Getting From SAS-A to SAS-B: Some differences between academic and business SAS use

INTRODUCTION

The ways in which SAS is used across industries varies greatly. For the SAS professional this can have significant implications throughout an individual's career. As job titles change, it is likely the job duties will as well, and in such situations professionals may be required to adapt to new ways of using SAS. One area where there appears to be a significant difference in SAS-based practices is the way that SAS is used for academic versus professional use.

The purpose of this paper is to present information relevant to SAS professionals who are making the transformation from academic SAS use to business SAS use. This information should also be helpful to other users who are new to business applications of SAS and have no prior academic experience with SAS. SAS programmers and managers who are not from academic backgrounds may find this information helpful in understanding the perspectives of their academically trained SAS users and employees.

The format of this paper will not be a "how to", but a "what to", as in "what to think about" and "what to learn" to make the transition from SAS for academics (SAS-A) to SAS for businesses (SAS-B) easier. The resources section provides references for learning more about and using these features.

There are several rough categories into which the differences between SAS for academics and SAS for businesses fall. I will discuss each of these in turn and provide some examples of how they are different in each area.

These categories are:

- the level of programming emphasis;
- the use of various display and output styles;
- knowledge of and use of SAS software's depth;
- community differences; and
- different means of furthering career progress.

USERS VS. PROGRAMMERS

The main distinction in how using SAS for academic purposes differs from using SAS for business purposes is likely related to the different levels of programming expertise required to complete job functions. For SAS-A use, SAS is essentially an analytical tool. Understanding the statistical methods involved and interpretation of the output is of greater importance than writing efficient code. The level of programming knowledge necessary to answer a given research question does not venture far beyond what is needed to recode variables in a flat file and run the relevant PROCs. For this reason, we might

say that SAS-A tasks require "user" level knowledge of SAS, or knowledge of how to use the basic tools SAS provides.

In contrast, SAS-B uses often emphasize efficient programming, data mining, and data manipulation capabilities. Being able to manipulate and develop basic SAS tools becomes more important. Because of this, as one moves from SAS-A to SAS-B, it will be beneficial to change how one approaches the writing of SAS syntax. The copy and paste style that one can get by on in academics makes for slow, difficult to read, and un-necessarily long programs when used in complex data mining scenarios. Two things that academic users rarely place an emphasis on, making SAS programs run efficiently and ensuring that they are readable by others, are likely to be of primary importance in SAS-B situations.

It may be helpful for the recently transitioned or beginning SAS-B user to read a short introductory programming book to expose themselves to programming concepts such as branching, iteration, logical programming structure, and programming best practices. Using macros via the SAS Macro language provides another way to simplify the writing, editing, and execution of SAS programs. SAS also integrates features of Structured Query Language (SQL) and of course, the SAS Component Language (SCL). Developing even a basic understanding of these options will quickly expand the selection of tools you have available when approaching a task.

DISPLAY/PRESENTATION OF OUTPUT

As academics, we were rarely concerned with the appearance of our SAS output. Compared to other statistical packages, the default SAS output is rather crude and esoteric. No journal in its right mind would take a frequency table from the default SAS output and as such, it was a given that any output in SAS would be as quickly as possible moved to some other method of display, such as Microsoft Excel. However necessary or well-suited this approach was for academic publishing, it leaves academic SAS users without exposure to quite a few capabilities of SAS that can be useful in the professional world.

In the business world, it is not always necessary to produce publication quality statistics, but it is still important that end-users are presented with information in a clear and concise way. In such a case PROC FORMAT, label statement, title statements, and footnote statements can go a long way in producing output that requires minimal post-analysis editing and is easily modified when changes to the analysis occur. Consider the amount of time and effort saved by being able to automate the output formatting process rather than exporting to Excel and manipulating the output there, especially when frequent changes are involved. Formatting output via SAS programs will save time and effort for repetitive tasks or when inevitable changes occur.

