

BASUG Quarterly Meeting Announcement

We are thrilled with the response we got to our Call-for-Papers for our Q1 Coders' Corner meeting. This meeting will feature many presenters – some first-timers, some seasoned – covering a variety of topics. Please join us to learn lots of tips 'n techniques from your fellow BASUG colleagues.

Topic: Coders' Corner 2008

When: Friday, March 28, 2008, 8:15 AM – Noon

Where: **Holiday Inn Brookline**

1200 Beacon Street
Brookline, MA 02446
(617) 277-1200

For directions, please visit: <http://www.basug.org/directions.html>

How: Individual, **On-Line Registration Required**. No Email!

To register: Please visit: <http://www.basug.org/register.php3>

Contact: If you have questions about the meeting contact Sally Skinner
(sally.skinner@tufts.edu)

Agenda*

8:15	Sign in and Refreshments
8:45	Announcements
9:00	PUT and INPUT functions: Making more use of formats and informat by James Zeitler, Harvard Business School
9:12	Using SAS® Graph to produce four graphs per page by Amanda Tweed, Millenium Pharmaceuticals
9:24	Population Pyramid Graphs by Meena Doshi, New England Research Institutes
9:36	Enterprise Guide: Tips for Programmers by John D. Chapman, PhD, Markcelian Analytics, Inc.
9:48	BREAK
10:10	Exporting the Data to EXCEL? Order the columns the way you want them! by Mark Martin, Siemens Healthcare Diagnostics
10:22	Tricks to create report quality WORD tables and figures in SAS® by Brendan Bartley, CBAR, Harvard School of Public Health
10:34	Compressing Output files in a SAS® job (on UNIX) by Leslie Somos, Paraxel
10:46	Writing SAS® Functions Without SAS® Toolkit by Jerry Lewis, Biogen Idec
10:58	BREAK
11:15	Comparing Values Across Observations: Be Careful by Prafulla Girase, IMS Health
11:27	How to automatically run SAS® procedures on a group of datasets within a single directory by Diana Ventura, CBAR, Harvard University
11:39	Go Where Few Have Gone Before (learn to use SAS® metadata) by Steven Ezzy, Ingenix

11:51	Closing remarks & announcements
Noon	Meeting Adjourned

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

Abstracts and Bios

Tricks to create report quality WORD tables and figures in SAS®

by Brendan Bartley

This paper describes some macros and tricks to enhance RTF output for reports. The first macro allows the user to move tables within a document and have the table numbering change corresponding to their location in the document. It also sets up bookmarks for reference in a report. Another macro eliminates all but the first section breaks in a RTF document. The last macro gives a table its own style name so that its formatting does not change when the formatting of the text in the document changes.

Brendan Bartley is a programmer in clinical trials at the Center for Biostatistics in AIDS Research (CBAR) at the Harvard School of Public Health. Brendan has been using SAS® software in health research since 2001.

Enterprise Guide: Tips for Programmers

by John D. Chapman, PhD

SAS® marketing of Enterprise Guide (EG) emphasizes its graphical nature, with tasks and wizards that offer “point-and-click interface to the full power of SAS.” But EG also has many features that make it an attractive alternative to the conventional Windows interface for programming. This Coders’ Corner presentation will present several such features and seek to motivate programmers who aren’t using EG to give it a try.

John D. Chapman provides data analysis and reporting services to the health care industry. SAS® is a valued tool in his arsenal. A long-time user of the SAS® Windows interface, he became an EG convert with the release of version 4.1 and would never choose to go back.

Population Pyramid Graphs

by Meena Doshi

The most important characteristics of a population—in addition to its size and the rate at which it is expanding or contracting—are the ways in which its members are distributed according to age, sex, ethnic or racial category, and residential status (urban or rural). Perhaps the most fundamental of these characteristics is the age distribution of a population. SAS® can be used to create these population pyramids that can depict the age and sex distribution of the population.

Meena Doshi is a Statistician at New England Research Institutes, Inc located in Watertown MA. She has a Masters in Applied Statistics from The Ohio State University and a Masters in Public Health (Epidemiology and Biostatistics) from Tufts University. She has worked as a Marketing Analyst with a startup in Boston and as a Statistical Analyst at Tufts University.

