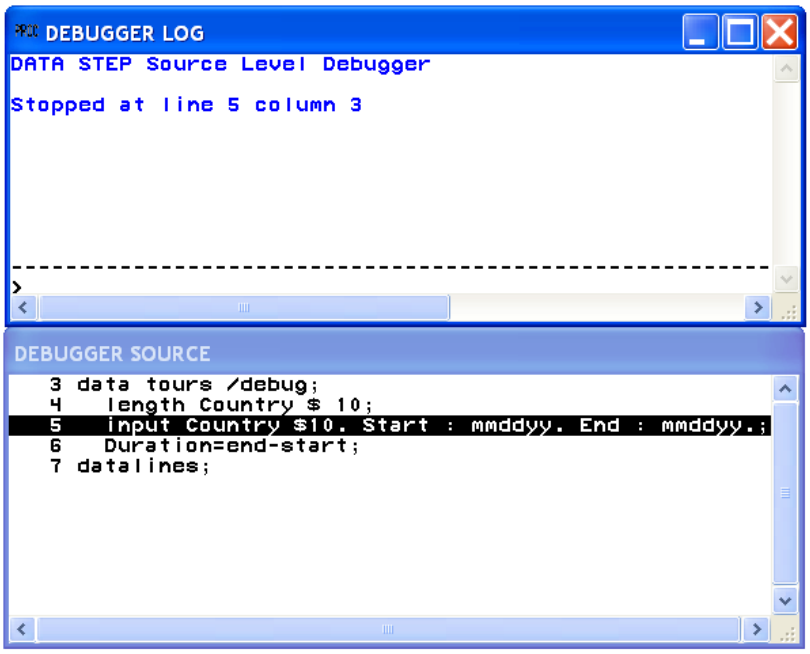
proc print data=tours;
   format start end date9.;
   title 'Tour Duration';
run;
```

```
                        Tour Duration                              1

     Obs    Country          Start           End    Duration

       1    Italy        30MAR2000    13APR2000        14
       2    Brazil       19FEB2000    28FEB2000         9
       3    Japan        22MAY2000    15JUN2000        24
       4    Venezuela    03NOV2000    18JAN2000      -290
       5    Australia    21DEC2000    15JAN2001        25
```

The value of Duration for the tour to Venezuela shows a negative number, -290 days. To help identify the error, run the DATA step again using the DATA step debugger. SAS displays the following debugger windows:



At the DEBUGGER LOG command line, issue the EXAMINE command to display the values of all variables in the program data vector before execution begins:

```
examine _all_
```

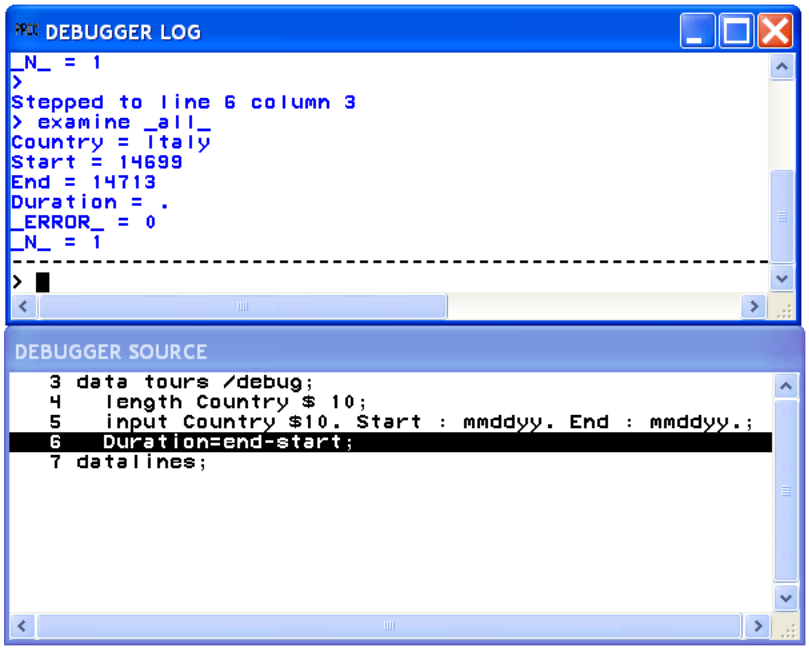
Initial values of all variables appear in the DEBUGGER LOG window. SAS has not yet executed the INPUT statement.

Press ENTER to issue the STEP command. SAS executes the INPUT statement, and the assignment statement is now highlighted.

Issue the EXAMINE command to display the current value of all variables:

```
examine _all_
```

The following display shows the results:



Because a problem exists with the Venezuela tour, suspend execution before the assignment statement when the value of Country equals Venezuela. Set a breakpoint to do this:

```
break 6 when country='Venezuela'
```

Execute the GO command to resume program execution:

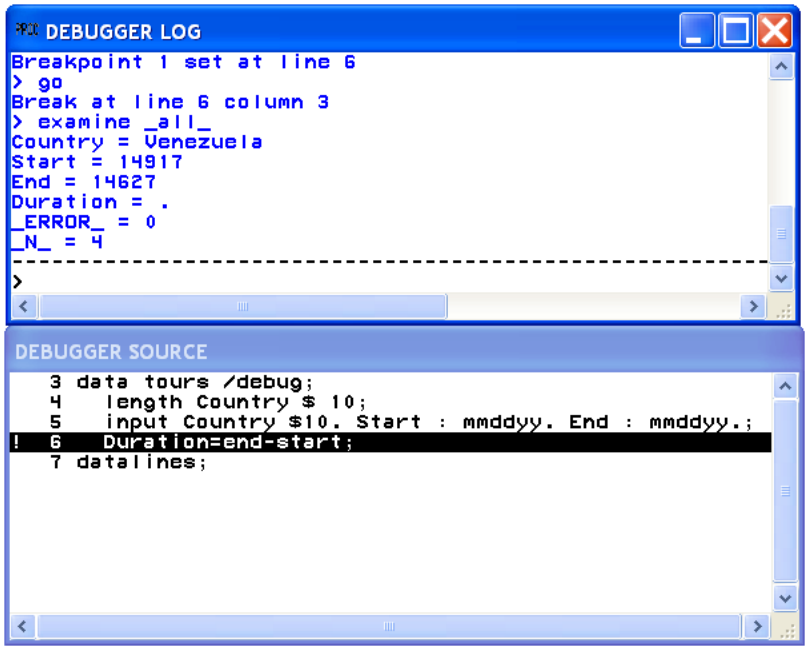
```
go
```

SAS stops execution when the country name is Venezuela. You can examine Start and End tour dates for the Venezuela trip. Because the assignment statement is highlighted (indicating that SAS has not yet executed that statement), there will be no value for Duration.

Execute the EXAMINE command to view the value of the variables after execution:

```
examine _all_
```

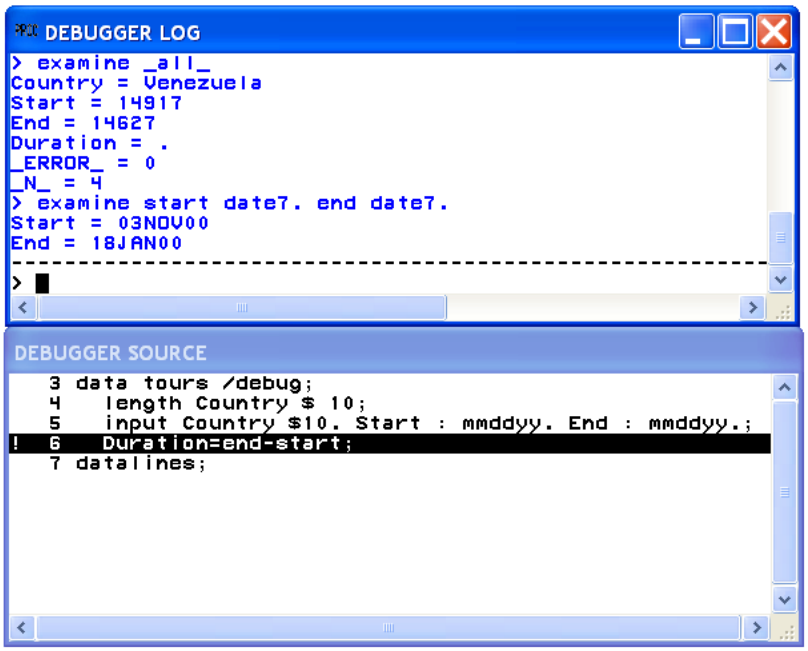
The following display shows the results:



To view formatted SAS dates, issue the EXAMINE command using the DATE*w*. format:

```
examine start date7. end date7.
```

The following display shows the results:



Because the tour ends on November 18, 2000, and not on January 18, 2000, there is an error in the variable End. Examine the source data in the program and notice that the value for End has a typographical error. By using the SET command, you can

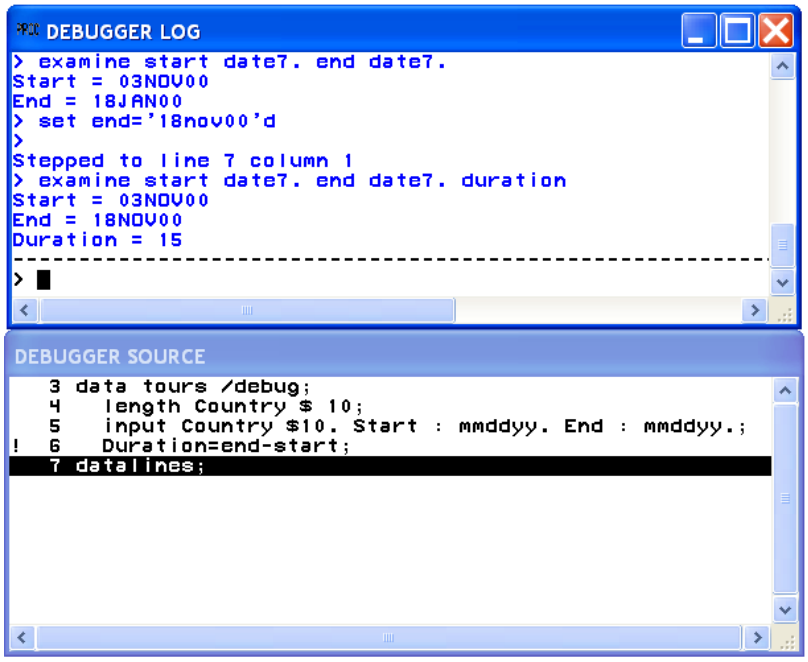
temporarily set the value of End to November 18 to see whether you get the anticipated result. Issue the SET command using the DDMMMYY*w*. format:

```
set end='18nov00'd
```

Press ENTER to issue the STEP command and execute the assignment statement. Issue the EXAMINE command to view the tour date and Duration fields:

```
examine start date7. end date7. duration
```

The following display shows the results:



The Start, End, and Duration fields contain correct data.

End the debugging session by issuing the QUIT command on the DEBUGGER LOG command line. Correct the original data in the SAS program, delete the DEBUG option, and resubmit the program.

```
   /* corrected version */
options yearcutoff=1920;

data tours;
  length Country $ 10;
  input Country $10. Start : mmddyy. End : mmddyy.;
  duration=end-start;
datalines;
Italy       033000 041300
Brazil      021900 022800
Japan       052200 061500
Venezuela   110300 111800
Australia   122100 011501
;

proc print data=tours;
```

```
   format start end date9.;
   title 'Tour Duration';
run;
```

Tour Duration 1

Obs	Country	Start	End	duration
1	Italy	30MAR2000	13APR2000	14
2	Brazil	19FEB2000	28FEB2000	9
3	Japan	22MAY2000	15JUN2000	24
4	Venezuela	03NOV2000	18NOV2000	15
5	Australia	21DEC2000	15JAN2001	25

Example 3: Debugging DO Loops

An iterative DO, DO WHILE, or DO UNTIL statement can iterate many times during a single iteration of the DATA step. When you debug DO loops, you can examine several iterations of the loop by using the AFTER option in the BREAK command. The AFTER option requires a number that indicates how many times the loop will iterate before it reaches the breakpoint. The BREAK command then suspends program execution. For example, consider this data set:

```
data new / debug;
   set old;
   do i=1 to 20;
      newtest=oldtest+i;
      output;
   end;
run;
```

To set a breakpoint at the assignment statement (line 4 in this example) after every five iterations of the DO loop, issue this command:

```
break 4 after 5
```

When you issue the GO commands, the debugger suspends execution when I has the values of 5, 10, 15, and 20 in the DO loop iteration.

In an iterative DO loop, select a value for the AFTER option that can be divided evenly into the number of iterations of the loop. For example, in this DATA step, 5 can be evenly divided into 20. When the DO loop iterates the second time, I again has the values of 5, 10, 15, and 20.

If you do not select a value that can be evenly divided (such as 3 in this example), the AFTER option causes the debugger to suspend execution when I has the values of 3, 6, 9, 12, 15, and 18. When the DO loop iterates the second time, I has the values of 1, 4, 7, 10, 13, and 16.

Example 4: Examining Formatted Values of Variables

You can use a SAS format or a user-created format when you display a value with the EXAMINE command. For example, assume that the variable BEGIN contains a SAS date value. To display the day of the week and date, use the SAS WEEKDATE*w*. format with EXAMINE:

```
examine begin weekdate17.
```

When the value of BEGIN is 033001, the debugger displays the following:

```
Sun, Mar 30, 2001
```

As another example, you can create a format named SIZE:

```
proc format;
   value size 1-5='small'
              6-10='medium'
             11-high='large';
run;
```

To debug a DATA step that applies the format SIZE. to the variable STOCKNUM, use the format with EXAMINE:

```
examine stocknum size.
```

For example, when the value of STOCKNUM is 7, the debugger displays the following:

```
STOCKNUM = medium
```

Commands

List of Debugger Commands

BREAK
CALCULATE
DELETE
DESCRIBE
ENTER
EXAMINE
GO
HELP
JUMP
LIST
QUIT
SET
STEP
SWAP
TRACE
WATCH

Debugger Commands by Category

Table A1.1 Categories and Descriptions of Debugger Commands

Category	DATA Step Debugger	Description
Controlling Program Execution	"GO" on page 2197	Starts or resumes execution of the DATA step
	"JUMP" on page 2198	Restarts execution of a suspended program
	"STEP" on page 2202	Executes statements one at a time in the active program
Controlling the Windows	"HELP" on page 2198	Displays information about debugger commands
	"SWAP" on page 2203	Switches control between the SOURCE window and the LOG window
Manipulating DATA Step Variables	"CALCULATE" on page 2193	Evaluates a debugger expression and displays the result
	"DESCRIBE" on page 2195	Displays the attributes of one or more variables
	"EXAMINE" on page 2196	Displays the value of one or more variables
	"SET" on page 2201	Assigns a new value to a specified variable
Manipulating Debugging Requests	"BREAK" on page 2190	Suspends program execution at an executable statement
	"DELETE" on page 2193	Deletes breakpoints or the watch status of variables in the DATA step
	"LIST" on page 2199	Displays all occurrences of the item that is listed in the argument
	"TRACE" on page 2203	Controls whether the debugger displays a continuous record of the DATA step execution
	"WATCH" on page 2204	Suspends execution when the value of a specified variable changes
Tailoring the Debugger	"ENTER" on page 2195	Assigns one or more debugger commands to the ENTER key
Terminating the Debugger	"QUIT" on page 2200	Terminates a debugger session

Dictionary

BREAK

Suspends program execution at an executable statement.

Category: Manipulating Debugging Requests

Alias: B

Syntax

BREAK *location* <AFTER *count*> <WHEN *expression*> <DO *group* >

Arguments

location
specifies where to set a breakpoint. *Location* must be one of these:

label	a statement label. The breakpoint is set at the statement that follows the label.
line-number	the number of a program line at which to set a breakpoint.
*	the current line.

AFTER *count*
honors the breakpoint each time the statement has been executed *count* times. The counting is continuous. That is, when the AFTER option applies to a statement inside a DO loop, the count continues from one iteration of the loop to the next. The debugger does not reset the *count* value to 1 at the beginning of each iteration.

If a BREAK command contains both AFTER and WHEN, AFTER is evaluated first. If the AFTER count is satisfied, the WHEN expression is evaluated.

Tip: The AFTER option is useful in debugging DO loops.

WHEN *expression*
honors a breakpoint when the expression is true.

DO *group*
is one or more debugger commands enclosed by a DO and an END statement. The syntax of the DO *group* is the following:

DO; *command-1* < ... ; *command-n*; >END;

command
specifies a debugger command. Separate multiple commands by semicolons.

A DO group can span more than one line and can contain IF-THEN/ELSE statements, as shown:

IF *expression* THEN *command*; <ELSE *command*;>
IF *expression* THEN DO *group*; <ELSE DO *group*;>

IF evaluates an expression. When the condition is true, the debugger command or DO group in the THEN clause executes. An optional ELSE command gives an alternative action if the condition is not true. You can use these arguments with IF:

expression
specifies a debugger expression. A non-zero, nonmissing result causes the expression to be true. A result of zero or missing causes the expression to be false.

command
specifies a single debugger command.

DO *group*
specifies a DO group.

Details

The BREAK command suspends execution of the DATA step at a specified statement. Executing the BREAK command is called *setting a breakpoint*.

When the debugger detects a breakpoint, it does the following:

- checks the AFTER *count* value, if present, and suspends execution if *count* breakpoint activations have been reached
- evaluates the WHEN expression, if present, and suspends execution if the condition that is evaluated is true
- suspends execution if neither an AFTER nor a WHEN clause is present
- displays the line number at which execution is suspended
- executes any commands that are present in a DO group
- returns control to the user with a > prompt.

If a breakpoint is set at a source line that contains more than one statement, the breakpoint applies to each statement on the source line. If a breakpoint is set at a line that contains a macro invocation, the debugger breaks at each statement generated by the macro.

Examples

- Set a breakpoint at line 5 in the current program:

```
b 5
```

- Set a breakpoint at the statement after the statement label **eoflabel**:

```
b eoflabel
```

- Set a breakpoint at line 45 that will be honored after every third execution of line 45:

```
b 45 after 3
```

- Set a breakpoint at line 45 that will be honored after every third execution of that line only when the values of both DIVISOR and DIVIDEND are 0:

```
b 45 after 3
       when (divisor=0 and dividend=0)
```

- Set a breakpoint at line 45 of the program and examine the values of variables NAME and AGE:

```
b 45 do; ex name age; end;
```

- Set a breakpoint at line 15 of the program. If the value of DIVISOR is greater than 3, execute STEP. Otherwise, display the value of DIVIDEND.

