



## BASUG Meeting 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	A little bit of this, a little bit of that
When	Wednesday, March 23, 2016 8:15AM - Noon
Where	Microsoft New England Research & Development Center (NERD) <sup>1</sup> One Memorial Drive Cambridge MA 857-453-6000
Directions	Please visit the meeting site directions page ( <a href="http://microsoftnewengland.com/About">http://microsoftnewengland.com/About</a> )
How to Register	Individual, on-line registration required. Sorry, NO WALK-INS. <a href="http://www.basug.org/index.php?p=eventreg">Register Now! (http://www.basug.org/index.php?p=eventreg)</a>
Payment	\$10 -- if paid on-line by Monday, March 14, 2016 \$15 -- if paid on-line by NOON on Tuesday, March 22, 2016 \$20 -- at the door (checks only)
Contact	If you have questions about the meeting, please contact the meeting organizers (mailto:Lori.Goldman@siemens.com)

## Agenda\*

Time	Activity
8:15AM	Sign-in and Refreshments
8:45AM	Announcements
9:00AM	<b>"Coding for Performance"</b> , by Robert Rosofsky, Health Info. Systems Consulting LLC and Malcolm Rucker, Harvard Pilgrim Healthcare Institute
9:50AM	Break

Time	Activity
10:10AM	<b>"Key Features in ODS Graphics for Efficient Clinical Graphing"</b> , by Yuxin (Ellen) Jiang, Biogen
10:25AM	<b>"I Got a Fake ID! A macro to de-identify user-specified variables in a dataset"</b> , by Jen Popovic, Harvard Pilgrim Healthcare Institute
10:55AM	Break
11:10AM	<b>"Creating Code Templates in the SAS Enhanced Editor using Abbreviations and User Defined Keywords"</b> , by Paul Grant, SAS Institute Inc. (retired)
11:25AM	<b>"Simplifying Complex Character Comparisons by Using the IN Operator and the Colon (:) Operator Modifier"</b> , by Paul Grant, SAS Institute Inc. (retired)
11:40AM	<b>"Using the Power of Proc Geocode to Make More of Your Geographic Datasets"</b> , by Kieran Shah, Abt Associates Inc.
11:50AM	<b>"Talk to me!"</b> , by Elizabeth Axelrod, Abt Associates Inc.
12:00PM	Networking Lunch
<p><i>* Note: Times (and sequence) are approximate and subject to change. Please re-visit the BASUG website (<a href="http://www.basug.org/index.php">http://www.basug.org/index.php</a>) for updated information.</i></p>	

## Abstracts and Speaker Biographies

### "Coding for Performance", by Robert Rosofsky and Malcolm Rucker

With more use of "big data" in organizations today, getting your SAS programs to run in an acceptable amount of time becomes more important and more difficult. We rewrote from scratch an application that creates multiple summary files, processing nearly all records from detail-level SAS tables, which in some cases, contain billions of rows. The application is routinely run at multiple sites after every database refresh is approved, and the highly summarized results are used by another application which supports rapid feasibility assessments (e.g., Do we have the sample size necessary to warrant a study?).

We increased the speed of execution of the application at sites by as little as 5 times to as much as 36 times, in some cases dropping elapsed runtimes from the 3-5 days range to the 2-20 hours range. It's our hope that any SAS programmer, from the beginner to the advanced, will be able to benefit from many of the techniques we will present.

To improve patient care and public health through the effective use of information systems, **Robert Rosofsky** has been developing and implementing health information systems for nearly three decades in both the public sector and as an independent consultant. He offers services in project management, systems analysis, SAS data management/statistical programming, and SAS training.

**Malcolm Rucker** specializes in designing and building automated SAS solutions to analytic problems using information generated from health claims and electronic medical records (EMR) data sources. In the past, Malcolm has worked at non-profit integrated provider/insurer health care organizations, for-profit health-industry consulting

firms, and he is now currently working at the Department of Population Medicine (DPM) of Harvard Pilgrim Healthcare Institute supporting the Sentinel project.

## "Key Features in ODS Graphics for Efficient Clinical Graphing", by Yuxin (Ellen) Jiang

High-quality effective graphs not only enhance understanding of the data but also facilitate regulators in the review and approval process. In recent SAS releases, SAS has made significant progress toward more efficient graphing in ODS Statistical Graphics (SG) procedures and Graph Template Language (GTL). A variety of graphs can be quickly produced using convenient built-in options in SG procedures. With graphical examples and comparison between SG procedures and traditional SAS/GRAPH procedure in reporting clinical trial data, this paper highlights several key features in ODS Graphics to efficiently produce sophisticated statistical graphs with more flexible and dynamic control of graphical presentation.