A common request these days is to make content available on the web. It is precisely the simplicity of SAS output that makes it well suited for such a task. The SAS output delivery system (ODS) makes moving information from SAS to the web quite simple

with its HTML output option. By using PROC TEMPLATE, you can produce customized tables in almost any style you need.

KNOWLEDGE OF SPECIAL SAS

As a SAS-A user you likely never needed to move beyond the capabilities of base SAS (with the exception of possibly the SAS/STAT package). As a SAS-B user you may also never need to move beyond the capabilities of base SAS, but you may find it helpful to do so. SAS produces a large number of products that extend its software's functionality for specific uses. Licenses for these packages often cost enough to make them too expensive to justify purchasing them for periodic individual use, but as a SAS-B user such software may already be part of your company's site installation, or at least within your company's budget. It is worthwhile to spend some time on the SAS website and SAS-L listserv researching the available packages and their uses when developing a course of action for your next major project.

A COMMUNITY SHIFT

An individual moving from SAS-A to SAS-B use might not expect it, but the SAS-B community closely parallels the structure and functions of the academic community. Topic specific conferences of academia are replaced with (or complemented by) tool-specific SAS conferences which serve the same purposes of networking, disseminating knowledge, and providing complimentary continental breakfasts. Likewise, the esoteric academics we previously turned to for assistance and wisdom become esoteric programmers.

In contrast to the sometimes secretive culture of academic research, members of the SAS community are relatively open to sharing knowledge. One of the best places to get answers to questions is via the web; specifically, the SAS-L listserv, local listservs, previous conference proceedings, and the personal web pages of fellow SAS users (see the "Resources" section for more information). For the SAS-B user who is the lone SAS user in their area, these resources can be exceptionally useful.

CONTINUING EDUCATION

Whereas career progress in academia is marked by publications, teaching evaluations, or the tenure process, certification is one way of providing evidence of continued development and continuing education in the professional arena. SAS offers certification tracks that focus on general base programming, advanced programming, database warehousing and web applications development. As a search for the term "certification" on the SAS-L archives shows, the SAS certification program had a somewhat slow path to acceptance among the SAS community. As of late though, the sentiment appears to be that obtaining a SAS certification is a worthwhile endeavor.

SUMMARY

In summary, there are five major areas that define the differences in what it means to use SAS in academic vs. business settings. Those moving from SAS-A to SAS-B may ease their transition into their new positions (or increase their potential of landing one) by taking the following steps:

- Familiarizing themselves with basic programming logic and best practices and practicing these skills. Exploring the capabilities of the SAS Macro language. Learning to think like a programmer.

- Becoming familiar with PROC FORMAT, within PROC formatting and labeling, the capabilities of the Output Delivery System (ODS), and PROC TEMPLATE.

- Learning about additional software tools other than base SAS and how they may be better suited to certain tasks.

- Joining the SAS-L and local SAS user group listservs.

- Obtaining SAS certifications.

CONCLUSION

Despite what your adviser told you, moving from SAS-A to SAS-B can be a rewarding transition. There are some difficulties to overcome, but being cognizant of the differences you will experience will help you better plan for them. For SAS-B programmers and managers who have never experienced the SAS-A side of the world, understanding the different skills that academically trained SAS users will bring to their work will allow you to adjust your programming and management accordingly.

RESOURCES

The Little SAS Book: A Primer. Lora D. Delwiche, Susan J. Slaughter. SAS Publishing; 3rd edition (November 2003)

SAS Certification Prep Guide: Base Programming. SAS Publishing. (September 2004)

Learn to Program. Chris Pine. Pragmatic Bookshelf (Jan 13, 2006)

SAS-L Listserv:
To read archives, subscribe, or post to the list visit:
http://listserv.uga.edu/archives/sas-l.html

SUGI archives:
http://www.lexjansen.com/sugi/

CONTACT

Bryan Shepherd
Texas Health Care Information Council
bryan.shepherd@austin.utexas.edu

Thanks to several members of the SAS-L listserv for their helpful and insightful comments on this topic, namely David Cassell, Ronald Fehd, John Gerstle, Sigurd Hermansen, and Arthur Tabachneck.