Choosing the Right Procedure

Functional Categories of Base SAS Procedures

Report Writing
These procedures display useful information, such as data listings (detail reports), summary reports, calendars, letters, labels, multipanel reports, and graphical reports:

- CALENDAR
- CHART*
- FREQ*
- MEANS*
- PLOT
- PRINT
- REPORT*
- SQL*
- SUMMARY*
- TABULATE*
- TIMEPLOT

* These procedures produce reports and compute statistics.

Statistics
These procedures compute elementary statistical measures that include descriptive statistics based on moments, quantiles, confidence intervals, frequency counts,
Utilities

These procedures perform basic utility operations. They create, edit, sort, and transpose data sets, create and restore transport data sets, create user-defined formats, and provide basic file maintenance such as to copy, append, and compare data sets:

- APPEND
- EXPORT
- PWENCODE
- BMDP
- FONTREG
- PRTEXP
- CATALOG
- FORMAT
- REGISTRY
- CIMPORT
- FSLIST
- RELEASE
- COMPARE
- IMPORT
- SORT
- CONTENTS
- OPTIONS
- SOURCE
- CONVERT
- OPTLOAD
- SQL
- COPY
- OPTSAVE
- TAPECOPY
- CPORT
- PDS
- TAPELABEL
- CV2VIEW
- PDSCOPY
- TEMPLATE
- DATASETS
- PMENU
- TRANSPOSE
- DBCSTAB
- PRINTTO
- TRANTAB
- DOCUMENT
- PRTDEF

See the SAS documentation for your operating environment for a description of these procedures.


See SAS/ACCESS for Relational Databases: Reference for a description of this procedure.


Report-Writing Procedures

Table 1.1 on page 5 lists report-writing procedures according to the type of report.
Table 1.1  Report-Writing Procedures by Task

<table>
<thead>
<tr>
<th>To produce...</th>
<th>Use this procedure...</th>
<th>Which...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detail reports</td>
<td>PRINT</td>
<td>produces data listings quickly; can supply titles, footnotes, and column sums.</td>
</tr>
<tr>
<td></td>
<td>REPORT</td>
<td>offers more control and customization than PROC PRINT; can produce both column and row sums; has DATA step computation abilities.</td>
</tr>
<tr>
<td></td>
<td>SQL</td>
<td>combines Structured Query Language and SAS features such as formats; can manipulate data and create a SAS data set in the same step that creates the report; can produce column and row statistics; does not offer as much control over output as PROC PRINT and PROC REPORT.</td>
</tr>
<tr>
<td>Summary reports</td>
<td>MEANS or SUMMARY</td>
<td>computes descriptive statistics for numeric variables; can produce a printed report and create an output data set.</td>
</tr>
<tr>
<td></td>
<td>PRINT</td>
<td>produces only one summary report; can sum the BY variables.</td>
</tr>
<tr>
<td></td>
<td>REPORT</td>
<td>combines features of the PRINT, MEANS, and TABULATE procedures with features of the DATA step in a single report writing tool that can produce a variety of reports; can also create an output data set.</td>
</tr>
<tr>
<td></td>
<td>SQL</td>
<td>computes descriptive statistics for one or more SAS data sets or DBMS tables; can produce a printed report or create a SAS data set.</td>
</tr>
<tr>
<td></td>
<td>TABULATE</td>
<td>produces descriptive statistics in a tabular format; can produce stub-and-banner reports (multidimensional tables with descriptive statistics); can also create an output data set.</td>
</tr>
<tr>
<td>Miscellaneous highly formatted reports</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calendars</td>
<td>CALENDAR</td>
<td>produces schedule and summary calendars; can schedule tasks around nonwork periods and holidays, weekly work schedules, and daily work shifts.</td>
</tr>
<tr>
<td>Multipanel reports (telephone book listings)</td>
<td>REPORT</td>
<td>produces multipanel reports.</td>
</tr>
<tr>
<td>Low-resolution graphical reports*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHART</td>
<td></td>
<td>produces bar charts, histograms, block charts, pie charts, and star charts that display frequencies and other statistics.</td>
</tr>
<tr>
<td>PLOT</td>
<td></td>
<td>produces scatter diagrams that plot one variable against another.</td>
</tr>
<tr>
<td>TIMEPLOT</td>
<td></td>
<td>produces plots of one or more variables over time intervals.</td>
</tr>
</tbody>
</table>

* These reports quickly produce a simple graphical picture of the data. To produce high-resolution graphical reports, use SAS/GRAPH software.
Available Statistical Procedures

Table 1.2 on page 6 lists statistical procedures according to task. Table A1.1 on page 1341 lists the most common statistics and the procedures that compute them.

