

BASUG Quarterly Meeting Announcement

Got a new dataset? How much missing data does it have? How do you know there are no missing values if you don't look *in* the dataset? Can you find "xxx" in at least one of the variables? Let Mike Zdeb tell you how approach these questions.

A hash of hashes? How does that work? Judy Loren will explain this advanced use of hash tables to rapidly link and process large datasets to solve a classic healthcare challenge.

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	Data Manipulation for the Adventurous Programmer
When	Tuesday, April 17, 2012 8:15 a.m. – Noon
Where	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 (obsolete)
Contact	If you have questions about the meeting, contact one of the organizers. - o Karen Olson - o Amy Baker - o Lisa Kang

Agenda

8:15	Sign in and Refreshments
8:45	Announcements
9:00	An easy route to a missing data report with ODS + PROC FREQ + a data step <i>by Mike Zdeb, University at Albany School of Public Health</i>
9:45	Break
10:00	Searching for variable values with CAT functions: An alternative to arrays and loops <i>by Mike Zdeb, University at Albany School of Public Health</i>
10:45	Break
11:00	Building provider panels: An application for the hash of hashes <i>by Judy Loren, Health Dialog</i>
12:00	Lunch

Speaker biographies and abstracts

An Easy Route to a Missing Data Report with ODS+PROC FREQ+A Data Step

by Mike Zdeb

A first step in analyzing data is making a decision on how to handle missing values. That decision could be deletion of observations and/or variables with excess missing data, substitution of imputed values for missing values, or taking no action at all if the amount of missing data is insignificant and not likely to affect the analysis. This presentation shows how to use an ODS OUTPUT statement, PROC FREQ, and some data step programming to produce a missing data report showing the percentage of missing data for each variable in a data set. Also shown is a method for identifying and dropping from a data set all variables with either all or a high percentage of missing values. The method of producing the missing data report is less complicated and more compact than several methods already proposed in other papers.

Searching for Variable Values with CAT Functions: An Alternative to Arrays and Loops

by Mike Zdeb

When one wants to search through a large number of variables for one or more specific values, the common approach is to first use an array to store the variable values and then use a DO loop to cycle through the array elements looking for the value(s) of interest. With the advent of the various CAT (concatenation) functions in version 9, there is an alternative to the array/loop approach that reduces the amount of SAS code needed to search for values, with the added benefit of taking approximately as much elapsed time as the array/loop approach. This presentation shows several scenarios where one of the CAT functions can be used in combination with version 9 FIND and COUNT functions to simplify the task of searching for variable values. The standard approach (an array plus a loop) is shown first, followed by the CAT function solution.

Mike Zdeb is an assistant professor in the Department of Epidemiology and Biostatistics at the University at Albany School of Public Health in Rensselaer, NY. He has been a SAS user for 20+ years and has presented papers and conducted workshops at SUGI, SAS Global Forum, NESUG, and numerous local user groups. He has written a SAS Press book, *Mapping Made Easy Using SAS*, and has been a reviewer for a number of SAS Press books. Among the courses he teaches are a 3-credit introductory SAS class and a 1-credit advanced topics in SAS class. Mike and his teaching are the subject of a "Customer Success Story" in the area of higher education on the SAS web site.

Building Provider Panels: An Application for the Hash of Hashes

by Judy Loren

Did you know that an element of a hash table can be another hash table? How cool is that? Kind of like a multidimensional array on steroids. You don't even have to know in advance how many you're going to need (not even in macro variables)!

The "hash of hashes" technique provides one solution to a classic challenge in healthcare: how to process claims data to determine which doctor within a given specialty is most likely overseeing a patient's care. The algorithm requires examining an unpredictable amount of information and making decisions at defined breakpoints.

This paper will demonstrate the code for a "hash of hashes" and compare it to other approaches to the same challenge (the double DOW and SQL). Along the way we'll make use of some

features of hash tables that were new in 9.2. If you love hash tables, you'll flip for the hash of hashes!

Judy Loren is a Senior Scientist at Health Dialog, which partners with payers and providers to achieve the Triple Aim: Better care for individuals, better health for populations, and lower per capita costs. Her current project leverages information and technology to help the new Accountable Care Organizations (ACOs) achieve those goals. An avowed SAS bigot, Judy is proud to follow in her grandmother's footsteps--the one who converted the State of Maine payroll to a computer-based system back when programming meant hard-wiring boards.