```
b 15 do; if divisor>3 then st;
           else ex dividend; end;
```

See Also

Commands:

"DELETE" on page 2193

"WATCH" on page 2204

CALCULATE

Evaluates a debugger expression and displays the result.

Category: Manipulating DATA Step Variables

Syntax

CALC *expression*

Arguments

expression
specifies any debugger expression.
Restriction: Debugger expressions cannot contain functions.

Details

The CALCULATE command evaluates debugger expressions and displays the result. The result must be numeric.

Examples

- Add 1.1, 1.2, 3.4 and multiply the result by 0.5:

```
calc (1.1+1.2+3.4)*0.5
```

- Calculate the sum of STARTAGE and DURATION:

```
calc startage+duration
```

- Calculate the values of the variable SALE minus the variable DOWNPAY and then multiply the result by the value of the variable RATE. Divide that value by 12 and add 50:

```
calc (((sale-downpay)*rate)/12)+50
```

See Also

"Working with Expressions" on page 2176 for information about debugger expressions

DELETE

Deletes breakpoints or the watch status of variables in the DATA step.

Category: Manipulating Debugging Requests
Alias: D

Syntax

DELETE BREAK *location*

DELETE WATCH *variable(s)* | _ALL_

Arguments

BREAK
deletes breakpoints.
Alias: B

location
specifies a breakpoint location to be deleted. *Location* can have one of these values:

ALL	all current breakpoints in the DATA step.
label	the statement after a statement label.
line-number	the number of a program line.
*	the breakpoint from the current line.

WATCH
deletes watched status of variables.
Alias: W

variable(s)
names one or more watched variables for which the watch status is deleted.

ALL
specifies that the watch status is deleted for all watched variables.

Examples

□ Delete the breakpoint at the statement label

```
eoflabel
:
d b eoflabel
```

□ Delete the watch status from the variable ABC in the current DATA step:

```
d w abc
```

See Also

Commands:

"BREAK" on page 2190
"WATCH" on page 2204

DESCRIBE

Displays the attributes of one or more variables.

Category: Manipulating DATA Step Variables
Alias: DESC

Syntax

DESCRIBE *variable(s)* | _ALL_

Arguments

variable(s)
identifies one or more DATA step variables

ALL
indicates all variables that are defined in the DATA step.

Details

The DESCRIBE command displays the attributes of one or more specified variables.

DESCRIBE reports the name, type, and length of the variable, and, if present, the informat, format, or variable label.

Examples

- Display the attributes of variable ADDRESS:

```
desc address
```

- Display the attributes of array element ARR{*i* + *j*}:

```
desc arr{i+j}
```

ENTER

Assigns one or more debugger commands to the ENTER key.

Category: Tailoring the Debugger

Syntax

ENTER <*command-1* <... ; *command-n*>>

Arguments

command
specifies a debugger command.
Default: STEP 1

Details

The ENTER command assigns one or more debugger commands to the ENTER key. Assigning a new command to the ENTER key replaces the existing command assignment. If you assign more than one command, separate the commands with semicolons.

Examples

- Assign the command STEP 5 to the ENTER key:

```
enter st 5
```

- Assign the commands EXAMINE and DESCRIBE, both for the variable CITY, to the ENTER key:

```
enter ex city; desc city
```

EXAMINE

Displays the value of one or more variables.

Category: Manipulating DATA Step Variables
Alias: E

Syntax

EXAMINE *variable-1 <format-1> <...variable-n <format-n>>*

EXAMINE _ALL_ *<format>*

Arguments

variable
identifies a DATA step variable.

format
identifies a SAS format or a user-created format.

ALL
identifies all variables that are defined in the current DATA step.

Details

The EXAMINE command displays the value of one or more specified variables. The debugger displays the value using the format currently associated with the variable, unless you specify a different format.

Examples

- Display the values of variables N and STR:

```
ex n str
```

- Display the element *i* of the array TESTARR:

```
ex testarr{i}
```

- Display the elements *i*+1, *j**2, and *k*-3 of the array CRR:

```
ex crr{i+1}; ex crr{j*2}; ex crr{k-3}
```

- Display the SAS date variable T_DATE with the DATE7. format:

```
ex t_date date7.
```

- Display the values of all elements in array NEWARR:

```
ex newarr{*}
```

See Also

Command:
"DESCRIBE" on page 2195

GO

Starts or resumes execution of the DATA step.

Category: Controlling Program Execution
Alias: G

Syntax

GO <*line-number* | *label*>

Without Arguments

If you omit arguments, GO resumes execution of the DATA step and executes its statements continuously until a breakpoint is encountered, until the value of a watched variable changes, or until the DATA step completes execution.

Arguments

line-number
gives the number of a program line at which execution is to be suspended next.

label
is a statement label. Execution is suspended at the statement following the statement label.

Details

The GO command starts or resumes execution of the DATA step. Execution continues until all observations have been read, a breakpoint specified in the GO command is reached, or a breakpoint set earlier with a BREAK command is reached.

Examples

- Resume executing the program and execute its statements continuously:

```
g
```

- Resume program execution and then suspend execution at the statement in line 104:

```
g 104
```

See Also

Commands:

"JUMP" on page 2198

"STEP" on page 2202

HELP

Displays information about debugger commands.

Category: Controlling the Windows

Syntax

HELP

Without Arguments

The HELP command displays a directory of the debugger commands. Select a command name to view information about the syntax and usage of that command. You must enter the HELP command from a window command line, from a menu, or with a function key.

JUMP

Restarts execution of a suspended program.

Category: Controlling Program Execution

Alias: J

Syntax

JUMP *line-number* | *label*

Arguments

line-number
indicates the number of a program line at which to restart the suspended program.

label
is a statement label. Execution resumes at the statement following the label.

Details

The JUMP command moves program execution to the specified location without executing intervening statements. After executing JUMP, you must restart execution with GO or STEP. You can jump to any executable statement in the DATA step.

CAUTION:
Do not use the JUMP command to jump to a statement inside a DO loop or to a label that is the target of a LINK-RETURN group. In such cases you bypass the controls set up at the beginning of the loop or in the LINK statement, and unexpected results can appear. △

JUMP is useful in two situations:

- when you want to bypass a section of code that is causing problems in order to concentrate on another section. In this case, use the JUMP command to move to a point in the DATA step after the problematic section.
- when you want to re-execute a series of statements that have caused problems. In this case, use JUMP to move to a point in the DATA step before the problematic statements and use the SET command to reset values of the relevant variables to the values they had at that point. Then re-execute those statements with STEP or GO.

Examples

- Jump to line 5: `j 5`

See Also

Commands:

"GO" on page 2197
"STEP" on page 2202

LIST

Displays all occurrences of the item that is listed in the argument.

Category: Manipulating Debugging Requests
Alias: L

Syntax

LIST _ALL_ | BREAK | DATASETS | FILES | INFILES | WATCH

Arguments

ALL
displays the values of all items.

BREAK
displays breakpoints.

Alias: B

DATASETS
displays all SAS data sets used by the current DATA step.

FILES
displays all external files to which the current DATA step writes.

INFILES
displays all external files from which the current DATA step reads.

WATCH
displays watched variables.

Alias: W

Examples

- List all breakpoints, SAS data sets, external files, and watched variables for the current DATA step:

```
l _all_
```

- List all breakpoints in the current DATA step:

```
l b
```

See Also

Commands:

"BREAK" on page 2190
"DELETE" on page 2193
"WATCH" on page 2204

QUIT

Terminates a debugger session.

Category: Terminating the Debugger

Alias: Q

Syntax

QUIT

Without Arguments

The QUIT command terminates a debugger session and returns control to the SAS session.

Details

SAS creates data sets built by the DATA step that you are debugging. However, when you use QUIT to exit the debugger, SAS does not add the current observation to the data set.

You can use the QUIT command at any time during a debugger session. After you end the debugger session, you must resubmit the DATA step with the DEBUG option to begin a new debugging session; you cannot resume a session after you have ended it.

SET

Assigns a new value to a specified variable.

Category: Manipulating DATA Step Variables

Alias: None

Syntax

SET *variable=expression*

Arguments

variable
specifies the name of a DATA step variable or an array reference.

expression
is any debugger expression.

Tip: *Expression* can contain the variable name that is used on the left side of the equal sign. When a variable appears on both sides of the equal sign, the debugger uses the original value on the right side to evaluate the expression and stores the result in the variable on the left.

Details

The SET command assigns a value to a specified variable. When you detect an error during program execution, you can use this command to assign new values to variables. This enables you to continue the debugging session.

Examples

- Set the variable A to the value of 3:

```
set a=3
```

- Assign to the variable B the value **12345** concatenated with the previous value of B:

```
set b='12345' || b
```

□ Set array element ARR{1} to the result of the expression a+3:

```
set arr{1}=a+3
```

□ Set array element CRR{1,2,3} to the result of the expression crr{1,1,2} + crr{1,1,3}:

```
set crr{1,2,3} = crr{1,1,2} + crr{1,1,3}
```

□ Set the variable A to the result of the expression a+c*3:

```
set a=a+c*3
```

STEP

Executes statements one at a time in the active program.

Category: Controlling Program Execution

Alias: ST

Syntax

STEP <*n*>

Without Arguments

STEP executes one statement.

Arguments

n
specifies the number of statements to execute.

Details

The STEP command executes statements in the DATA step, starting with the statement at which execution was suspended.

When you issue a STEP command, the debugger:

□ executes the number of statements that you specify

□ displays the line number

□ returns control to the user and displays the > prompt.

Note: By default, you can execute the STEP command by pressing the ENTER key.
△

See Also

Commands:

"GO" on page 2197

"JUMP" on page 2198

SWAP

Switches control between the SOURCE window and the LOG window.

Category: Controlling the Windows

Alias: None

Syntax

SWAP

Without Arguments

The SWAP command switches control between the LOG window and the SOURCE window when the debugger is running. When you begin a debugging session, the LOG window becomes active by default. While the DATA step is still being executed, the SWAP command enables you to switch control between the SOURCE and LOG window so that you can scroll and view the text of the program and also continue monitoring the program execution. You must enter the SWAP command from a window command line, from a menu, or with a function key.

TRACE

Controls whether the debugger displays a continuous record of the DATA step execution.

Category: Manipulating Debugging Requests

Alias: T

Default: OFF

Syntax

TRACE <ON | OFF>

Without Arguments

TRACE displays the current status of the TRACE command.

Arguments

ON
prepares for the debugger to display a continuous record of DATA step execution. The next statement that resumes DATA step execution (such as GO) records all actions taken during DATA step execution in the DEBUGGER LOG window.

OFF
stops the display.

Examples

- Determine whether TRACE is ON or OFF:

```
trace
```

- Prepare to display a record of debugger execution:

```
trace on
```

WATCH

Suspends execution when the value of a specified variable changes.

Category: Manipulating Debugging Requests
Alias: W

Syntax

WATCH *variable(s)*

Arguments

variable(s)
specifies one or more DATA step variables.

Details

The WATCH command specifies a variable to monitor and suspends program execution when its value changes.

Each time the value of a watched variable changes, the debugger does the following:

- suspends execution
- displays the line number where execution has been suspended
- displays the variable's old value
- displays the variable's new value
- returns control to the user and displays the > prompt.

Examples

- Monitor the variable DIVISOR for value changes:

```
w divisor
```

APPENDIX

2

Perl Regular Expression (PRX) Metacharacters

Tables of Perl Regular Expression (PRX) Metacharacters

General Constructs

Table A2.1 General Constructs

Metacharacter	Description
( )	indicates grouping.
non-metacharacter	matches a character.
{ } [ ]( )^ $. | * + ? \	to match these characters, override (escape) with \.
\	overrides the next metacharacter.
n	matches capture buffer *n*.
(?:...)	specifies a non-capturing group.

Basic Perl Metacharacters

The following table lists the metacharacters that you can use to match patterns in Perl regular expressions.

Table A2.2 Basic Perl Metacharacters and Their Descriptions

Metacharacter	Description
\a	matches an alarm (bell) character.
\A	matches a character only at the beginning of a string.
\b	matches a word boundary (the position between a word and a space): □ "er\b" matches the "er" in "never" □ "er\b" does not match the "er" in "verb"
\B	matches a non-word boundary: □ "er\B" matches the "er" in "verb" □ "er\B" does not match the "er" in "never"
\cA-\cZ	matches a control character. For example, \cX matches the control character control-X.
\C	matches a single byte.
\d	matches a digit character that is equivalent to [0–9].
\D	matches a non-digit character that is equivalent to [^0–9].
\e	matches an escape character.
\E	specifies the end of case modification.
\f	matches a form feed character.
\l	specifies that the next character is lowercase.
\L	specifies that the next string of characters, up to the \E metacharacter, is lowercase.
\n	matches a newline character.
\num $num	matches capture buffer *num*, where *num* is a positive integer. Perl variable syntax ($num) is valid when referring to capture buffers, but not in other cases.
\Q	escapes (places a backslash before) all non-word characters.
\r	matches a return character.
\s	matches any white space character, including space, tab, form feed, and so on, and is equivalent to [\f\n\r\t\v].
\S	matches any character that is not a white space character and is equivalent to [^\f\n\r\t\v].
\t	matches a tab character.
\u	specifies that the next character is uppercase.
\U	specifies that the next string of characters, up to the \E metacharacter, is uppercase.
\w	matches any word character or alphanumeric character, including the underscore.
\W	matches any non-word character or nonalphanumeric character, and excludes the underscore.
ddd	matches the octal character *ddd*.

Metacharacter	Description
\x*dd*	matches the hexadecimal character *dd*.
\z	matches a character only at the end of a string.
\Z	matches a character only at the end of a string or before newline at the end of a string.

Metacharacters and Replacement Strings

You can use the following metacharacters in both a regular expression and in replacement text, when you use a substitution regular expression:

\l

\u

\L

\E

\U

\Q

These metacharacters are useful in replacement text for controlling the case of capture buffers that are used within replacement text. For an example of how these metacharacters can be used, see “Replacing Text: Example 3” on page 335

For a description of these metacharacters, see Table A2.2 on page 2206.

Other Quantifiers

The following table lists other qualifiers that you can use in Perl regular expressions. The descriptions of the metacharacters in the table include examples of how the metacharacters can be used.

Table A2.3 Other Quantifiers

Metacharacter	Description
\	marks the next character as either a special character, a literal, a back reference, or an octal escape: □ “\n” matches a newline character □ “\\” matches “\” □ “\(“ matches”(“
|	specifies the *or* condition when you compare alphanumeric strings. For example, the construct x|y matches either x or y: □ "z|food" matches either "z" or "food" □ “(z|f)ood” matches “zood” or “food”
^	matches the position at the beginning of the input string.
$	matches the position at the end of the input string.

Metacharacter	Description
period (.)	matches any single character except newline. To match any character including newline, use a pattern such as "[.\n]".
(pattern)	specifies grouping. Matches a pattern and creates a capture buffer for the match. To retrieve the position and length of the match that is captured, use CALL PRXPOSN. To retrieve the value of the capture buffer, use the PRXPOSN function. To match parentheses characters, use "\(" or "\)".

Greedy and Lazy Repetition Factors

Perl regular expressions support “greedy” repetition factors and “lazy” repetition factors. A repetition factor is considered greedy when the repetition factor matches a string as many times as it can when using a specific starting location. A repetition factor is considered lazy when it matches a string the minimum number of times that is needed to satisfy the match. To designate a repetition factor as lazy, add a ? to the end of the repetition factor. By default, repetition factors are considered greedy.

The following table lists the greedy repetition factors. The descriptions of the repetition factors in the table include examples of how they can be used.

Table A2.4 Greedy Repetition Factors

Metacharacter	Description
*	matches the preceding subexpression zero or more times: □ zo* matches "z" and "zoo" □ * is equivalent to {0,}
+	matches the preceding subexpression one or more times: □ "zo+" matches "zo" and "zoo" □ "zo+" does not match "z" □ + is equivalent to {1,}
?	matches the preceding subexpression zero or one time: □ "do(es)?" matches the "do" in "do" or "does" □ ? is equivalent to {0,1}
{n}	matches at least *n* times.
{n,}	matches a pattern at least *n* times.
{n,m}	m and n are non-negative integers, where n<=m. They match at least *n* and at most *m* times: □ "o{1,3}" matches the first three o’s in "fooooood" □ "o{0,1}" is equivalent to "o?" *Note:* You cannot put a space between the comma and the numbers. △

The following table lists the lazy repetition metacharacters.

Table A2.5 Lazy Repetition Factors

Metacharacter	Description
*?	matches a pattern zero or more times.
+?	matches a pattern one or more times.
??	matches a pattern zero or one time.
{n}?	matches exactly *n* times.
{n,}?	matches a pattern at least *n* times.
{n,m}?	matches a pattern at least *n* times but not more than *m* times.

Class Groupings

The following table lists character class groupings. You specify these classes by enclosing characters inside brackets. These metacharacters share a set of common properties. To be successful, the character class must always match a character. The negated character class must always match a character that is not in the list of characters that are designated inside the brackets. The descriptions of the metacharacters in the table include examples of how the metacharacters can be used.

Table A2.6 Character Class Groupings

Metacharacter	Description
[...]	specifies a character set that matches any one of the enclosed characters: □ "[abc]" matches the "a" in "plain"
[^...]	specifies a negative character set that matches any character that is not enclosed: □ "[^abc]" matches the "p" in "plain"
[a-z]	specifies a range of characters that matches any character in the range: □ "[a-z]" matches any lowercase alphabetic character in the range "a" through "z"
[^a-z]	specifies a range of characters that does not match any character in the range: □ "[^a-z]" matches any character that is not in the range "a" through "z"
[[:alpha:]]	matches an alphabetic character.
[[:^alpha:]]	matches a nonalphabetic character.
[[:alnum:]]	matches an alphanumeric character.
[[:^alnum:]]	matches a nonalphanumeric character.
[[:ascii:]]	matches an ASCII character. Equivalent to [\0–\177].
[[:^ascii:]]	matches a non-ASCII character. Equivalent to [^\0–\177].

Metacharacter	Description
[[:blank:]]	matches a blank character.
[[:^blank:]]	matches a non-blank character.
[[:cntrl:]]	matches a control character.
[[:^cntrl:]]	matches a character that is not a control character.
[[:digit:]]	matches a digit. Equivalent to \d.
[[:^digit:]]	matches a non-digit character. Equivalent to \D.
[[:graph:]]	is a visible character, excluding the space character. Equivalent to [[:alnum:][:punct:]].
[[:^graph:]]	is not a visible character. Equivalent to [^[:alnum:][:punct:]].
[[:lower:]]	matches lowercase characters.
[[:^lower:]]	does not match lowercase characters.
[[:print:]]	prints a string of characters.
[[:^print:]]	does not print a string of characters.
[[:punct:]]	matches a punctuation character or a visible character that is not a space or alphanumeric.
[[:^punct:]]	does not match a punctuation character or a visible character that is not a space or alphanumeric.
[[:space:]]	matches a space. Equivalent to \s.
[[:^space:]]	does not match a space. Equivalent to \S.
[[:upper:]]	matches uppercase characters.
[[:^upper:]]	does not match uppercase characters.
[[:word:]]	matches a word. Equivalent to \w.
[[:^word:]]	does not match a word. Equivalent to \W.
[[:xdigit:]]	matches a hexadecimal character.
[[:^xdigit:]]	does not match a hexadecimal character.

Look-Ahead and Look-Behind Behavior

Look-ahead and look-behind are ways to look ahead or behind a match to see whether a particular text occurs. The text that is found with look-ahead or look-behind is not included in the match that is found. For example, if you want to find names that end with “Jr.”, but you do not want “Jr.” to be part of the match, you could use the regular expression /.*(?=Jr\.)/. For the value "John Wainright Jr.", the regular expression will find "John Wainright" as a match because it is followed by "Jr."

Table A2.7 Look-Ahead and Look-Behind Behavior

Metacharacter	Description
(?=...)	specifies a zero-width, positive, look-ahead assertion. For example, in the expression *regex1 (?=regex2)*, a match is found if both *regex1* and *regex2* match. *regex2* is not included in the final match.
(?!...)	specifies a zero-width, negative, look-ahead assertion. For example, in the expression *regex1 (?!regex2)*, a match is found if *regex1* matches and *regex2* does not match. *regex2* is not included in the final match.