<table>
<thead>
<tr>
<th>To produce...</th>
<th>Use this procedure...</th>
<th>Which...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive statistics</td>
<td>CORR</td>
<td>computes simple descriptive statistics.</td>
</tr>
<tr>
<td></td>
<td>MEANS or SUMMARY</td>
<td>computes descriptive statistics; can produce printed output and output data sets. By default, PROC MEANS produces printed output and PROC SUMMARY creates an output data set.</td>
</tr>
<tr>
<td></td>
<td>REPORT</td>
<td>computes most of the same statistics as PROC TABULATE; allows customization of format.</td>
</tr>
<tr>
<td></td>
<td>SQL</td>
<td>computes descriptive statistics for data in one or more DBMS tables; can produce a printed report or create a SAS data set.</td>
</tr>
<tr>
<td></td>
<td>TABULATE</td>
<td>produces tabular reports for descriptive statistics; can create an output data set.</td>
</tr>
<tr>
<td></td>
<td>UNIVARIATE</td>
<td>computes the broadest set of descriptive statistics; can create an output data set.</td>
</tr>
<tr>
<td>Frequency and cross-tabulation tables</td>
<td>FREQ</td>
<td>produces one-way to n-way tables; reports frequency counts; computes chi-square tests; computes tests and measures of association and agreement for two-way to n-way cross-tabulation tables; can compute exact tests and asymptotic tests; can create output data sets.</td>
</tr>
<tr>
<td></td>
<td>TABULATE</td>
<td>produces one-way and two-way cross-tabulation tables; can create an output data set.</td>
</tr>
<tr>
<td></td>
<td>UNIVARIATE</td>
<td>produces one-way frequency tables.</td>
</tr>
<tr>
<td>Correlation analysis</td>
<td>CORR</td>
<td>computes Pearson’s, Spearman’s, and Kendall's correlations and partial correlations; also computes Hoeffding’s D and Cronbach’s coefficient alpha.</td>
</tr>
<tr>
<td>Distribution analysis</td>
<td>UNIVARIATE</td>
<td>computes tests for location and tests for normality.</td>
</tr>
<tr>
<td></td>
<td>FREQ</td>
<td>computes a test for the binomial proportion for one-way tables; computes a goodness-of-fit test for one-way tables; computes a chi-square test of equal distribution for two-way tables.</td>
</tr>
<tr>
<td>Robust estimation</td>
<td>UNIVARIATE</td>
<td>computes robust estimates of scale, trimmed means, and Winsorized means.</td>
</tr>
<tr>
<td>Data transformation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computing ranks</td>
<td>RANK</td>
<td>computes ranks for one or more numeric variables across the observations of a SAS data set and creates an output data set; can produce normal scores or other rank scores.</td>
</tr>
</tbody>
</table>
Choosing the Right Procedure

### Utility Procedures

<table>
<thead>
<tr>
<th>To produce...</th>
<th>Use this procedure...</th>
<th>Which...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardizing data</td>
<td>STANDARD</td>
<td>creates an output data set that contains variables that are standardized to a given mean and standard deviation.</td>
</tr>
<tr>
<td>Low-resolution graphics</td>
<td>CHART</td>
<td>produces a graphical report that can show one of the following statistics for the chart variable: frequency counts, percentages, cumulative frequencies, cumulative percentages, totals, or averages.</td>
</tr>
<tr>
<td></td>
<td>UNIVARIATE</td>
<td>produces descriptive plots such as stem and leaf, box plot, and normal probability plot.</td>
</tr>
</tbody>
</table>

* To produce high-resolution graphical reports, use SAS/GRA F software.