BASUG Membership

We are in the process of transitioning our event registration and membership system. Although we are still using EventBee to handle our meeting registration, we are not currently processing any memberships. We will address new memberships and renewals for current members at a later date, when the new system is ready.

BASUG Contacts

Mailing Address:

BASUG
PO Box 170253
Boston, MA 02117

Email our Webmaster: basugwm@basug.org

-
- (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.

BASUG Quarterly Meeting Announcement

Visualize your data with SAS maps. We are pleased to have 3 experts on SAS maps present at our next quarterly meeting: Darrell Massengill, Mike Zdeb and Louise Hadden.

Please join us for these informative talks, and consider staying for the afternoon class (separate event), taught by Mike Zdeb. *For information on the training, please visit [the 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 SAS Mapping

When Tuesday, June 26, 2012
8:15AM – Noon

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! You MUST pre-register to attend.

Payment \$10 – if paid on-line by June 15, 2012
\$15 – if paid on-line by 5pm on June 25, 2012
\$20 – at-the-door - check only

Contact If you have questions about the meeting, contact the meeting organizers:
[Bruce Haimowitz](#), [Joan Yu](#), or [Eric Hamann](#)

Agenda*

8:15	Sign in and Refreshments
8:45	Announcements
	Poster Presentation
9:00	ZIP Code 411: Decoding SASHELP.ZIPCODE and Other SAS® Maps Online Mysteries by Louise Hadden, Abt Associates
9:10	What's New in Mapping by Darrell Massengill, SAS Institute
9:55	Break

10:05	Google Maps and SAS/GRAPH <i>by Darrell Massengill, SAS Institute</i>
10:45	Break
11:00	Driving Distances and Times using SAS & Google Maps <i>by Mike Zdeb, University at Albany</i>
12:00	Networking Lunch

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

Speaker biographies and abstracts

Poster Presentation

ZIP Code 411: Decoding SASHELP.ZIPCODE and Other SAS® Maps Online Mysteries

by Louise Hadden, Abt Associates, and Mike S. Zdeb, University at Albany

SAS® provides a very useful file called SASHELP.ZIPCODE. This file is relatively unknown and under-utilized. This paper explores current and potential uses of the file, as well as how to update the file on a quarterly basis by accessing SAS MAPS ONLINE. The authors also provide a walk-through of several programming projects that they have recently conducted. Data and programming resources provided by SAS and other sources will be highlighted.

Louise Hadden first fell in love with SAS when the earth was a primordial soup. She has a special interest in graphic presentations using SAS, especially mapping. Much of her programming these days is in support of Health Economics Research at Abt Associates Inc. She is a SAS certified programmer.

What's New in Mapping

by Darrell Massengill, SAS Institute

SAS has a number of mapping technologies available. This presentation examines what is new with these technologies, focusing on SAS/GRAPH mapping. "New" means different things to each person, so the presentation will include items that may be new to you, items that are relatively new, items recently shipped or about to ship, and future items in the development pipeline.

Darrell Massengill is a Senior Development Manager in the Data Visualization R&D group and has been a development manager with SAS since 1983. He has a BS from UNC-CH and a master's degree in Computer Science from NC State University.

Google Maps and SAS/GRAPH

by Darrell Massengill, SAS Institute

Google Maps have become one of the most widely recognized means of handling maps and map data. This presentation introduces tools and shows you how to create your own Google Maps from within SAS and SAS/GRAPH. In addition, we will briefly discuss the legal and technical limitations and implications of using these maps.

Driving Distances and Times using SAS & Google Maps

by Mike Zdeb, University at Albany School of Public Health

SAS 9.2 contains new functions (e.g., ZIPCITYDISTANCE, GEODIST) that allow a user to compute geodesic distance (the shortest distance between two points on the surface of a sphere). Both functions use the Vincenty distance formula. Prior to SAS 9.2, a data step was used to compute such distances, commonly with the Haversine formula. Vincenty and Haversine distance estimates are straight line distances and there are occasions that type of estimate is what you need. There are other occasions what you want is not the straight line distance, but a driving distance. Given only one combination of locations, using Google Maps to get the driving distance and time is no problem. However, if you have a large number of location pairs, a FILENAME statement and the URL access method within SAS can be used to access Google Maps multiple times and extract both the driving distance and time each time the site is accessed.