Go Where Few Have Gone Before (learn to use SAS® metadata)

by Stephen Ezzy

As has often been said, there are many ways to accomplish the same thing in SAS®. Many SAS® users are unaware of the ‘data about their data’ (better known as metadata) that the SAS® system maintains in the SASHelp library. This presentation suggests ways to use these data to think outside the box in terms of accomplishing SAS® tasks. (It may even suggest a few tasks that might not be possible any other way!)

Steve Ezzy has been with Ingenix in Waltham MA since May, 2005 and has 19 years

experience in SAS® programming. He also has some expertise in SQL Server, Visual Basic and geographic information systems (GIS). Steve has used SAS® to solve problems in the areas of public health, environmental epidemiology, clinical trials and fishery biology. He even wrote a SAS® program to solve sudoku puzzles and was pretty pleased with himself about that until he discovered that every Tom, Dick and Harry on the SAS-L web site has done the same thing.

Comparing Values Across Observations: Be Careful

by Prafulla Girase

There are several techniques that can be used to compare values across observations in SAS®. It is important to understand how each technique works in order to pick the most appropriate technique and get correct results. This paper discusses three of the techniques that can be used. The techniques discussed in this paper include the LAG function, one-to-one reading, and one-to-one merging. To illustrate appropriate usage and highlight advantages of each technique, a business question is proposed and answered using all three techniques. After reading this paper, you will be better prepared to select the most appropriate method in instances where values must be compared across observations. The intended audience for this paper is someone with a basic understanding of BASE SAS®.

Prafulla Girase is a Statistical Programmer at IMS Health in Watertown, MA. He has been using SAS® software with various data including epidemiology, clinical trials, & medical claims since 2000.

Writing SAS® Functions Without SAS® Toolkit

by Jerry Lewis

Where a frequently repeated calculation can be written in closed form, a parameterized SAS® macro can be written to produce the non-terminated SAS® expression. Since the expression is non-terminated, the macro effectively behaves like a SAS® function (that can be used within a larger expression in a SAS® data step). As an example, a macro function for exact Clopper-Pearson binomial confidence limits will be discussed (published implementations, such as Daly [1992] used a dedicated data step). A seed list of some other useful expressions (exact Poisson confidence limits, 1-sided normal tolerance factors, Wald-Wolfowitz approximate 2-side normal tolerance factors, etc) will be given to stimulate thinking about other possible applications.

Jerry Lewis is a Lead Biostatistician (nonclinical) at Biogen Idec in Cambridge. He has been using SAS® in academia and industry since 1976.

Exporting the Data to EXCEL? Order the columns the way you want them!

by Mark Martin

Do you spend time re-ordering the variable columns in Excel after exporting data from SAS® datasets? You can specify the variable order before Proc EXPORT in either of two ways: the DATA step RETAIN statement, or the Proc SQL SELECT statement. Code examples will show how to do this. You'll also learn subtle Proc Export differences between SAS® 8 and SAS® 9 for Excel output.

Mark Martin is a Senior Biostatistician at Siemens Healthcare Diagnostics in Norwood MA. His internal clients include product development scientists, who often like to use Excel for graphics and data exchange. Mark has been using SAS® since 1990.

Compressing Output files in a SAS® job (on UNIX)

by Leslie J. Somos

When your disk space is limited, here is one way to write your flat file from SAS® in an already-compressed format, instead of writing it out and then compressing it outside of SAS® (which requires space for both the original file and the compressed version on disk simultaneously).

Leslie has used SAS® for many years, on many platforms. He misses the documentation for Version 5, when it all fit into a single (~2-inch thick) book. At Boston University School of Public Health, he recently earned a Master of Public Health with concentration in Biostatistics. He currently works at PAREXEL in Lowell.

Using SAS® Graph to produce four graphs per page

by Amanda Tweed

SAS® Graph provides many options for enhancing graphical display. While simple bivariate plots can be generated using very basic code, producing numerous graphs based on data subsets can lead to an overwhelming volume of graphical output that is cumbersome to review. Condensing the output facilitates comparison between plots, aids in the review of voluminous output and reduces paper use. PROC GREPLAY allows the programmer to define a template and position multiple graphs per page. Used in conjunction with the ODF RTF destination, the code is not complex. This session demonstrates the use of the GREPLAY procedure to display four separate graphs created

using PROC GPLOT on a single page and highlights the benefits of this form of output.