### Efficiency Issues

#### Quantiles

For a large sample size $n$, the calculation of quantiles, including the median, requires computing time proportional to $n \log(n)$. Therefore, a procedure, such as UNIVARIATE, that automatically calculates quantiles may require more time than other data summarization procedures. Furthermore, because data is held in memory, the procedure also requires more storage space to perform the computations. By default, the report procedures PROC MEANS, PROC SUMMARY, and PROC TABULATE require less memory because they do not automatically compute quantiles. These procedures also provide an option to use a new fixed-memory quantiles estimation method that is usually less memory intense. See “Quantiles” on page 555 for more information.

#### Computing Statistics for Groups of Observations

To compute statistics for several groups of observations, you can use any of the previous procedures with a BY statement to specify BY-group variables. However, BY-group processing requires that you previously sort or index the data set, which for very large data sets may require substantial computer resources. A more efficient way to compute statistics within groups without sorting is to use a CLASS statement with one of the following procedures: MEANS, SUMMARY, or TABULATE.

### Additional Information about the Statistical Procedures

Appendix 1, “SAS Elementary Statistics Procedures,” on page 1339 lists standard keywords, statistical notation, and formulas for the statistics that base SAS procedures compute frequently. The individual statistical procedures discuss the statistical concepts that are useful to interpret the output of a procedure.

### Utility Procedures

Table 1.3 on page 8 groups utility procedures according to task.
## Table 1.3  Utility Procedures by Task