Mike Zdeb is an assistant professor in the Department of Epidemiology and Biostatistics at the University at Albany School of Public Health in Rensselaer, NY. He has been a SAS user for 20+ years and has presented papers and conducted workshops at SUGI, SAS Global Forum, NESUG, and numerous local user groups. He has written a SAS Press book, *Mapping Made Easy Using SAS*, and has been a reviewer for a number of SAS Press books. Among the courses he teaches are a 3-credit introductory SAS class and a 1-credit advanced topics in SAS class. Mike and his teaching are the subject of a "Customer Success Story" in the area of higher education on the SAS web site.

New Policies and Procedures for Membership and Registration

For at least the duration of 2012, BASUG will not sell new memberships, or support renewal of membership fees. (We may return to annual memberships at some point in the future.) For more information on our new policy, [click here](#).

BASUG Contacts

Mailing Address:

BASUG
PO Box 170253
Boston, MA 02117

Email [Our 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.

BASUG Half-Day Training Announcement

Have you ever wanted to learn how to make informative and beautiful maps with your data? Learn how from Mike Zdeb, the SAS Books by Users author of *Maps Made Easy Using SAS*.

Please join us for this informative training, and consider coming to our morning quarterly meeting as well (separate event). For information on the morning meeting, please visit [the Quarterly Meeting Announcement](#)

BASUG is hosting an informal light buffet lunch between the morning and afternoon sessions. We hope you can join us for this opportunity to network and socialize with your fellow SAS users.

Topic	Maps Made Easy Using SAS
Summary	This course will teach you the basics of using PROC GMAP to create maps in SAS. It will then proceed to cover some more advanced topics such as adding annotations to your maps and introduce the newer map-related SAS procedures PROC MAPIMPORT, PROC GEOCODE and PROC GINSIDE. SAS users at all levels will find this course useful and informative. <i>Please see detailed topics below.</i>
Instructor	Mike Zdeb, Albany School of Public Health, University of Albany, State University of New York
When	Tuesday, June 26, 2012 1:30 – 5:30
Where	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
Price	\$155 - if paid online by June 15, 2012 \$180 - if paid online before by 5:00pm on June 25, 2012 \$195 - at-the-door – check only <i>Please see details below</i>
Audience / Prerequisites	The intended audience can range from those with no knowledge of using any SAS/Graph procedures to those who have already tried to create maps but might not have ventured to using map annotation.

How to register Individual, on-line registration required.
Pre-payment guarantees you a seat and handouts.
Please visit the [event registration page](#)

Contact If you have questions about the course, please contact [the training coordinators](#)

Course Description

You can create maps with SAS by using PROC GMAP, one of the procedures available within SAS/GRAPH. Like other SAS procedures, PROC GMAP can be used on a number of levels. At a beginning level, you can produce a number of different types of maps using very little SAS code and no procedure options. At a more advanced level, you can create maps with labeled areas and hyperlinks to other information.

This course starts with basic topics: the data needed to create maps, the various types of maps that can be produced, and some other SAS/GRAPH procedures that complement PROC GMAP. These basics will rely on using procedure options to control the appearance of maps. .

Once you understand map basics, the course will move onto map customization using the annotate facility. Map annotation is a powerful tool that allows a great deal of flexibility in map appearance, but once again the basics will be covered first before moving onto one or two more complex examples. Examples will include adding popup information that appears as you scroll across map areas and adding hyperlinks that allow you drill-down to additional information when you click on map area or symbol.

The final portion of the class will demonstrate the use of some of the newer map-related procedures. You will learn how to use PROC MAPIMPORT to create map data sets from shapefiles, how to use PROC GEOCODE to add latitude and longitude to your data sets, and given you know latitude and longitude, how to use PROC GINSIDE to determine what geographic area (for example, census tract) contains that location. At various points throughout the class, some map-related macros and functions will also be used and discussed.