Amanda Tweed is a Statistical Programmer at Millennium Pharmaceuticals in Cambridge. She has been in industry for nearly 3 years. Prior to that time, she worked in academia—both at Mass General Hospital and the University of Pennsylvania. She has been a SAS® user for four years, with prior experience programming in STATA. Currently, she is learning to program in R.

How to automatically run SAS® procedures on a group of datasets within a single directory

by Diana Ventura

Sometimes it is necessary to analyze a group of datasets. In this case it makes sense to create a macro variable that represents the dataset names. The macro variable allows SAS® procedures to be run on a group of datasets without typing the names of each dataset, this streamlines repetitive code and eliminates needless typing. It might seem complicated to capture the name of datasets in a selected directory into a macro variable but the process actually requires only the use of the contents option within the Proc Datasets procedure, the use of call symput to put the variable into a macro variable, and setting up a iterative macro do loop to initiate use of the macro. All of this can be accomplished in a few lines of code.

Diana Ventura is the Head of Standardization for the Center for Biostatistics in AIDS Research, where she manages meta-data analysis, clinical data quality assurance, and CRF forms programming. She has been using SAS® in the public health arena since 1992.

PUT and INPUT functions: Making more use of formats and informats

by James Zeidler

Using the INPUT statements with informats is the standard way to read into SAS® from an external data file. And the PUT statement with formats is the analogous way to write from SAS® to an external data file. In this talk we'll see how to use informats and formats with the INPUT and PUT functions to transform variables in a SAS® dataset. It's a

good thing.

James Zeitler is a Research Database Analyst at Harvard Business School, working primarily to assist faculty doing research in a variety of business-related disciplines. He has been using SAS® to do this sort of thing for at least twenty years.

BASUG Membership

Keep your BASUG Membership up-to-date! Print out a membership form from our website, and bring it with you to the meeting (along with your check). Please note that we have discontinued Corporate Memberships – we now provide only Individual Membership at \$30 annually. For more information on our membership policy, or to print out a form visit: <http://www.basug.org/basugj.shtml>

Directions and Parking

Please visit: <http://www.basug.org/directions.html>

BASUG Contacts

Mailing Address:

BASUG
PO Box 253
Boston, MA 02117

To email our Webmaster: basugwm@basug.org

BASUG Quarterly Meeting Announcement

SAS® 9.2 is out! Please join us for this informative meeting on all the cool new stuff in this new version. We are delighted to host Jason Secosky, a current developer in R&D at SAS, and Chris Olinger, an original developer of ODS.

Topic: New Features in SAS® 9.2

When: Wednesday, May 7, 2008, 8:15 AM – Noon

Where: **Holiday Inn Brookline**

1200 Beacon Street

Brookline, MA 02446

(617) 277-1200

For directions, please visit: <http://www.basug.org/directions.html>

How: Individual, **On-Line Registration Required**. No Email!

To register: Please visit: <http://www.basug.org/register.php3>

Contact: If you have questions about the meeting contact Lori Goldman (Lori.Goldman@ingenix.com) or Asli Memisoglu (asli_memisoglu@comcast.net)

Agenda*

8:15 - 8:45	Sign in and Refreshments
8:45 - 9:00	Announcements
9:00 - 09:50	User-Written Data Step Functions, by Jason Secosky, SAS Institute
09:50 - 10:00	Leg stretch
10:00 - 10:45	Changes and Enhancements to ODS for 9.2 by Chris Olinger, d-Wise Technologies
10:45 - 11:05	Break
11:05 - 12:00	A Sampler of What's New in Base SAS 9.2, by Jason Secosky, SAS Institute

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

Abstracts and Bios

User-Written Data Step Functions

By Jason Secosky, SAS Institute

For years, users have asked for the ability to write functions in DATA step syntax. The day has come when this is possible. In SAS®9, the FCMP procedure provides the ability to write functions using DATA step syntax. In SAS®9.2, these functions can be called from a DATA step. This paper explains how to write functions with PROC FCMP and how these functions are invoked from a DATA step. It also contains caveats when calling PROC FCMP functions from a DATA step.

