

Paper 118-29

Version 9 Programming Enhancements – Vive la difference?

Marje Fecht, Prowerk Consulting Ltd, Mississauga, Ontario, Canada

ABSTRACT

Performance improvements are the well-publicized enhancement to Version 9, but what else has changed that impacts your “SAS programs”? This Hands On Workshop explores many new Version 9 features including:

- new functions that may eliminate the need for complex expressions
- changes and additions to informats and formats
- interface additions and improvements
- new Export capabilities
- procedure enhancements
- and more. . .

This workshop is beneficial for Version 8 SAS users who are getting started with Version 9.

INTRODUCTION

The new high-performance architecture of SAS Version 9 enables performance improvements that are clearly the “biggest” Version 9 success story. However, there are many other additions to Version 9 that will streamline your programming, and reduce lengthy coding and logic. This hands-on-workshop focuses on the software enhancements that can change your daily coding routine.

This Proceedings paper provides a brief synopsis of the topics that will be covered in the hands-on-workshop. The detailed hands-on-workshop PowerPoint presentation is available at the SUGI 29 presentation, or from the author following SUGI 29.

INTERFACE ENHANCEMENTS

When you first access Version 9, you will notice few changes to the windowing environment. This means that you will instead be able to focus your energies on performance and programmatic enhancements.

A few interface enhancements include:

- additional icons in the Explorer window
- more integrated help facility, including SAS Online Doc®
- Import / Export wizard support of
 - Microsoft Excel 2002 spreadsheets
 - Identification of a specific sheet to import from or export to
 - Ability to export to multiple sheets.

FORMATS AND INFORMATS

You can improve the readability of your programs by using longer format and informat names. Character format and informat names can be a maximum of 31 characters. The maximum length for numeric format and informat names is now 32.

COMPATIBILITY WITH PREVIOUS SOFTWARE VERSIONS

If you require access to your Version 9 SAS data sets from Version 8, you should avoid using the longer format and informat names. In many situations, Version 8 can read Version 9 SAS data sets that use 8 byte or shorter format and informat names. For details on compatibility of SAS files across versions, see the Migration section of this paper.

FUNCTIONS

Prior to Version 9, you probably used complex sequences of functions or steps to accomplish standard tasks. When you concatenated 2 character strings, you likely had to left justify, trim, and include an extra blank to get the proper result. If you needed to determine the *second highest value* in a list of variables for each observation, you likely used quite a few comparative statements or maybe even used a PROC.

More than 60 analytic and data manipulation functions were added to Version 9. Many functions were added to Version 9 to reduce code complexity. Additionally, the logic behind some functions was revised to improve

performance (LENGTH and TRIM functions).

REDUCE CODE COMPLEXITY

In Version 9, several character functions were added to reduce the complexity of common tasks. For example, the CATX function provides a straightforward solution to joining 2 strings so the result removes leading and trailing blanks and includes a designated character as a separator.

Version 8 Solution:

```
Combine = trim(left(Char1)) || ' ' || left(Char2);
```

Version 9 Solution:

```
Combine = CATX(' ', Char1, Char2);
```

If you have used repetitive INDEX and SUBSTR functions to determine how many times a “word” occurs in a character string, that task can also be accomplished more easily in Version 9. The COUNT function determines the number of times *argument 2* occurs in *argument 1*.

SUBSTR has always been an effective function for extracting (or replacing) a portion of a character string. Two variations of SUBSTR are now available:

CHAR - to extract a single character from a text string

```
JustOne = CHAR(String , 5);
```

FIRST – to grab just the first character from a text string

```
Initial = FIRST(name);
```

If you need to strip both leading and trailing blanks from a string, you can skip using the LEFT and TRIM functions and just use the new STRIP function.

COMPUTING STATISTICS OR QUANTITIES

If you work frequently with Medians or Percentiles, you can now obtain those quantities in the DATA step. The MEDIAN function determines the median of non-missing values. The PCTL function accepts a percentage for *argument 1* to specify the percentile of interest. It then computes that percentile for the list of values.