Though there are a lot of topics to cover in one afternoon, you will learn to enough create maps on your own with PROC GMAP and how to, at a minimum, add area labels and symbols showing locations of cities. And remember, the course comes with a "lifetime warranty".

Instructor Bio

Mike Zdeb is an assistant professor in the Department of Epidemiology and Biostatistics at the University at Albany School of Public Health in Rensselaer, NY. He has been a SAS user for 20+ years and has presented papers and conducted workshops at SUGI, SAS Global Forum, NESUG, and numerous local user groups. He has written a SAS Press book, Mapping Made Easy Using SAS, and has been a reviewer for a number of SAS Press books. Among the courses he teaches are a 3-credit introductory SAS class and a 1-credit advanced topics in SAS class. Mike and his teaching are the subject of a "Customer Success Story" in the area of higher education on the SAS web site.

Training Registration and Payment Instructions

Please read this ENTIRE section carefully!

1. Pricing

\$ 155 if paid online by June 15, 2012
\$ 180 if paid online before by 5:00pm on June 25, 2012
\$ 195 if paid at-the-door (check only)

2. Register

INDIVIDUAL, ON-LINE REGISTRATION IS REQUIRED.

You *must register* for this training (even if you plan to pay by check at-the-door). To register and purchase tickets for the class please visit the [event registration page](#).

3. Payment Methods

Please register early! Seating and handouts are guaranteed only for pre-paid registrants.

Credit Card: We urge you to pay by credit card, using our online system. Make sure to purchase your ticket by June 15, 2012 to get the early-bird price.

Check: **Do NOT mail a check to us prior to the workshop.** Pre-register online, and bring a check with you for the at-the-door price.

4. Refund Policy

Full payment is due by the day of the class. There will be NO EXCEPTIONS. We never accept cash. We *do* accept credit card payments through our online registration service.

To receive a refund for the training, please send an email to [our training coordinators](#) by 5PM on June 15, 2012. After June 15, 2012, we will refund your payment (less a \$10 processing fee) only if we can fill your seat with other attendees.

New Policies and Procedures for Membership and Registration

For *at least* the duration of 2012, BASUG will *not* sell new memberships, or support renewal of membership fees. In essence, we are phasing out membership, at least through 2012. (We may return to annual memberships at some point in the future.) For more information on our new policy, click [here](#).

BASUG Contacts

Mailing Address:

BASUG
PO Box 170253
Boston, MA 02117

Email [Our 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.

BASUG Quarterly Meeting Announcement

We are very happy to present a meeting focused on healthcare data. And with 2 experts in the field, this meeting will be a real treat! Come to our morning meeting to learn about just a few of the healthcare databases available and also some programming techniques to clean and analyze healthcare data in general.

Please join us for these informative talks, and consider staying for the afternoon class (separate event), taught by Paul Gorrell. *For information on the training, please visit the [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	Analyzing Healthcare Data Using SAS
When	Thursday, September 20, 2012 8:15AM – Noon
Where	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 9/7/12 \$15 – if paid on-line (through 5pm 9/19/12) \$20 – at-the-door (checks only)
Contact	If you have questions about the meeting, contact the meeting organizers Asli Memisoglu and Bridget Neville

Agenda*

8:15	Sign in and Refreshments
8:45	Announcements
9:00	Data Hygiene Routines for Administrative Healthcare Data <i>by Craig Dickstein, Tamarack Professional Services, LLC</i>
9:50	Break

9:55	Calculating the Hospital Readmission Interval <i>by Craig Dickstein, Tamarack Professional Services, LLC</i>
10:55	Break
11:05	Using SAS to Generate Estimates of U.S. Prescription Drug Cost and Use <i>by Paul Gorrell, IMPAQ International</i>
12:00	Networking Lunch

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

Speaker biographies and abstracts

Data Hygiene Routines for Administrative Healthcare Data

by Craig Dickstein

How clean is your data? Do you know what a valid medical claim code should look like? This talk will discuss the architecture of the various medical coding systems as this is the most important feature of beginning to construct data hygiene routines. Several SAS routines to exercise this understanding will be discussed and demonstrated.

Calculating the Hospital Readmission Interval

by Craig Dickstein

Hospital readmission rates are an important metric in understanding utilization and quality of healthcare. This talk will describe the challenges associated with using SAS and administrative healthcare data to evaluate possible hospital readmissions. Several Base SAS programming solutions to arrive at the interval between admissions will be discussed and demonstrated.