A Sampler of What's New in Base SAS® 9.2

By Jason Secosky, SAS Institute

This presentation shows how coding with SAS is easier than ever with SAS 9.2. We highlight the top new features and performance improvements in DATA step, PROC SQL, and PROC SORT. Included are several "Top 10" SASware Ballot items, writing functions with DATA step syntax, and more intuitive and culturally acceptable sorting with PROC SORT. Come see what's new in the next version of Base SAS.

Jason Secosky is a Principal Software Developer at SAS and is responsible for extending and maintaining the DATA step language. He has been at SAS since 1997. Prior to working at SAS, Jason received a B.S. in Computer Engineering from the University of California, San Diego and a M.S. in Computer Science from the University of Washington. Jason is interested in discussing how you use SAS and the DATA step.

Changes and Enhancements to ODS for 9.2

By Chris Olinger, d-Wise Technologies

Version 9.2 of the SAS system provides some long awaited features to the ODS system. This talk will focus on some of the major changes available in the release to make the SAS programmer's life a little easier such as CSS support, better symbol control, retirement of the REPLACE statement, and more.

Chris Olinger, President of d-Wise Technologies, has been developing enterprise software for over 20 years. While employed at SAS Institute he was the manager of the Base SAS reporting group and lead developer of ODS as well as being involved in the design and development of the SAS Drug Development platform. More recently, Chris has been involved in the development of SAS and web based applications within the framework of his software company.

BASUG Membership

Keep your BASUG Membership up-to-date! Print out a membership form from our website, and bring it with you to the meeting (along with your check). Please note that we have discontinued Corporate Memberships – we now provide only Individual Membership at \$30 annually. For more information on our membership policy, or to print out a form visit:

<http://www.basug.org/basugj.shtml>

Directions and Parking

Please visit: <http://www.basug.org/directions.html>

BASUG Contacts

Mailing Address:

BASUG
PO Box 253
Boston, MA 02117

To email our Webmaster: basugwm@basug.org

Advanced ODS!

BASUG Half-Day Training Session

Topic: Using the Advanced Features of the SAS Output Delivery System
Learn it from the person who developed ODS!

Summary: This seminar will cover topics relating to getting the most out of ODS Styles, reporting techniques using the Base SAS reporting procedures, ODS Tagsets, and using SAS to interface with alternative reporting environments. It will also touch on some of the newer technology such as Statistical Graphics and ODS enhancements to make life just a little easier.

Instructor: Chris Olinger of d-Wise Technologies

Chris Olinger, President of d-Wise Technologies, has been developing enterprise software for over 20 years. While employed at SAS Institute he was the manager of the Base SAS reporting group and lead developer of ODS as well as being involved in the design and development of the SAS Drug Development platform. More recently, Chris has been involved in the development of SAS and web based applications within the framework of his software company.

When: Wednesday, May 7, 2008, 1:00-5:00

Price: \$125 (if paid by check postmarked by Thursday, April 24, 2008)
\$150 (paid at-the-door)
See details below

Audience: This course is intended for SAS programmers and others writing reports using SAS.

Prerequisites: The attendees should have a basic knowledge of ODS.

Benefits: Students will have the opportunity to meet the person who was responsible for the development and design of the ODS portion of the SAS system. Bring your questions about ODS topics and get them answered by the person with firsthand knowledge of how it functions. Ask questions and learn how to accomplish the output format that has you stumped. Learn about new features that can make tasks easier or reusable for additional projects.

Contact: If you have questions about the course contact: Bob Virgile (rvirgile@comcast.net) or Brian Saper (briansaper@yahoo.com)

Registration: ***Pre-registration is required. Please see detailed information below.***

Course Description

This half-day advanced ODS seminar will cover topics relating to getting the most out of ODS Styles, including understanding how to create and manage styles, the difference between REPLACE and STYLE inheritance, and how to build a style from scratch. Also covered are reporting techniques using the Base SAS reporting procedures, specifically around the use of inline styles and references via PROC REPORT. We will dedicate at least an hour to ODS Tagsets and their possible uses as an interface with alternative reporting environments, as well as some unusual uses for alternative programming.

Topics

- Advanced Features of ODS
 - ODS Styles
 - Reporting Techniques using the Base SAS reporting procedures
 - ODS Tagsets
 - Using SAS to interface with alternative reporting environments
 - Statistical Graphics
 - ODS enhancements that make life a little easier
-

Registration and Payment Instructions for the Course

Please read this entire section carefully!

