BASUG Quarterly Meeting Announcement

Please join us for these informative talks, and consider staying for the afternoon class (separate event), taught by Russ Lavery. *For information on the training, please visit Training Announcement*

After the meeting, we will provide an informal light buffet lunch for all attendees. We hope you can stay for this opportunity to network and socialize with your fellow SAS users.

Topic	**Mixed Bag!**
When	**June 23, 2017** 8:15AM – 12.30PM
Where	Description: Description: MCS_NERD_Logo - Copy Microsoft New England Research and Development Center (NERD)[1] One Memorial Drive Conference Center, First Floor Cambridge, MA 02142 857-453-6000
Directions	Please visit the meeting site directions page
How to register	Individual, on-line registration required. Please visit the event registration page NO WALK-INs!
Payment	$10 – if paid on-line by Tuesday, June 13 $15 – if paid on-line by **NOON** Thursday, June 22 $20 – at-the-door (checks only)
Contact	If you have questions about the meeting, contact the meeting organizers: Karen Olson Bridget Neville

Agenda*

8:15AM	Sign in and Refreshments
8:45AM	Announcements
9:00AM	The SQL Optimizer Project *by Russ Lavery, Independent Consultant*
10:00AM	Break
10:15AM	Some _FILE_ Magic *by Mike Zdeb, University at Albany, School of Public Health*
10:40AM	PROC Freq and Tables with Zeroes as Marginal Totals *by Mike Zdeb, University at Albany, School of Public Health*
11:10AM	Break

11:25AM	An Annotated Guide: Using PROC Tabulate and PROC Summary to Validate SAS Code *by Russ Lavery, Independent Consultant*
12:00PM	Learning the PDV through the Data Step Debugger *by Russ Lavery, Independent Consultant*
12:30PM	Networking Lunch

**Note: Times (and sequence) are approximate and subject to change. Please re-visit the BASUG website (www.basug.org) for updated information.*

Abstracts and Speaker Biographies

The SQL Optimizer Project

by Russ Lavery

This paper discusses two little known options of SAS® PROC SQL: _method and _tree. This paper will also attempt to describe some of the logic that the Optimizer employs. PROC SQL has a powerful subroutine, the SQL Optimizer, that examines submitted SQL code and the state of the system (file size, index presence, buffersize, sort order etc). The SQL Optimizer creates a run plan for optimally running the query. Run plans describe executable programs that the Optimizer will create to produce the desired output. These executable programs can be quite complicated and often involve the creating, sorting and merging of many temporary files. Consistent with the Optimizer s goal of minimizing run times, the executable programs will trim variables and observations from the input file(s)/working file(s) as soon as they can be removed. Many details of the run plan can be determined by using two PROC SQL options (_Method and _Tree) and this paper explains output from these two options. _Method and _Tree produce different output that present different aspects of the run plan. Learning to interpret _Method and _Tree can help programmers explain why small variations in code, or system conditions, can cause substantial variations in run time.

Russ Lavery has over twenty years' experience using SAS on both windows and UNIX machines. He has both built and managed data marts and then created and managed production reports produced from the data marts. He is active in the SAS community and has presented over sixty papers at SAS conferences/universities and companies all over the US, in Europe and China. He has won three best paper awards and one of his talks, with his voice, was burned onto a CD and is included in the back of Art Carpenter's book on PROC Report (which is the training he will be providing at BASUG). He is one of sixty "SAS Partners" in the North East section of the US.

Some _FILE_ Magic

by Mike Zdeb

The use of the SAS® automatic variable _INFILE_ has been the subject of several published papers. However, discussion of possible uses of the automatic variable _FILE_ has been limited to postings on the SAS-L listserv and on the SAS Support Communities web site. This paper shows several uses of the variable _FILE_, including: creating a new variable in a data set by concatenating the formatted values of other variables; recoding variable; searching for variable values.

PROC Freq and Tables with Zeroes as Marginal Totals

by Mike Zdeb

Producing consistent looking tables over various iterations of data is sometimes a problem. Variable values that occur at one point in time and not in another can result in tables that look different given that SAS PROCs can only count values that appear in the data. Three procedures (REPORT, SUMMARY, TABULATE) permit the use of preloaded formats that can overcome that problem. This paper show how using PROC SUMMARY to pre-count data extends that capability to PROC FREQ allowing a user to produce tables with zero counts as marginal totals.

Mike Zdeb is an assistant professor in the epidemiology/biostatistics department at the University at Albany School of Public Health in Rensselaer, NY (and also "semi-retired"). He has used SAS for 25+ years and has presented papers and conducted workshops at SUGI, SAS Global Forum, NESUG, SESUG, and at numerous local user group meetings in the US and Canada. He has been a contributor to the SAS Community WIKI, the SAS Support Communities, and the SAS-L listserv. Mike is the author of the SAS Press book "Mapping Made Easy Using SAS" and has also been a reviewer for a number of SAS Press books.

An Annotated Guide: Using PROC Tabulate and PROC Summary to Validate SAS Code

by Russ Lavery

This paper discusses how PROC Tabulate and PROC Summary can be used to help a programmer avoid errors. This paper might be especially useful for new programmers. A programmer should know where “every observation goes”. Additionally, Good Programming Practice requires that a programmer “exercise” (check in some way) every combination of the clauses in every IF statement to see how observations are classified by the IF. PROC Tabulate is an excellent, fast and simple tool for doing that checking. Importantly, PROC Tabulate produces a pattern, in its output, that makes checking fast. Additionally, using PROC Tabulate allows a programmer to account for every observation in the data set, a “comforting” fact. Many programming tasks require taking subsets of the data, based on logical rules. Programmers are often told to check the number of observations in the starting data set and not to lose, or create, observations without knowing why. With PROC Summary we can also suggest that a programmer also check the sums of important variables (Dollars, Scripts, units sold) in the source data set and never change the sum without knowing why.

Learning the PDV through the Data Step Debugger (DSD)

by Russ Lavery

The Data Step Debugger (DSD) simplifies debugging Data Steps and not whole programs. Using the debugger requires some understanding of how SAS works because the DSD does not issue any error messages. When using the DSD, a programmer mentally compares what s/he sees, in the Program Data Vector (PDV), with what s/he expected to see. Critical, to use of the DSD, is an understanding of the PDV and, as a side issue, the DSD is an excellent way to learn the PDV. The goal of a programmer should be to learn to combine DSD commands and to link those combinations to a key or a macro. Linking a series of commands to a key, or a macro makes the DSD much more powerful.

Event Registration Page
Top of Meeting Announcement
Training Announcement
Return to BASUG

BASUG Contacts

Mailing Address:

BASUG
PO Box 170253
Boston, MA 02117

Email the BASUG Webmaster

(1) The Microsoft New England Research & Development Center (NERD) is a research and software innovation campus located in the heart of Cambridge, Massachusetts. The NERD vertical campus spans two buildings with its primary presence and conference center located at One Memorial Drive and a recently renovated and expanded space located at One Cambridge Center. NERD is home to some of Microsoft's most strategic teams including Microsoft Research New England, Microsoft Application Virtualization (App-V), SharePoint Workspace, Microsoft Technical Computing, Microsoft Advertising, Microsoft Lync, Microsoft Office 365 and more. NERD has become a hub of activity for the local tech community and has hosted more than 500 events and welcomed more than 40,000 visitors during the past two years.