<table>
<thead>
<tr>
<th>To perform these utility tasks...</th>
<th>Use this procedure...</th>
<th>Which...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply information</td>
<td>COMPARE</td>
<td>compares the contents of two SAS data sets.</td>
</tr>
<tr>
<td></td>
<td>CONTENTS</td>
<td>describes the contents of a SAS data library or specific library members.</td>
</tr>
<tr>
<td></td>
<td>OPTIONS</td>
<td>lists the current values of all SAS system options.</td>
</tr>
<tr>
<td></td>
<td>SQL</td>
<td>supplies information through dictionary tables on an individual SAS data set as well as all SAS files active in the current SAS session. Dictionary tables can also provide information about macros, titles, indexes, external files, or SAS system options.</td>
</tr>
<tr>
<td>Manage SAS system options</td>
<td>OPTIONS</td>
<td>lists the current values of all SAS system options.</td>
</tr>
<tr>
<td></td>
<td>OPTLOAD</td>
<td>reads SAS system option settings that are stored in the SAS registry or a SAS data set.</td>
</tr>
<tr>
<td></td>
<td>OPTSAVE</td>
<td>saves SAS system option settings to the SAS registry or a SAS data set.</td>
</tr>
<tr>
<td>Affect printing and Output</td>
<td>DOCUMENT$^*$</td>
<td>manipulates procedure output that is stored in ODS documents.</td>
</tr>
<tr>
<td>Delivery System output</td>
<td>FONTREG</td>
<td>adds system fonts to the SAS registry.</td>
</tr>
<tr>
<td></td>
<td>FORMAT</td>
<td>creates user-defined formats to display and print data.</td>
</tr>
<tr>
<td></td>
<td>PRINTTO</td>
<td>routes procedure output to a file, a SAS catalog entry, or a printer; can also redirect the SAS log to a file.</td>
</tr>
<tr>
<td></td>
<td>PRTDEF</td>
<td>creates printer definitions.</td>
</tr>
<tr>
<td></td>
<td>PRTEXP</td>
<td>exports printer definition attributes to a SAS data set.</td>
</tr>
<tr>
<td></td>
<td>TEMPLATE$^*$</td>
<td>customizes ODS output.</td>
</tr>
<tr>
<td>Create, browse, and edit data</td>
<td>FSLIST</td>
<td>browses external files such as files that contain SAS source lines or SAS procedure output.</td>
</tr>
<tr>
<td></td>
<td>SQL</td>
<td>creates SAS data sets using Structured Query Language and SAS features.</td>
</tr>
<tr>
<td>Transform data</td>
<td>DBCSTAB$^*$</td>
<td>produces conversion tables for the double-byte character sets that SAS supports.</td>
</tr>
<tr>
<td></td>
<td>FORMAT</td>
<td>creates user-defined informats to read data and user-defined formats to display data.</td>
</tr>
<tr>
<td></td>
<td>SORT</td>
<td>sorts SAS data sets by one or more variables.</td>
</tr>
<tr>
<td></td>
<td>SQL</td>
<td>sorts SAS data sets by one or more variables.</td>
</tr>
<tr>
<td></td>
<td>TRANSPOSE</td>
<td>transforms SAS data sets so that observations become variables and variables become observations.</td>
</tr>
<tr>
<td></td>
<td>TRANTTAB$^*$</td>
<td>creates, edits, and displays customized translation tables.</td>
</tr>
<tr>
<td>Manage SAS files</td>
<td>APPEND</td>
<td>appends one SAS data set to the end of another.</td>
</tr>
<tr>
<td></td>
<td>BMDP$^*$</td>
<td>invokes a BMDP program to analyze data in a SAS data set.</td>
</tr>
<tr>
<td>To perform these utility tasks...</td>
<td>Use this procedure...</td>
<td>Which...</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------------------</td>
<td>----------</td>
</tr>
<tr>
<td>CATALOG</td>
<td>manages SAS catalog entries.</td>
<td></td>
</tr>
<tr>
<td>CIMPORT</td>
<td>restores a transport sequential file that PROC CPORT creates (usually under another operating environment) to its original form as a SAS catalog, a SAS data set, or a SAS library.</td>
<td></td>
</tr>
<tr>
<td>CONVERT</td>
<td>converts BMDP system files, OSIRIS system files, and SPSS portable files to SAS data sets.</td>
<td></td>
</tr>
<tr>
<td>COPY</td>
<td>copies a SAS data library or specific members of the library.</td>
<td></td>
</tr>
<tr>
<td>CPORT</td>
<td>converts a SAS catalog, a SAS data set, or a SAS library to a transport sequential file that PROC CIMPORT can restore (usually under another operating environment) to its original form.</td>
<td></td>
</tr>
<tr>
<td>CV2VIEW</td>
<td>converts SAS/ACCESS view descriptors to PROC SQL views.</td>
<td></td>
</tr>
<tr>
<td>DATASETS</td>
<td>manages SAS files.</td>
<td></td>
</tr>
<tr>
<td>EXPORT</td>
<td>reads data from a SAS data set and writes them to an external data source.</td>
<td></td>
</tr>
<tr>
<td>IMPORT</td>
<td>reads data from an external data source and writes them to a SAS data set.</td>
<td></td>
</tr>
<tr>
<td>PDS</td>
<td>lists, deletes, and renames the members of a partitioned data set.</td>
<td></td>
</tr>
<tr>
<td>PDSCOPY</td>
<td>copies partitioned data sets from disk to tape, disk to disk, tape to tape, or tape to disk.</td>
<td></td>
</tr>
<tr>
<td>REGISTRY</td>
<td>imports registry information to the USER portion of the SAS registry.</td>
<td></td>
</tr>
<tr>
<td>RELEASE</td>
<td>releases unused space at the end of a disk data set under the z/OS environment.</td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>provides an easy way to back up and process source library data sets.</td>
<td></td>
</tr>
<tr>
<td>SQL</td>
<td>concatenates SAS data sets.</td>
<td></td>
</tr>
<tr>
<td>TAPECOPY</td>
<td>copies an entire tape volume or files from one or more tape volumes to one output tape volume.</td>
<td></td>
</tr>
<tr>
<td>TAPELABEL</td>
<td>lists the label information of an IBM standard-labeled tape volume under the z/OS environment.</td>
<td></td>
</tr>
</tbody>
</table>

Control windows PMENU creates customized pull-down menus for SAS applications.

Miscellaneous PWENCODE encodes passwords for use in SAS programs.

* See the SAS documentation for your operating environment for a description of these procedures.
@ See SAS/ACCESS for Relational Databases: Reference for a description of this procedure.
APPEND procedure
adds observations from one SAS data set to the end of another SAS data set.