**Ellen Jiang** is a Senior Analyst III in Biogen's Clinical Data Science group. She has worked in Biotech/Pharmaceutical industry for more than 8 years, supporting statistical analyses, reporting, and submission for Phase I-III clinical trials. Ellen has MPH in Biostatistics from Boston University School of Public Health. She is SAS certified programmer.

## "I Got a Fake ID! A macro to de-identify user-specified variables in a dataset", by Jen Popovic

This presentation introduces a SAS macro that functions to de-identify a dataset by assigning a randomly generated 'caseid' to one or more user-specified variables. This program has application in projects for which individual-level data are needed but actual values of person-level identifying variables are required to be masked. For context, this presentation also includes background about different data de-identification methods and why they are important. Although the focus is on the use and de-identification of healthcare data, the ideas and methods presented are conceptually applicable to data from any source.

**Jen Popovic** is the Assistant Director for Scientific Solutions for the Sentinel project at Harvard Medical School and the Harvard Pilgrim Health Care Institute. Sentinel is a US Food and Drug Administration-funded project to create an active surveillance system to monitor the safety of FDA-regulated medical products. She oversees the strategic direction for data- and analytic-development initiatives relating to Sentinel, as well as the team of analytic programmers who implement the initiatives. Jen is also a licensed veterinarian who maintains academic interests in the clinical sciences and the pathologic basis of disease.

## "Creating Code Templates in the SAS Enhanced Editor using Abbreviations and User Defined Keywords", by Paul Grant

Two features of the SAS Enhanced Editor make it possible for programmers to create templates for SAS code. Code templates are portions of code which are pre-programmed with the statements, options and formatting most useful to you and which can be inserted into your programs by typing a keyword. The two features of the Enhanced Editor which make this possible are abbreviations and user defined keywords. This presentation will show you how to create your own code templates using these two features.

## "Simplifying Complex Character Comparisons by Using the IN Operator and the Colon (:) Operator Modifier", by Paul Grant

Complex character comparisons - comparisons of character values in a data set against a number of character constants which vary in length - are easier specified than coded. For example, how would you select the records of customers whose last names begin with 'Mc' or 'Mac'? How would you select the records of customers who live in ZIP codes beginning with '010' through '0131'? This paper will show you how to use the IN operator and the colon (:) operator modifier to code these comparisons simply and clearly.

Until retiring in 2015, **Paul Grant** was a Senior Systems Engineer and Technical Architect with SAS Institute for 17 years. His customers included the U.S Treasury Department, the FDIC, the U.S. Joint Chiefs of Staff, and the FBI. He chaired SUGI 21 in Chicago in 1996 and co-chaired NESUG '94 in Philadelphia. He has been active in the Boston Area SAS Users Group since 1983. He originated Coders' Corner and the Sunday Workshops at SUGI and NESUG. He has been developing systems with SAS since 1981.

## "Using the Power of Proc Geocode to Make More of Your Geographic Datasets", by Kieran Shah

Survey and administrative data often offer a rich variety of geographic-specific data, which if used in conjunction with SAS's Proc Geocode, and Proc GMAP can yield interesting and unexpected new insights. However, such data will often lack the specific geographic variable needed for your analysis, or may lack the appropriate variable desired for a mapping analysis. With the addition of the Street-level geocoding in SAS 9.2, it is possible to attain census tract, census block, zip, and county level information from an address, along with the important longitude and latitude. This presentation will demonstrate how to combine these two procedures to produce interesting information, along with detailing where to find the most useful U.S. maps.

**Kieran Shah** is a Programmer Analyst working at Abt Associates in Cambridge. He has worked on a variety of projects using geographical data, creating maps and providing distance analysis for clients such as the State of Minnesota, CMS, and the Federal Justice and Statistics. Prior to working at Abt Associates, he completed an MPH in biostatistics at Boston University.

## "Talk to me!", by Elizabeth Axelrod

Wouldn't it be nice if your long-running program could tap you on the shoulder and say 'Okay, I'm all done now'. It can! This quick tip will show you how easy it is to have your SAS® program send you (or anyone else) an email during program execution. Once you've got the simple basics down, you'll come up with all sorts of uses for this great feature, and you'll wonder how you ever lived without it.

**Elizabeth Axelrod** has been using SAS for over 34 years, and she's been a SAS fan(atic) for almost that long. She is a Lead Programmer Analyst at Abt Associates Inc., where she constructs large-scale research files in the Health Policy area. As a co-director of Abt's Data Management and Analytic Computing Methods Center, she develops guidelines for quality assurance, reproducibility, and documentation. She is currently the President of BASUG – the Boston Area SAS Users Group.

## **BASUG Contacts**

Mailing Address:

BASUG

PO Box 170253  
Boston MA 02117

Email Our Webmaster (<mailto:basugwm@basug.org>)

*<sup>1</sup>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.*