If you need to find the *second largest* value in a list of values, or perhaps the *third smallest* value, the LARGEST and SMALLEST functions are available. In addition to supplying a list of values, you supply *which* placement you are looking for.

```
ThirdLargest = LARGEST(3 , a, b, c, d, e, f);
```

CREATING MACRO VARIABLES

Users of CALL SYMPUT commonly include LEFT, TRIM and PUT functions to make sure that the macro variable value is constructed properly.

```
call symput( 'Max' , trim(left(put(maxValue , 8.))) );
```

The Version 9 CALL SYMPUTX handles left justification, trimming, and also character conversion.

```
call symputx( 'Max' , maxValue);
```

DATA STEP ENHANCEMENTS

Several DATA step enhancements may reduce your coding and enhance your results.

LOG MESSAGING

If you use the log for custom debugging or messaging, you may find the Version 9 PUTLOG statement useful. PUTLOG always writes to the LOG regardless of the current FILE destination. Additionally, if the message begins with “ERROR:”, “WARNING:”, or “NOTE” then the message will be appropriately colored in the LOG.

USING AN IN OPERATOR

The introduction of the IN operator many versions ago reduced the need for extensive IF statement comparisons. Now, the flexibility of the IN operator has been further enhanced to reduce manual input. Integer ranges can be specified with an IN operator:

```
if code in (3 7:11 15 19:25) then put 'Found';
```

ADDITION WORTH EXPLORING

Beyond the scope of this paper, two additional enhancements well worth exploring include:

- PERL regular expressions for fast search, extract, and replace
- Hash tables for searching a collection of values on a key and for sorting values in the DATA step.

See support.sas.com for excellent explanations and examples.

PROCEDURE ENHANCEMENTS

The Version 9 enhancements to BASE SAS procedures focus primarily on ODS formatting and on Multi-thread support.

ENHANCED ODS FEATURES

The following BASE procedures were enhanced with additional features usable with ODS:

- CONTENTS
- DATASETS
- CORR
- FREQ
- UNIVARIATE.

MULTI-THREADED PROCESSING

The following BASE procedures were enhanced to support multi-threaded processing. This feature can be deactivated by using the NOTHREADS Procedure option.

- SORT
- SUMMARY / MEANS
- SQL
- TABULATE.

MIGRATION / COMPATIBILITY OF SAS FILES

To determine the compatibility of your SAS files between Version 8 and Version 9 (on the same platform), SAS has tools and assistance on the web. As of the date this paper was written, the following are some of the tools available through

<http://support.sas.com/rnd/migration/index.html>

- Compatibility calculator
- Validation tools
- Data representation / encoding determination
- PROC MIGRATE details.

If you do have to migrate your files from Version 8 to make them usable by Version 9, PROC MIGRATE is quite simple to use. The MIGRATE procedure supports most SAS file types (exceptions are well documented).

CONCLUSION

In addition to the performance enhancements with SAS Version 9, there are many opportunities to improve your coding techniques using the new functions and statements added to BASE software. Removing complex coding and replacing with the more logical new functions can streamline coding. Usage of longer format and informat names improves program readability.

The list of enhancements is extremely long; this paper's focus on day-to-day programming tasks is just a start towards effective use of Version 9.

REFERENCES

The SAS website support.sas.com has extensive and excellent documentation on the new features available with Version 9. In addition to understandable overviews of the new architecture and resulting performance enhancements, you will also find product-by-product details of all new features.

The SAS website (including the SAS Online Doc) served as the reference material for this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Marje Fecht
Prowerk Consulting

Email: marje.fecht@prowerk.com
Web: www.prowerk.com

CANADA:
PO Box 190, Port Credit Station
Mississauga, ON L5G 4L7 CANADA

USA:
PO Box 100549
Cape Coral, FL 33904 USA

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.