New SAS® System Features for Experienced Programmers  
(Or, Teaching Old Dogs New Tricks)  
Earl Westerlund, Eastman Kodak Company, Rochester, NY

ABSTRACT
The SAS® System has always been easy to use and flexible in its ability to read, manipulate and present data and information. That, combined with SAS Institute's constant additions and refinements, is why so many users are so loyal.

There is one drawback to this constant addition of new products, features and options, however: it is now nearly impossible to keep up with the SAS System. Between the flood of "Changes and Enhancements" technical reports, and the day to day pressures of getting work done, there is a strong temptation to rely on the tried and true methods of program development. After all, who needs PROC REPORT when you have PUT statements - and know how to use them? Yet it's worth the effort to keep up, because many of the new features can improve the efficiency of SAS applications, or the time it takes to develop them.

The purpose of this paper is to provide an overview of recent additions to the SAS System that are alternatives, and possibly improvements, to traditional methods of SAS programming. Topics include new techniques for data extraction and manipulation, report writing, graphics annotation, and interactive programming. The specifics of these new features will not be discussed.

This paper is intended for people who have been using SAS software longer than they care to admit, but want to keep up to date. Youngsters are welcome to attend, as long as they remain respectful of their elders.

ORGANIZATION OF THIS PAPER
The first part of this paper will discuss new tools and enhancements to old tools that have appeared in SAS software over the past few releases. "New" is a relative term. They may have appeared in the first major release of Version 6, but most of these features were buried in technical reports and were missed by a lot of us.

A reference to a SAS software manual or technical report is included for each of the topics mentioned. The numbers in brackets indicate the number of the manual in the "References" section at the end of the paper. For example, [2] pp. 30-32 is a reference to pages 30 to 32 of Technical Report P-222, which is reference number [2] in the list.

While there is not enough room to include much information or any meaningful examples in this paper, you can look at the reference cited for more information.

At the end of the paper, there are some suggestions for keeping up with the changes to the SAS System.

DATA ACCESS
Much of the work that has been done on the SAS System in recent releases has been in the area of data access. Below are some of the more significant developments.

PROC SQL
[7]; [10], pp. 194 ff.
PROC SQL is probably the most significant advance in data access. Introduced in release 6.06, PROC SQL is the SAS System's implementation of Structured Query Language.

PROC SQL can extract and combine data in ways that are very difficult to do in a DATA step. Also, as SQL is an industry standard, SAS users have a common language with users of other products that support SQL.

SAS Data Views
[4], pp. 200-201; [2], pp. 149-155
Views allow you to give users a "logical" SAS data set which may be a subset of a physical SAS data file (subset by variables, observations or both). They can also be combinations of multiple physical SAS data sets.

Views can be generated in PROC SQL, in a SAS/ACCESS® view descriptor, or in a DATA step. A flat (ASCII) file can be used as a SAS data set, through the use of a DATA step view.

The WHERE Statement
[4], pp. 498-504; [2], pp. 45-46
The WHERE statement is available in most procedures that read SAS data sets. It saves you the bother of creating an intermediate temporary data set for processing a subset of data. If the variable
on which you are extracting is an index (see below), the extract may take significantly less time. The WHERE statement also has special operators that are not available in the IF statement, allowing you to easily do substring and wildcard searches, and even search for values that "sound like" another value. Beginning with Release 6.07, you can use SAS functions in a WHERE statement.

The IN Operator of the WHERE and IF Statements
[4], pp. 123-124
The IN operator allows you to select data or do conditional processing if the value of the variable is in a list of specified values. The old style of selection required you to specify a separate equality for each value (for example, IF STATE = 'PA' OR STATE = 'NY' OR STATE = 'NJ' THEN ...). The IN operator lets you combine them into one easier to read group (for example, IF STATE IN ('PA', 'NY', 'NJ') THEN ...).

New DATA Access Options: DSD, KEY=, indexing
Index variables can significantly decrease the time to access data, but at a cost of increased disk space. The KEY= option of the SET statement can be used to directly access an observation in a SAS data set, based on the value of an index variable. It is similar in nature to the POINT= option, except that the POINT= option accesses an observation based on its physical position in the data set.

The DSD option of the INFILE statement provides greater flexibility in reading data from external files. Its purpose is to read delimited data such that two consecutive delimiters implies a missing data value.