BMDP procedure
invokes a BMDP program to analyze data in a SAS data set. See the SAS documentation for your operating environment for more information.

CALENDAR procedure
displays data from a SAS data set in a monthly calendar format. PROC CALENDAR can display holidays in the month, schedule tasks, and process data for multiple calendars with work schedules that vary.

CATALOG procedure
manages entries in SAS catalogs. PROC CATALOG is an interactive, nonwindowing procedure that enables you to display the contents of a catalog, copy an entire catalog or specific entries in a catalog, and rename, exchange, or delete entries in a catalog.

CHART procedure
produces vertical and horizontal bar charts, block charts, pie charts, and star charts. These charts provide a quick visual representation of the values of a single variable or several variables. PROC CHART can also display a statistic associated with the values.

CIMPORT procedure
restores a transport file created by the CPORT procedure to its original form (a SAS data library, catalog, or data set) in the format appropriate to the operating environment. Coupled with the CPORT procedure, PROC CIMPORT enables you to move SAS data libraries, catalogs, and data sets from one operating environment to another.

COMPARE procedure
compares the contents of two SAS data sets. You can also use PROC COMPARE to compare the values of different variables within a single data set. PROC COMPARE produces a variety of reports on the comparisons that it performs.

CONTENTS procedure
prints descriptions of the contents of one or more files in a SAS data library.

CONVERT procedure
converts BMDP system files, OSIRIS system files, and SPSS portable files to SAS data sets. See the SAS documentation for your operating environment for more information.

COPY procedure
copies an entire SAS data library or specific members of the library. You can limit processing to specific types of library members.

CORR procedure
computes Pearson product-moment and weighted product-moment correlation coefficients between variables and descriptive statistics for these variables. In addition, PROC CORR can compute three nonparametric measures of association (Spearman’s rank-order correlation, Kendall’s tau-b, and Hoeffding’s measure of dependence, D), partial correlations (Pearson’s partial correlation, Spearman’s partial rank-order correlation, and Kendall’s partial tau-b), and Cronbach’s coefficient alpha.
CPORR procedure
writes SAS data libraries, data sets, and catalogs in a special format called a transport file. Coupled with the CIMPORT procedure, PROC CPORR enables you to move SAS libraries, data sets, and catalogs from one operating environment to another.

CV2VIEW procedure
converts SAS/ACCESS view descriptors to PROC SQL views. Starting in SAS System 9, conversion of SAS/ACCESS view descriptors to PROC SQL views is recommended because PROC SQL views are platform independent and enable you to use the LIBNAME statement. See SAS/ACCESS for Relational Databases: Reference for details.

DATASETS procedure
lists, copies, renames, and deletes SAS files and SAS generation groups, manages indexes, and appends SAS data sets in a SAS data library. The procedure provides all the capabilities of the APPEND, CONTENTS, and COPY procedures. You can also modify variables within data sets, manage data set attributes, such as labels and passwords, or create and delete integrity constraints.

DBCSTAB procedure
produces conversion tables for the double-byte character sets that SAS supports.

DOCUMENT procedure
manipulates procedure output that is stored in ODS documents. PROC DOCUMENT enables a user to browse and edit output objects and hierarchies, and to replay them to any supported ODS output format. See SAS Output Delivery System: User’s Guide for details.

EXPORT procedure
reads data from a SAS data set and writes it to an external data source.

FONTREG procedure
adds system fonts to the SAS registry.

FORMAT procedure
creates user-defined informats and formats for character or numeric variables. PROC FORMAT also prints the contents of a format library, creates a control data set to write other informats or formats, and reads a control data set to create informats or formats.

FREQ procedure
produces one-way to n-way frequency tables and reports frequency counts. PROC FREQ can compute chi-square tests for one-way to n-way tables, tests and measures of association and of agreement for two-way to n-way cross-tabulation tables, risks and risk difference for 2x2 tables, trends tests, and Cochran-Mantel-Haenszel statistics. You can also create output data sets.

FSLIST procedure
displays the contents of an external file or copies text from an external file to the SAS Text Editor.