1. You **must pre-register** for this training, by going to this site:
<http://www.basug.org/register.php3>

INDIVIDUAL, ON-LINE PRE-REGISTRATION IS REQUIRED FOR TRAINING. NO EMAIL!

All registrants will receive an immediate email confirmation of their pre-registration. Within 72 hours you will then receive another e-mail, with a customized invoice for the training, based on your membership and registration status.

Seats are limited, so register early! Seating is guaranteed only for **pre-paid** registrants.

2. Complete the invoice form and mail it along with your check, postmarked by Thursday, April 24, 2008 to our Treasurer at:

**BASUG
PO Box 253
Boston, MA 02117**

3. Note that this class is open to BASUG members only. This means that your 2008 \$30 individual membership dues must be paid up in order to attend! If you are not currently a member, your invoice will include the \$30 fee for 2008 BASUG membership dues.

If you have any questions about your BASUG membership status, please contact our Membership Coordinator (Robert Rosofsky) at: Robert.Rosofsky@Verizon.net

4. **Please do not mail any checks after Thursday, April 24, 2008.** If you miss the mail-in deadline, bring your completed invoice form and a check made payable to BASUG to cover the at-the-door tuition rate of \$150, plus any \$30 membership fees due, as specified in your invoice form. For walk-ins, we will only be able to accommodate people on a first-come, first-served basis until seats are filled.

Full payment is due by the day of the class. There will be NO EXCEPTIONS. As a small voluntary organization, we cannot accept credit cards at this time. We also will not handle cash, so checks only please.

- 5 **REFUND POLICY:** To receive a full refund, notify either of the meeting coordinators (see "Contact" above) by 5PM on April 24, 2008. ***After April 24, we will refund your payment only if we can fill your seat with attendees at the door.***

MEMBERSHIP

Keep your BASUG Membership up-to-date! Print out a membership form from our website, and bring it with you to the meeting (along with your check). Please note that we have discontinued Corporate Memberships - we now provide only Individual Membership at \$30 annually. For more information on our membership policy, or to print out a form visit:

<http://www.basug.org/basugj.shtml>

DIRECTIONS TO THE TRAINING

Directions to the Brookline Holiday Inn can be found at <http://www.basug.org/directions.html>.

BASUG CONTACTS

BASUG's Mail Address: BASUG PO Box 253 Boston, MA 02117

To email our Webmaster: basugwm@basug.org

BASUG Quarterly Meeting Announcement

Learn how to get more from your data! Please join us for an informative meeting on data modeling and exploration.

We are delighted to host **Daniel Bauer**, Associate Professor of Psychology in the LL Thurstone Psychometric Laboratory at the University of North Carolina at Chapel Hill, **Dominique Haughton**, Professor of Mathematical Sciences at Bentley College, **Maria Skaletsky** from the Academic Teaching Center at Bentley College, and **Timothy C. D'Auria** from Creative Computing, Inc.

Topics: Multilevel Modeling in SAS
SAS Text Miner: Introduction and Case study

When: Wednesday, September 24, 2008, 8:45 AM – Noon

Where: **Holiday Inn Brookline**
1200 Beacon Street, Brookline, MA 02446, 617-277-1200

For directions, please visit: <http://www.basug.org/directions.html>

How: **Individual, On-Line Registration Required.** No e-mail please!

To register: Please visit: <http://www.basug.org/register.php3>

Contact: If you have questions about the meeting contact:
Karen Olson (Karen.Olson@childrens.harvard.edu) or
Eric Hamann (ehamann@wyeth.com)

Agenda

- 8:15 - 8:45 Sign-in and refreshments
- 8:45 - 9:00 Announcements
- 9:00 - 10:15 **Multilevel Modeling in SAS**, Daniel Bauer, University of North Carolina at Chapel Hill
- 10:15 - 10:35 Break
- 10:35 - 11:30 **SAS Text Miner: Introduction and Case study**, Dominique Haughton and Maria Skaletsky, Bentley College
- 11:30 - 12:00 **Using SAS Text Mining to Save Lives**, Timothy D'Auria, Creative Computing, Inc.

Note: Times are approximate. Please re-visit the BASUG website (www.basug.org) for updated information.