Craig Dickstein, an independent Consultant, works with clients and select project teams to implement customized business solutions for the healthcare industry. He began his career in medical research, gaining an appreciation for the rigors of applied research methodology. A SAS user since 1978, he has significant experience managing and developing SAS applications, statistical analysis, and data management. With a Master's degree in statistics, Craig served as Director of Statistical Services with Blue Cross Blue Shield of New Hampshire. Craig has a long history of SAS user group involvement including a stint on the Executive Committee of the NorthEast SAS Users Group, Inc. Craig is a co-author of *Health Care Data and the SAS System* for SAS Publishing and a Business Knowledge Series course based on the book for SAS Education Division.

Using SAS to Generate Estimates of U.S. Prescription Drug Cost and Use

by Paul Gorrell

This talk will present an overview of the MEPS Prescribed Medicines file. This annual public-use file provides detailed information on household-reported prescribed medicines for a nationally representative sample of the civilian, non-institutionalized, population of the United States and can be used to make estimates of prescribed medicine cost and use. Each record represents one household-reported prescribed medicine that was purchased during the relevant calendar year.

The talk will focus on using SAS to generate estimates of prescription drug cost and use in aggregate, as well as by particular sources of payment (e.g. private insurance, Medicare), and by persons in selected demographic groups, or with particular conditions. Prescribed medicine variables to be discussed include national drug code (NDC), medicine name, Multum Lexicon variables, medical conditions, date variables, expenditures, sources of payments, pharmacy type, etc. In addition, the talk will discuss (i) the use of

weight variables for generating population estimates, (ii) merging with other MEPS files to obtain demographic or medical condition information, and (iii) pooling annual files to increase sample size for subpopulation or trends analysis. Also discussed will be variance estimation with MEPS. MEPS is a complex survey and the use of the SAS survey PROCs (e.g. SURVEYMEANS) is required to account for the MEPS survey design properties.

Paul Gorrell is a Principal Research Scientist and Director of Analytic Programming at IMPAQ International. His use of SAS focuses on health services research using survey and administrative data. He is active in the SAS user community, having presented numerous times at DCSUG, NESUG, and SAS Global Forum. He is on the NESUG Executive Committee and was NESUG conference co-chair in 2006 and 2011. He is also a regular Section Chair at SAS Global Forum. For the 2013 conference he is co-chairing the Statistics and Analysis section.

New Policies and Procedures for Membership and Registration

For at least the duration of 2012, BASUG will not sell new memberships, or support renewal of membership fees. (We may return to annual memberships at some point in the future.) For more information on our new policy, [click here](#).

BASUG Contacts

Mailing Address:

BASUG
PO Box 170253
Boston, MA 02117

Email [Our 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.

BASUG Half-Day Training Announcement

Just 1% of the U.S. population accounts for more than 20% of total healthcare costs, and 5% account for almost 50% of total costs.¹ In this half-day training, you can learn about the data sources underlying this type of observation, and you can learn how to use SAS survey procedures to generate these estimates and extend the analysis over medical conditions, insurance coverage & demographic information.

Please join us for this informative training, and consider coming to our morning quarterly meeting as well (separate event). *For information on the morning meeting, please visit the [Quarterly Meeting Announcement](#)*

BASUG is hosting an informal light buffet lunch between the morning and afternoon sessions. We hope you can join us for this opportunity to network and socialize with your fellow SAS users.

Topic	Analyzing U.S. Healthcare Data with SAS
Summary	This half-day class is for programmers and healthcare analysts who want to learn how to analyze data from nationally representative healthcare data released each year by the U.S. government. The goal is to provide sufficient data-source background and analytic resources so that participants can independently conduct analyses on a wide variety of topics in healthcare research. <i>Please see detailed topics below.</i>
Instructor	Paul Gorrell, IMPAQ International, LLC
When	Thursday, September 20, 2012 1:00pm – 5:00pm
Where	Microsoft New England Research and Development Center (NERD) ² One Memorial Drive Conference Center, First Floor Cambridge, MA 02142 857-453-6000
Directions	Please visit the meeting site directions page
Price	\$155 - if paid online by September 7, 2012 \$180 - if paid online before 5:00pm September 19, 2012 \$195 - at-the-door – check only <i>Please see details below</i>
Audience / Prerequisites	Level of SAS programming: - Data step familiarity - Using PROCs to generate descriptive statistics
Benefits	• Learn how to download & analyze national healthcare data files containing medical, cost, and demographic & insurance information • Learn the use of SAS survey procedures