(?<=...)	specifies a zero-width, positive, look-behind assertion. For example, in the expression *(?<=regex1) regex2*, a match is found if both *regex1* and *regex2* match. *regex1* is not included in the final match. Works with fixed-width look-behind only.
(?<!...)	specifies a zero-width, negative, look-behind assertion. Works with fixed-width look-behind only.

Comments and Inline Modifiers

The metacharacters in this table contain a question mark as the first element inside the parentheses. The characters after the question mark indicate the extension.

Table A2.8 Comments and Inline Modifiers

Metacharacter	Description
(?#text)	specifies a comment in which the text is ignored.
(?imsx)	specifies one or more embedded pattern-matching modifiers. If the pattern is case insensitive, you can use (?i) at the front of the pattern. An example is `$pattern="(?i)foobar";`. Letters that appear after a hyphen (-) turn the modifiers off.

Selecting the Best Condition by Using Combining Operators

The elementary regular expressions (for example, \a and \w) that are described in the preceding tables can match at most one substring at the given position in the input string. However, operators that perform combining in typical regular expressions combine elementary metacharacters to create more complex patterns. In an ambiguous situation, these operators (see Table A1.9) can determine the best match or the worst match. The match that is the best is always chosen.

Table A2.9 Best Match Using Combining Operators

Metacharacter	Description
ST	in the following example, specifies that AB and A'B', and A and A' are substrings that can be matched by S, and that B and B' are substrings that can be matched by T: □ If A is a better match for S than A', then AB is a better match than A'B'. □ If A and A' coincide, then AB is a better match than AB' if B is a better match for T than B'.
S|T	specifies that when S can match, it is a better match than when only T can match. The ordering of two matches for S is the same as for S. Similarly, the ordering of two matches for T is the same as for T.
S{repeat-count}	matches as SSS . . . S (repeated as many times as necessary).
S{min,max}	matches as S{max}|S{max-1}| . . . |S{min+1}|S{min}.
S{min,max}?	matches as S{min}|S{min+1}| . . . |S{max-1}|S{max}.
S?, S*, S+	same as S{0,1}, S{0, big-number}, S{1,big-number}, respectively.
S??, S*?, S+	same as S{0,1}?, S{0, big-number}?, S{1,big-number}?, respectively.
(?=S), (?<=S)	considers the best match for S. (This is important only if S has capturing parentheses, and back references are used elsewhere in the whole regular expression.)
(?!S), (?<!S)	unnecessary to describe the ordering for this grouping operator because only whether S can match is important.

APPENDIX

3

SAS Utility Macro

%DS2CSV Macro

Converts SAS data sets to comma-separated value (CSV) files.

Restriction: Must be in open code. Cannot be used in a DATA step.

Syntax

%DS2CSV(*argument=value*, *argument=value*,...)

Arguments That Affect Input/Output

csvfile=*external-filename*
specifies the name of the CSV file where the formatted output is to be written. If the file that you specify does not exist, then it is created for you.

Note: Do not use the CSVFILE argument if you use the CSVFREF argument. △

csvfref=*fileref*
specifies the SAS fileref that points to the location of the CSV file where the formatted output is to be written. If the file that you specify does not exist, then it is created for you.

Note: Do not use the CSVFREF argument if you use the CSVFILE argument. △

openmode=REPLACE|APPEND
indicates whether the new CSV output overwrites the information that is currently in the specified file or if the new output is appended to the end of the existing file. The default value is REPLACE. If you do not want to replace the current contents, then specify OPENMODE=APPEND to add your new CSV-formatted output to the end of an existing file.

Note: OPENMODE=APPEND is not valid if you are writing your resulting output to a partitioned data set (PDS) on z/OS. △

Arguments That Affect MIME/HTTP Headers

For more information about MIME and HTTP headers, refer to the Internet Request for Comments (RFC) documents RFC 1521 (http://asg.web.cmu.edu/rfc/rfc1521.html) and RFC 1945 (http://asg.web.cmu.edu/rfc/rfc1945.html), respectively.

conttype=Y | N
indicates whether to write a content type header. This header is written by default.

Restriction: This argument is valid only when RUNMODE=S.

contdisp=Y | N
indicates whether to write a content disposition header. This header is written by default.

Note: If you specify CONTDISP=N, then the SAVEFILE argument is ignored. △

Restriction: This argument is valid only when RUNMODE=S.

mimehdr1=*MIME/HTTP-header*
specifies the text that is to be used for the first MIME or HTTP header that is written. This header is written after the content type and disposition headers. By default, nothing is written for this header.

Restriction: This argument is valid only when RUNMODE=S.

mimehdr2=*MIME/HTTP-header*
specifies the text that is to be used for the second MIME or HTTP header that is written. This header is written after the content type and disposition headers. By default, nothing is written for this header.

Restriction: This argument is valid only when RUNMODE=S.

mimehdr3=*MIME/HTTP-header*
specifies the text that is to be used for the third MIME or HTTP header that is written when RUNMODE=S is specified. This header is written after the content type and disposition headers. By default, nothing is written for this header.

Restriction: This argument is valid only when RUNMODE=S.

mimehdr4=*MIME/HTTP-header*
specifies the text that is to be used for the fourth MIME or HTTP header that is written. This header is written after the content type and disposition headers. By default, nothing is written for this header.

Restriction: This argument is valid only when RUNMODE=S.

mimehdr5=*MIME/HTTP-header*
specifies the text that is to be used for the fifth MIME or HTTP header that is written. This header is written after the content type and disposition headers. By default, nothing is written for this header.

runmode=S | B
specifies whether you are running the %DS2CSV macro in batch or server mode. The default setting for this argument is RUNMODE=S.

- *Server mode* (RUNMODE=S) is used with Application Dispatcher programs and streaming output stored processes. Server mode causes DS2CSV to generate appropriate MIME or HTTP headers. For more information about Application Dispatcher, refer to the Application Dispatcher documentation at **`http://support.sas.com/rnd/web/intrnet/dispatch.html`**.
- *Batch mode* (RUNMODE=B) means that you are submitting the DS2CSV macro in the SAS Program Editor or that you included it in a SAS program.

Note: No HTTP headers are written when you specify batch mode. △

Restriction: RUNMODE=S is valid only when used within the SAS/IntrNet and Stored Process servers.

savefile=*filename*
specifies the filename to display in the Web browser's **`Save As`** dialog box. The default value is the name of the data set plus ".csv".

Note: This argument is ignored if CONTDISP=N is specified. △

Restriction: This argument is valid only when RUNMODE=S.

Arguments That Affect CSV Creation

colhead=Y | N
indicates whether to include column headings in the CSV file. The column headings that are used depend on the setting of the LABELS argument. By default, column headings are included as the first record of the CSV file.

data=*SAS-data-set-name*
specifies the SAS data set that contains the data that you want to convert into a CSV file. This argument is required. However, if you omit the data set name, DS2CSV attempts to use the most recently created SAS data set.

formats=Y | N
indicates whether to apply the data set's defined variable formats to the values in the CSV file. By default, all formats are applied to values before they are added to the CSV file. The formats must be stored in the data set in order for them to be applied.

labels=Y | N
indicates whether to use the SAS variable labels that are defined in the data set as your column headings. The DS2CSV macro uses the variable labels by default. If a variable does not have a SAS label, then use the name of the variable. Specify labels=N to use variable names instead of the SAS labels as your column headings. See colhead on page 2215 argument for more information about column headings.

pw=*password*
specifies the password that is needed to access a password-protected data set. This argument is required if the data set has a READ or PW password. (You do not need to specify this argument if the data set has only WRITE or ALTER passwords.)

sepchar=*separator-character*
specifies the character that is used for the separator character. Specify the two-character hexadecimal code for the character or omit this argument to get the default setting. The default settings are 2C for ASCII systems and 6B for EBCDIC systems. (These settings represent commas (,) on their respective systems.)

var=*var1 var2 ...*
specifies the variables that are to be included in the CSV file and the order in which they should be included. To include all of the variables in the data set, do not specify this argument. If you want to include only a subset of the variables, then list each variable name and use single blank spaces to separate the variables. Do not use a comma in the list of variable names.

where=*where-expression*
specifies a valid WHERE clause that selects observations from the SAS data set. Using this argument subsets your data based on the criteria you supply for *where-expression*.

Details

The DS2CSV macro converts SAS data sets to comma-separated value (CSV) files. You can specify the hexadecimal code for the separator character if you want to create some other type of output file (for example, a tab-separated value file).

Example

The following example uses the %DS2CSV macro to convert the SASHELP.RETAIL data set to a comma-separated value file:

```
%ds2csv (data=sashelp.retail, runmode=b, csvfile=c:\temp\retail.csv);
```

APPENDIX

4

Recommended Reading

Recommended Reading

Here is the recommended reading list for this title:

- *An Array of Challenges—Test Your SAS Skills*
- *Base SAS Glossary*
- *Base SAS Procedures Guide*
- *Cody's Data Cleaning Techniques Using SAS Software*
- *Combining and Modifying SAS Data Sets: Examples*
- *Debugging SAS Programs: A Handbook of Tools and Techniques*
- *Health Care Data and SAS*
- *The Little SAS Book: A Primer*
- *Output Delivery System: The Basics*
- *Quick Results with the Output Delivery System*
- *SAS Companion for OpenVMS on HP Integrity Servers*
- *SAS Companion for UNIX Environments*
- *SAS Companion for Windows*
- *SAS Companion for z/OS*
- *SAS Guide to Report Writing: Examples*
- *SAS Language Reference: Concepts*
- *SAS Metadata LIBNAME Engine: User's Guide*
- *SAS National Language Support (NLS): Reference Guide*
- *SAS Output Delivery System: User's Guide*
- *SAS Programming by Example*
- *SAS Scalable Performance Data Engine: Reference*
- *The SAS Workbook*
- *SAS XML LIBNAME Engine: User's Guide*
- *Step-by-Step Programming with Base SAS Software*
- *Using the SAS Windowing Environment: A Quick Tutorial*

For a complete list of SAS publications, go to **support.sas.com/bookstore**. If you have questions about which titles you need, please contact a SAS Publishing Sales Representative at:

SAS Publishing Sales
SAS Campus Drive
Cary, NC 27513
Telephone: 1-800-727-3228
Fax: 1-919-531-9439
E-mail: **sasbook@sas.com**
Web address: **support.sas.com/bookstore**

Customers outside the United States and Canada, please contact your local SAS office for assistance.

Index

Numbers

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z

Your Turn

We welcome your feedback.

- If you have comments about this book, please send them to **yourturn@sas.com**. Include the full title and page numbers (if applicable).
- If you have comments about the software, please send them to **suggest@sas.com**.

SAS® Publishing Delivers!

SAS Publishing provides you with a wide range of resources to help you develop your SAS software expertise.
Visit us online at **support.sas.com/bookstore**.

SAS® PRESS

SAS Press titles deliver expert advice from SAS® users worldwide. Written by experienced SAS professionals, SAS Press books deliver real-world insights on a broad range of topics for all skill levels.

support.sas.com/saspress

SAS® DOCUMENTATION

We produce a full range of primary documentation:

- Online help built into the software
- Tutorials integrated into the product
- Reference documentation delivered in HTML and PDF formats—free on the Web
- Hard-copy books

support.sas.com/documentation

SAS® PUBLISHING NEWS

Subscribe to SAS Publishing News to receive up-to-date information via e-mail about all new SAS titles, product news, special offers and promotions, and Web site features.

support.sas.com/spn

SOCIAL MEDIA: JOIN THE CONVERSATION!

Connect with SAS Publishing through social media. Visit our Web site for links to our pages on Facebook, Twitter, and LinkedIn. Learn about our blogs, author podcasts, and RSS feeds, too.

support.sas.com/socialmedia

sas® THE POWER TO KNOW®