Abstracts and Bios

Multilevel Modeling in SAS

Daniel Bauer, PhD, University of North Carolina at Chapel Hill

Most statistical models assume independent observations, yet this assumption is often not met in practice. Dependence commonly arises when data are clustered or collected repeatedly over time. For instance, for a given complaint, patients who see the same physician will likely receive more similar treatment than patients who see different physicians, resulting in patient data that is positively correlated within physicians. Such dependence compromises the statistical tests provided by classical modeling approaches, often leading to spuriously significant effects. In contrast, multilevel models (a subclass of mixed models) are specifically designed to model these types of dependent data structures, providing accurate statistical tests and additional information about the outcome under study. This talk will introduce both the key concepts of multilevel modeling and the SAS procedures that can be used to fit linear (MIXED), generalized linear (GLIMMIX), and nonlinear (NLMIXED) multilevel models.

Daniel Bauer, PhD, is an Associate Professor of Psychology in the LL Thurstone Psychometric Laboratory at the University of North Carolina at Chapel Hill. His research focuses on the quantitative modeling of

adolescent behavioral health data, including longitudinal data on alcohol and substance abuse, antisocial and criminal behavior, and risky sex. His areas of specialization include multilevel models, structural equation models, and mixture models. In addition to teaching classes in these topics at the university, he developed and delivered a 3-day workshop titled "Multilevel Modeling of Hierarchical and Longitudinal Data Using SAS" through the SAS Education Business Knowledge Series from 2006 to 2008.

SAS Text Miner: Introduction and Case Study

Dominique Haughton, Bentley College and Toulouse School of Economics, and Maria Skaletsky, Bentley College

An Introduction to SAS Text Miner is presented in the context of an analysis of free-text responses to a questionnaire administered to 416 undergraduates in a business university about their perceptions of unfair grading practices. This talk will demonstrate how to cluster responses with SAS Text Miner, and in the process, identify preliminary themes within the responses. Different features of SAS Text Miner will be described as we go along.

A SAS Text Miner Case Study is presented. This study involves a text analysis performed in the context of the database marketing group of a leading provider of information infrastructure. At several contact points, customers provided their job title, but have not always filled in the field "Key Player Description", which consists of a brief (a few words) job description. The problem here is to impute the missing brief job description for a file of about a million such customers on the basis of available job titles. This talk will demonstrate how this problem was addressed with SAS Text Miner, using a sample of 10,000 customers.

Dominique Haughton, PhD, is a Professor of Mathematical Sciences at Bentley in Waltham, Massachusetts, near Boston. Major areas of interest are applied statistics, statistics and marketing, the analysis of living standards surveys, data mining, and model selection. She is currently researching multilevel models and living standards in Vietnam, clustering methods, models of gambling expenditures and social networks, notably in Senegal. Dr. Haughton is the Editor-in-chief of Case Studies in Business, Industry and Government Statistics (CSBIGS) and Co-editor of The Vietnamese Household: Explorations Using the Living Standards Measurement Survey (1992-1993) and Health and Wealth in Vietnam: An Analysis of Household Living Standards (1998).

Over forty of her articles have appeared in journals such as The American Statistician, BizEd, Telecommunications Policy, Journal of Higher Education Management, Computational Statistics and Data Analysis, Journal of Interactive Marketing, Economic Development and Cultural Change, Studies in Family Planning, Journal of Population Economics, Journal of Biosocial Science, Annals of Statistics, Sankhya, Journal of Statistical Computation and Simulation, Communications in Statistics, Statistica Sinica. She has four current PhD students in Business Analytics.

Maria Skaletsky, MBA, has focused her research on data mining, in particular, web usage and web content mining. Since 2003, she has held the position of research consultant at Bentley's Academic Technology Center. She is responsible for consulting faculty in all aspects of research approach, design, and implementation, in the selection of appropriate methodology for data gathering and statistical analysis. She also assists faculty in the use of large research databases such as CRSP, Compustat, and Execu Comp, as well as the use of various quantitative and qualitative software such as SAS, SPSS, NVivo and other. Before joining the ATC, Maria received her MBA with a concentration in Business Data Analysis from Bentley College.

Using SAS Text Mining to Save Lives

Timothy C. D'Auria, Creative Computing, Inc.