How to register	Individual, on-line registration required. Pre-payment guarantees you a seat and handouts. Please visit the event registration page
Contact	If you have questions about the course, please contact the training coordinators

Course Description

A central component of discussions of healthcare reform in the U.S. are estimations of healthcare cost and use at the national or State level, as well as for subpopulation analyses for individuals with certain demographic properties or medical conditions. For example, a striking, but persistent, observation is that just 1% of the U.S. population accounts for more than 20% of total healthcare costs, and 5% account for almost 50% of total costs. Participants in this half-day training will not only learn about the data sources underlying this type of observation, but also how to use SAS to generate these estimates and to extend the analysis in various ways, e.g. to investigate costs for people with chronic conditions or racial/ethnic disparities in access to care.

This half-day class is for programmers and healthcare analysts who want to learn how to analyze data from nationally representative healthcare data released each year by the U.S. government. The goal is to provide sufficient data-source background and analytic resources so that participants can independently conduct analyses on a wide variety of topics in healthcare research. For selected examples, such as the estimates above, we will concretely show how to download the data from Federal Websites, replicate published estimates, and then extend the analysis.

Course Details

General discussion topics will include:

- An overview of publicly available healthcare data and its applicability to healthcare research topics
- Downloading data files and creating SAS analytic datasets
- The type of information (variables) available:
 - o Cost (by person, by condition, by type of medical event)
 - o Use (by person, by condition, by type of medical event)
 - o Medical conditions
 - o Insurance coverage
 - o Demographic Information (socio-economic status, race/ethnicity)
- The use of survey weights for population estimates
- Variance estimation for estimates from complex (i.e. non-random) surveys
- The use of SAS survey procedures
- Sample size issues (including strategies for pooling data years to increase sample size)

Using concrete examples, we will demonstrate how to generate national and State estimates from the following data sources (and show how to replicate published estimates):

- Behavioral Risk Factor Surveillance System (BRFSS)
 - o State-level estimates of mammogram rates for women 40 and over
 - o State-level estimates of diabetes prevalence
- Medical Expenditure Panel Survey (MEPS)
 - o National estimates of healthcare expenditures, overall, and for specific populations such as the top 1% and 5% of the population, ranked by healthcare expenditures
 - o National estimates of Medicare expenditures for prescribed medicines before and after implementation of Medicare Part D (the prescription drug benefit)
 - o Comparisons with Medicare claims data
- National Health Examination and Nutrition Survey (NHANES)
 - o Serum Vitamin D levels of the U.S. population by age and race/ethnicity

Throughout, the emphasis will be on illustrating general and specific properties of the data sources being discussed, as well as use of the SAS DATA step and PROCs as they relate to healthcare research goals. The format of the training will be lecture/discussion, with the emphasis on discussion.

Instructor Bio

Paul Gorrell is a Principal Research Scientist and Director of Analytic Programming at IMPAQ International. His use of SAS focuses on health services research using survey and administrative data. He is active in the SAS user community, having presented numerous times at DCSUG, NESUG, and SAS Global Forum. He is on the NESUG Executive Committee and was NESUG conference co-chair in 2006 and 2011. He is also a regular Section Chair at SAS Global Forum. For the 2013 conference he is co-chairing the Statistics and Analysis section.

Training Registration and Payment Instructions

Please read this ENTIRE section carefully!

- | | |
|-------------|---|
| 1. Pricing | \$ 155 if paid online by September 7, 2012
\$ 180 if paid online before 5:00pm September 19, 2012
\$ 195 if paid at-the-door (check only) |
| 2. Register | INDIVIDUAL, ON-LINE REGISTRATION IS REQUIRED. |

You *must* register for this training (even if you plan to pay by check). To register and purchase tickets for the class please visit the [event registration page](#).