IMPORT procedure
reads data from an external data source and writes them to a SAS data set.

MEANS procedure
computes descriptive statistics for numeric variables across all observations and within groups of observations. You can also create an output data set that contains specific statistics and identifies minimum and maximum values for groups of observations.
OPTIONS procedure
   lists the current values of all SAS system options.

OPTLOAD procedure
   reads SAS system option settings from the SAS registry or a SAS data set, and
   puts them into effect.

OPTSAVE procedure
   saves SAS system option settings to the SAS registry or a SAS data set.

PDS procedure
   lists, deletes, and renames the members of a partitioned data set. See the SAS
documentation for your operating environment for more information.

PDSCOPY procedure
   copies partitioned data sets from disk to tape, disk to disk, tape to tape, or tape to
disk. See the SAS documentation for your operating environment for more
information.

PLOT procedure
   produces scatter plots that graph one variable against another. The coordinates of
each point on the plot correspond to the two variables' values in one or more
observations of the input data set.

PMENU procedure
   defines menus that you can use in DATA step windows, macro windows, and
SAS/AF windows, or in any SAS application that enables you to specify customized
menus.

PRINT procedure
   prints the observations in a SAS data set, using all or some of the variables.
PROC PRINT can also print totals and subtotals for numeric variables.

PRINTTO procedure
   defines destinations for SAS procedure output and the SAS log.

PRTDEF procedure
   creates printer definitions for individual SAS users or all SAS users.

PRTEXP procedure
   exports printer definition attributes to a SAS data set so that they can be easily
replicated and modified.

PWENCOD procedure
   encodes passwords for use in SAS programs.

RANK procedure
   computes ranks for one or more numeric variables across the observations of a
SAS data set. The ranks are written to a new SAS data set. Alternatively, PROC
RANK produces normal scores or other rank scores.

REGISTRY procedure
   imports registry information into the USER portion of the SAS registry.

RELEASE procedure
   releases unused space at the end of a disk data set in the z/OS environment. See
the SAS documentation for this operating environment for more information.

REPORT procedure
   combines features of the PRINT, MEANS, and TABULATE procedures with
features of the DATA step in a single report-writing tool that can produce both
detail and summary reports.
SORT procedure
sorts observations in a SAS data set by one or more variables. PROC SORT stores the resulting sorted observations in a new SAS data set or replaces the original data set.

SOURCE procedure
provides an easy way to back up and process source library data sets. See the SAS documentation for your operating environment for more information.

SQL procedure
implements a subset of the Structured Query Language (SQL) for use in SAS. SQL is a standardized, widely used language that retrieves and updates data in SAS data sets, SQL views, and DBMS tables, as well as views based on those tables. PROC SQL can also create tables and views, summaries, statistics, and reports and perform utility functions such as sorting and concatenating.

STANDARD procedure
standardizes some or all of the variables in a SAS data set to a given mean and standard deviation and produces a new SAS data set that contains the standardized values.

SUMMARY procedure
computes descriptive statistics for the variables in a SAS data across all observations and within groups of observations and outputs the results to a new SAS data set.

TABULATE procedure
displays descriptive statistics in tabular form. The value in each table cell is calculated from the variables and statistics that define the pages, rows, and columns of the table. The statistic associated with each cell is calculated on values from all observations in that category. You can write the results to a SAS data set.

TAPECOPY procedure
copies an entire tape volume or files from one or more tape volumes to one output tape volume. See the SAS documentation for your operating environment for more information.

TAPELABEL procedure
lists the label information of an IBM standard-labeled tape volume under the z/OS environment. See the SAS documentation for this operating environment for more information.

TEMPLATE procedure

TIMEPLOT procedure
produces plots of one or more variables over time intervals.

TRANSPOSE procedure
transposes a data set that changes observations into variables and vice versa.

TRANTAB procedure
creates, edits, and displays customized translation tables.

UNIVARIATE procedure
computes descriptive statistics (including quantiles), confidence intervals, and robust estimates for numeric variables. Provides detail on the distribution of numeric variables, which include tests for normality, plots to illustrate the distribution, frequency tables, and tests of location.