What if there was a way for a computer to read clinician notes and predict patients most at risk for serious, life-threatening complications before they happen? Unstructured textual data often comprises the vast majority, upwards of 80% in cases, of information available for organizational analysis, particularly in the healthcare industry. During this presentation, we will explore how unstructured data can be incorporated into analyses alongside structured sources to predict future patient emergency room usage. We will then discuss why text mining may be a preferable alternative to structured healthcare data analysis using logistic regression and other statistical procedures. The presentation will conclude with a summary of how SAS text mining capabilities can ultimately be leveraged to save patient lives.

Timothy C. D'Auria leads the SAS Analytics Practice at Creative Computing, Inc., a SAS Silver Consulting Partner. As a patented and published author with distinction in predictive modeling technologies, Timothy's interest is in the application of data mining to improve patient care and achieve competitive advantage in healthcare industries. His work in business data mining has been featured in leading industry publications

and has been incorporated into educational programs at NYU, the University of Denver, and Cornell University. He also led the successful development of three analysis-driven businesses from the ground-up, one of which has become a leading provider of hospitality forecasting technologies. His 3-Day Analytical Discovery Workshop helps organizations identify and develop opportunities to strategically leverage statistics and data mining to create change and improve service. Timothy received his Bachelors of Science degree in Statistics and Biology with Distinction in Research from Cornell University and is a certified Emergency Medical Technician. His most recent article, "What are Analytics?" is currently in-press with Managed Healthcare Executive.

BASUG Membership

Keep your BASUG Membership up-to-date! Bring a completed membership form to the meeting, along with your check. The Individual Membership fee is \$30 per year. Corporate memberships were discontinued. For more information on our membership policy, or to print a membership form visit:

<http://www.basug.org/basugj.shtml>

Directions and Parking <http://www.basug.org/directions.html>

BASUG Contacts

To email our Webmaster: basugwm@basug.org

Mailing Address: BASUG
 PO Box 253
 Boston, MA 02117

BASUG Quarterly Meeting Announcement

Expand your SAS Tool Box!

Our morning meeting is full of varied tips and techniques that you'll want to add to your SAS Tool Box, to help you to improve your SAS code and generate very high-quality output. We are pleased to bring back two previous Coders Corner speakers, who will present expanded versions of their talks. And our afternoon instructor will also speak at our meeting, showing us lots of new tools.

Topic: SAS Tool Box

When: Wednesday, December 10, 2008, 8:45 AM – Noon

Where: **Holiday Inn Brookline**
1200 Beacon Street
Brookline, MA 02446
617-277-1200

See directions below

How: **Individual, On-Line Registration Required.** No e-mail please!

To register: Please visit: <http://www.basug.org/register.php3>

Contact: If you have questions about the meeting contact:
Bruno Berszoner (BBerszoner@rcn.com) or
Louis Sinoff (lsinoff@mail.com)

Agenda

- 8:15 - 8:45 Sign-in and refreshments
- 8:45 - 9:00 Announcements
- 9:00 - 9:50 **Enterprise Guide: Tips for Programmers**
John Chapman, Markcelian Analytics, Inc.
- 9:50 - 10:00 Break
- 10:00 - 10:30 **SAS Macros to Help Create Report Quality Tables in Word Documents,**
Brendan Bartley, CBAR at Harvard University
- 10:30 - 10:45 Break
- 10:45 - 12:00 **Hot Tips from SAS-L**
Mike Zdeb, University of Albany, School of Public Health

Note: Times are approximate. Please re-visit the BASUG website (www.basug.org) for updated information.

Abstracts and Bios

Enterprise Guide: Tips for Programmers

John Chapman, Ph.D., Markcelian Analytics, Inc.

SAS marketing of Enterprise Guide (EG) emphasizes its graphical nature, with tasks and wizards that offer “point-and-click interface to the full power of SAS”. But EG also has many features that make it an attractive alternative to the conventional Windows interface for programming. This presentation will describe some of these features and seek to motivate programmers who aren't using EG to give it a try.

John D. Chapman, Ph.D. provides data analysis and reporting services to the health care industry. SAS is a

valued tool in his arsenal. A long-time user of the SAS Windows interface, he became an EG convert with the release of version 4.1 and would never choose to go back.

SAS Macros to Help Create Report Quality Tables in Word Documents

Brendan Bartley, Center for Biostatistics in AIDS Research (CBAR) at Harvard University

This paper describes a few macros to enhance RTF output for reports. The first macro allows the user to move tables within a document and have the table numbering change corresponding to their location in the document. It also sets up bookmarks for reference in a report. Another macro eliminates all but the first section breaks in a RTF document. The last macro gives a table its own style name so that its formatting is retained even when text formatting in other parts of the document change.