- | | |
|--------------------|--|
| 3. Payment Methods | <p>Please register early! Seating and handouts are guaranteed only for pre-paid registrants.</p> <p><u>Credit Card</u>: We urge you to pay by credit card, using our online system. Make sure to purchase your ticket by 09/07/2012 to get the early-bird price.</p> |
|--------------------|--|

Check: **Do NOT mail a check to us prior to the workshop.** Pre-register online, and bring a check with you for the at-the-door-price.

Full payment is due by the day of the class. There will be NO EXCEPTIONS. We never accept cash. We *do* accept credit card

4. Refund Policy

payments through our online registration service.

To receive a refund for the training, please send an email to [our training coordinators](#) on 09/07/2012. After 09/07/2012, we will refund your payment (less a \$10 processing fee) only if we can fill your seat with other attendees.

New Policies and Procedures for Membership and Registration

For *at least* the duration of 2012, BASUG will *not* sell new memberships, or support renewal of membership fees. In essence, we are phasing out membership, at least through 2012. (We may return to annual memberships at some point in the future.) For more information on our new policy, click [here](#).

BASUG Contacts

Mailing Address:

BASUG
PO Box 170253
Boston, MA 02117

Email [Our Webmaster](#)

- (1) Stanton MW, Rutherford MK. The high concentration of U.S. health care expenditures. Rockville (MD): Agency for Healthcare Research and Quality; 2005. Research in Action Issue 19. AHRQ Pub. No. 06-0060 http://meps.ahrq.gov/mepsweb/data_files/publications/ra19/ra19.pdf Accessed August 9, 2012
- (2) 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.

BASUG Quarterly Meeting Announcement

Got Data? Big or small, we've got tips and techniques for you! Our Q4 meeting is jam-packed with two full-length presentations, a coders corner, and an open forum discussion you won't want to miss. Please join us for this informative meeting.

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	Tips and Techniques for Data Big and Small
When	Tuesday, December 11, 2012 8:15AM – Noon
Where	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, PLEASE.
Payment	\$10 – if paid on-line by Friday, November 30th \$15 – if paid on-line (through 5pm Monday, December 10 th) \$20 – at-the-door (checks only)
Contact	If you have questions about the meeting, contact the meeting organizers Lori Goldman and Brian Saper

Agenda*

8:15	Sign in and Refreshments
8:45	Announcements
9:00	PROC DATASETS; The Swiss Army Knife of SAS® Procedures <i>by Michael Raithe, Westat</i>
10:00	Break
10:15	Common Sense Tips and Clever Tricks for Programming with Extremely Large SAS® Data Sets <i>by Kathy Hardis Fraeman, United BioSource Corporation</i>
11:05	Break
11:20	A Quick View of SAS Views <i>by Elizabeth Axelrod, Abt Associates Inc.</i>

11:30	Open Forum Discussion on How to Become a Top SAS Programmer <i>Michael Raithel, Westat and Kathy Hardis Fraeman, United BioSource Corporation</i>
12:00	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

PROC DATASETS; The Swiss Army Knife of SAS® Procedures

by Michael Raithel, Westat

The DATASETS procedure provides the most diverse selection of capabilities and features of any of the SAS procedures. It is the prime tool that programmers can use to manage SAS data sets, indexes, catalogs, etc. Many SAS programmers are only familiar with a few of PROC DATASETS's many capabilities. Most often, they only use the data set updating, deleting, and renaming capabilities. However, there are many more features and uses that should be in a SAS programmer's toolkit.

This paper highlights many of the major capabilities of PROC DATASETS. It discusses how it can be used as a tool to update variable information in a SAS data set; provide information on data set and catalog contents; delete data sets, catalogs, and indexes; repair damaged SAS data sets; rename files; create and manage audit trails; add, delete, and modify passwords; add and delete integrity constraints; and more. The paper contains examples of the various uses of PROC DATASETS that programmers can cut and paste into their own programs as a starting point. After reading this paper, a SAS programmer will have practical knowledge of the many different facets of this important SAS procedure.