"Client-Server" Type Applications:
SAS/ACCESS, SAS/CONNECT® and SAS/SHARE® Software
(See the respective products' usage and reference guides)
If your data reside in a corporate data base, consider leaving it where it is and accessing it via SAS/ACCESS software, rather than extracting it. SAS/ACCESS software allows you to define logical views to the data, giving users secure access to the data in a form that looks and acts just like a SAS data set.

If your application resides on one platform and the data are on another, SAS/CONNECT software may allow you to access. With SAS/CONNECT software, you establish a connection between SAS sessions on both platforms. Requests can be submitted from the client to the server, with resulting transfer of raw data or results.

Finally, if you control the data and many users require simultaneous access, SAS/SHARE software can help. SAS/SHARE software sets up a SAS data library in a server mode, allowing many users to access and update it simultaneously.

DATA MANIPULATION
New DATA Step Statements: MODIFY, REMOVE and REPLACE
The MODIFY statement allows updating of a SAS data set in place without rewriting. This can be useful if, for example, you need to simply recalculate the value of a variable in a data set, and do not need to add or drop variables. The REMOVE and REPLACE statements allow for deletion and replacement of observations without rewriting the data set.

New Functions: INDEXW, TRANWRD, COMPBL, and Run-time Format Specifications
There are many new functions. INDEXW works like INDEX, except that it returns the starting location of a specified string of characters, rather than a single character. TRANWRD works like TRANSLATE, except that it translates by word rather than by character. COMPBL compresses multiple blanks between words in a character string to a single blank. It is similar to the %CMPRES macro function.

The functions INPUTC, INPUTN, PUTC and PUTN work like the INPUT and PUT functions, except that allow you to specify the informat or format to use at run time.

DATA ANALYSIS AND REPORTING
The REPORT Procedure
[6]; [2] pp. 269-274
PROC REPORT is a very flexible tool for report writing. With PROC REPORT available, much DATA step report writing is no longer necessary. Some of the key features of PROC REPORT are:
- Ability to define the report layout interactively or through a command language
- Ability to store the report definition in a catalog entry for subsequent reuse
- Computation of summary statistics to be included in the report
- Multiple column output ("phone book style")
The FSVIEW Procedure
The FSVIEW procedure allows you to view and edit your SAS data set in a tabular format. FSQUERY and FSEDIT display one observation at a time. FSVIEW allows you to define a formula for a computed field such as a total.

The DATA Step Graphics Interface
[12] pp. 597-736
If the procedures in SAS/GRAPH® software or the annotate facility do not give you the graphical output you need, the DATA Step Graphics interface is available. DSGI enhances the DATA step with primitive graphics calls, allowing you to generate custom graphs that can be displayed on any supported device and stored in SAS catalogs as any SAS/GRAPH output.

Enhancements to the FORMAT Procedure
[8] pp. 300-304
The CNTLN= option of the FORMAT procedure allows you to generate a format using a SAS data set as input, rather than SAS program statements. Conversely, the CNTLOUT= option will generate a SAS data set containing the values of a user-defined format. The FMTLIB option displays the defined values of a format.

The CLASS Statement of the MEANS Procedure
[8] p. 370
Beginning in Version 6 of the SAS System, the MEANS and SUMMARY procedures are essentially the same. Because of this, the CLASS statement is now available in PROC MEANS. This saves processing time, because input data sets do not have to be sorted if subtotals are desired.

#BYVAR and #BYVAL in TITLE Statements
Special codes #BYVAR and #BYVAL are available to customize TITLE statements when BY group processing is being done. The result is a report with the value of the BY variable embedded in the title, rather than having to specify a generic title with a BY line below.

The PLOT Procedure: The $ Operator of the PLOT Statement and the VTOH= option
This is a small one, but it illustrates the gems that are hidden in the Changes and Enhancements books. This option allows two variables to be plotted, with the value of a third variable replacing the plotting symbol. The VTOH= option allows you to define the ratio of vertical lines to horizontal print positions — essential for producing square plots. Finding these two options allowed me to literally replace a 230 line SAS macro with three statements (including the RUN statement)!

APPLICATIONS DEVELOPMENT

WINDOW / DISPLAY and %WINDOW / %DISPLAY Statements
Sometimes using SAS/AF® software is overkill for a small application that requires a minimal amount of input and error checking. At these times, consider the WINDOW and %WINDOW statements to define windows and the DISPLAY and %DISPLAY statements to show them.

PROC SQL DICTIONARY-Tables (and SASHELP Views)
The PROC SQL DICTIONARY-Tables have much of the information that is in many of the Display Manager windows. Through them, you can find out system information such as:
- what SAS data libraries are allocated
- available catalogs, data sets, and variables
- current settings for SAS system options
They are also available as SAS data views in the SASHELP data library.

TIPS FOR KEEPING UP

The previous pages are just a sampling of some features that some of us have found to be very helpful. There are many more — far too many to fit into a paper like this. That is not really the point of this paper, anyway. The point is that you need to keep up with the SAS System. But given that there is so much, how do you keep up? You need a game plan for staying current.

Following is a summary of some tips to keep track of changes.

Know What You Need to Know
Given that there is no way you can keep up with everything, make sure you keep up with those areas of the SAS System that are most important to you. If you work in a group of data analysts — engineers, scientists, statisticians — you should be watching for enhancements to products like SAS/STAT®, SAS/QC®, and SAS/INSIGHT®. If you develop applications that rely heavily on user interfaces to corporate data bases, you should be tracking PROC PMENU, SAS/AF frame entries, and SAS/ACCESS software. Everybody should monitor base SAS software.
RTFM ("Read Those Fabulous Manuals")
Sounds boring, doesn't it? But that's where the information is. You do not have to read manuals cover to cover. Most reference manuals have "Changes and Enhancements" sections at the beginning. Read them. Skim through options of statements and procedures that you use often. Finally, be sure that you get and read the "Changes and Enhancements" technical reports. That is where you will find options to old procedures that allow you to get rid of a 230-line macro in favor of a three-line program.

Any group that uses SAS software should consider subscribing to the SAS Subscription service. You will receive manuals and technical reports that you do not want, but it's worth it to receive the manuals and technical reports that you need.

Play (I Mean, Research): Try Out New Things
When you get a new release of the SAS System, set aside some time to test it. Work through some of the examples in the manuals. Become familiar with the new features before you must. If you know in general terms how a new feature works, you'll be less likely to fall back on the "tried and true" the next time you're facing a tight deadline.

Bring in and Look at New Products
Take advantage of SAS Institute's trial periods to bring in and test new products. It is hard to cost-justify a product if you have never really used it, compared to seeing the benefits and hidden costs first-hand.

Look at SAS Software in Combination with Other Software
The computer technology explosion is not just limited to SAS software. Software you have on hand, used in combination with SAS software, may suit your needs better than a "straight SAS" solution. Perhaps the answer to your Windows application is not a SAS/AF user interface but a Visual Basic™ one, linking to a SAS application via DDE.

Keep in Touch with Peers
No matter how hard you try, you will not be able to keep up to date by yourself. Do not try. Join a local user group and participate in the meetings. Your peers in other organizations will have a different perspective, and will find other items that you have overlooked. Be prepared to share what you have learned, too. You know more than you think you do!

If you can, join SAS-L. It's a worldwide electronic user group. The participants are very happy to share information and offer advice and opinions on all types of subjects relating to SAS software. Many of the ideas in this paper are the result of a note I posted on SAS-L, asking for suggestions for topics to include in this paper. If you do not know much about going "on line," you might refer to Michael Davis's paper in the SUGI 18 Proceedings [1].

CONCLUSION
The SAS System has grown to the point that it is impossible to keep track of all new features. This paper has provided a brief summary of some of these new features, as well as some suggestions for developing a strategy for keeping up with the SAS System in the most effective manner.

REFERENCES
ACKNOWLEDGMENTS

I would like to thank all the people who responded to my SAS-L plea for suggestions:

Deborah Cannon  Lawrence H. Muhlbaier
Ben Conner    Bob Reimer
Kemon Gibes    Pierre Traissac
Keith Kaiser  Mike Whitney
Joseph Kelley  Jay Jacob Wind
Melvin Klassen  Mike Zdeb

COPYRIGHTS

SAS, SAS/ACCESS, SAS/AF, SAS/CONNECT, SAS/GRAPH, SAS/INSIGHT, SAS/QC, SAS/SHARE and SAS/STAT are registered trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Visual Basic is a trademark of Microsoft Corporation.

AUTHOR

Earl Westerlund
Eastman Kodak Company
1669 Lake Ave., Mail Code 24608
Rochester, NY 14652-4608
(716) 722-1983
Internet EarlW@Kodak.COM