Michael A. Raithel is a Senior Systems Analyst for Westat, a leading research firm headquartered in Rockville, Maryland. He has been using SAS software for over 25 years and is an internationally recognized expert on using SAS in mainframe and UNIX environments. Michael has authored over twenty-five SAS technical papers and is a popular lecturer at SAS Global Forum, regional SAS conferences, and at local SAS User Group meetings. He has taught SAS classes at Westat, at American University, at SAS Global Forum meetings, and at SAS headquarters in Marlow, England. Michael has been a section chair at SUGI, SESUG, and NESUG, and he co-chaired NESUG in 1995. He is a frequent poster to the SAS-L listserv and a contributor to SASCommunity.org. Michael has written three books on SAS software; the most current one being ***The Complete Guide to SAS Indexes***. A copy of his first SAS book, ***Tuning SAS Applications in the MVS Environment***, resides in the Smithsonian Institution of American History's Permanent Research Collection of Information Technology.

Common Sense Tips and Clever Tricks for Programming with Extremely Large SAS® Data Sets

by Kathy Hardis Fraeman, United BioSource Corporation

Working with extremely large SAS data sets – where the numbers of observations are in the hundreds of millions – can pose many challenges to the SAS programmer. Simple SAS code, including programs that involve basic DATA step programming and ordinary PROC SORTs, can become extremely difficult to run on such large data sets. Encountered problems range from extremely long run times, exceeding available memory, and seemingly unexplainable and random I/O errors that cause previously working programs to crash unexpectedly. This paper discusses common sense tips and clever tricks developed in Base SAS to improve the programming process with these extremely large SAS data sets. SORT/MERGEs on large data sets can be avoided with a trick with using SAS formats and hash tables, and the SAS options SGIO (Scatter-read/Gather-write Input Output) and BUFNO can be combined to avoid I/O errors that sometimes occur when reading extremely large SAS data sets. These and other programming tips and techniques were developed over many years' experience of writing program to analyze extremely large SAS

databases consisting of medical insurance claims, inpatient hospitalization, and electronic medical records (EMR) data.

Kathy Fraeman, SM, is the Director of Data Analytics for UBC's Epidemiology and Database Analytics (EDA) group. Ms. Fraeman has over 25 years of experience in analytic SAS programming and data management for biomedical research studies, including pharmaceutical clinical trials, epidemiologic studies, survey research, cost-of-drug-treatment studies, and drug safety and marketing studies. She is experienced in analyzing large and complex relational health care databases, including insurance claims files. Ms. Fraeman has co-authored several publications in the fields of epidemiology and health care issues. Before joining UBC, Ms. Fraeman was a self-employed independent SAS contractor who worked on over a dozen projects with UBC. She received her Bachelor of Science in Biology from the Massachusetts Institute of Technology and her Master's of Science in Environmental Health Sciences from the Harvard University School of Public Health.

A Quick View of SAS Views

by Elizabeth Axelrod, Abt Associates Inc.

SAS Views are a handy technique to have in your toolkit – especially if you work with large datasets. They enable you to pre-process your data – without actually creating a dataset. This Coders Corner presentation will introduce you to SAS Views, and provide some usage guidelines... so you'll always have a great view!

Elizabeth Axelrod is a Lead Programmer Analyst at Abt Associates Inc. in the Health Policy Research Division. She has been using SAS for over 30 years, and has been a SAS fan(atic) for almost that long. She is currently the President of BASUG - the Boston Area SAS Users Group.

Open Forum Discussion on How to Become a Top SAS Programmer

Moderated by Michael Raithel, Westat

BASUG is going to take the opportunity of having two top SAS programmers as guests to hold a discussion on how to become a top SAS programmer. Michael Raithel will kick off this discussion with a brief overview of his upcoming SAS Press book: *How to Become a Top SAS Programmer*. Kathy Fraeman will give an overview of her own evolution of becoming a top SAS professional. Then, Michael and Kathy will answer questions from the audience in an open forum. You are sure to leave this presentation with some new ideas you can use to grow your own SAS programming career.

Policies and Procedures for Membership and Registration

For *at least* the duration of 2012, BASUG will *not* sell new memberships, or support renewal of membership fees. (We may return to annual memberships at some point in the future.) For more information on our policy, [click here](#)

BASUG Contacts

Mailing Address:

BASUG
PO Box 170253
Boston, MA 02117

Email [Our 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.