Brendan Bartley currently works as a programmer at the Center for Biostatistics in AIDS Research (CBAR) at Harvard University and has been using SAS for 7 years. His flexible work schedule allows him to watch several European soccer matches in the afternoons, especially The Champions League, The World Cup, and The European Championships.

HOT Tips from SAS-L

Mike Zdeb, University of Albany, School of Public Health

One way to improve your SAS skills is to spend some time with SAS-L. If you are already a SAS-L participant, you know that it is a wonderful resource. If you have never heard of SAS-L, this presentation will give you a brief introduction and then use a number of recent postings to demonstrate why it is worth at least an occasional look.

Have you ever used two SET statements to read the same data set twice within one data step?

Have you have ever used the same data set name twice within a SET statement (the wonders of the self-interleave)?

Have you ever used the MOD function to help populate the elements of an array?

Ever use CALL EXECUTE or a data step hash object?

You probably have used a DO loop, but what about a DOW loop?

If you are saying to yourself that you've gotten along fine without any of the above, perhaps a look at how all of the above (and more) have been used to solve problems posed on SAS-L will convince you that SAS-L is well worth a look.

Mike Zdeb is an assistant professor in the epidemiology department of the University@Albany School of Public Health. He has been a SAS user for over 20 years and is the author of "Maps Made Easy Using SAS", a book in the SAS books-by-users series. Mike has presented papers and conducted workshops at SUGI (aka SAS Global Forum) and NESUG, and also at meetings of various local SAS user groups in the US and Canada. Among courses taught by Mike at the School of Public Health are a three credit introductory SAS course and a one credit course in advanced topics in SAS. Mike (and his students) have been the subject of several SAS e-newsletter articles and links those articles are available on Mike's school-related web site: <http://www.albany.edu/~msz03>.

BASUG Membership

Keep your BASUG Membership up-to-date! Bring a completed membership form to the meeting, along with your check. The Individual Membership fee is \$30 per year. Corporate memberships were discontinued. For more information on our membership policy, or to print a membership form visit:

<http://www.basug.org/basugj.shtml>

Directions and Parking

Directions to the Brookline Holiday Inn can be found at: <http://www.basug.org/directions.html>

BASUG Contacts To email our Webmaster: basugwm@basug.org

Mailing Address: BASUG
PO Box 253
Boston, MA 02117

BASUG Half-Day Training Session

We are delighted to host Mike Zdeb, who will be teaching a half-day course for BASUG on Combining Data. If you find yourself needing to combine datasets, check out this course – it's packed with tools, tips, techniques and loads of examples!

And by the way... The topic of combining data fits right in with our SAS Challenge... the basis for our March 2009 Quarterly meeting.

Topic	Combining Data: From SET to HASH
Course Description	This course is intended to guide the SAS programmer through the myriad ways of combining SAS data sets. It will cover a broad range of data set handling subjects, from the very basic (combining of data sets with a SET statement) through the more complex (combining observations using a data step hash object). Most of the time will be spent using the data step, though some PROC SQL alternatives to data step solutions will also be discussed (in the context of many-to-many merges and fuzzy merging). The course is intended for new (less than intermediate) level SAS users, though some more advanced users might learn a few new applications in the later stages of the course that discuss the use of arrays, formats, and hash tables. The entire course will be crammed with examples, and we will run lots of SAS jobs and look at outcomes, rather than just text slides. In addition to a handout, students will receive a copy of the example SAS code used throughout the course. Please see detailed topics below.
Instructor	Mike Zdeb, University of Albany, School of Public Health
When	Wednesday, December 10, 2008 1:30PM – 5:00PM
Where	Brookline Holiday Inn 1200 Beacon Street Brookline, MA 02446 (617) 277-1200 (see directions below)
Price	\$125 (if paid by check postmarked by Tuesday, November 25 th) \$150 (paid at-the-door) Please see details below.
Audience	Beginning to Intermediate SAS programmers looking to extend their skills in SAS data set manipulation.
Prerequisites	Familiarity with the very basic DATA step operations will help in understanding the more complex parts of this training, but is not entirely necessary.
Benefits	Attendees will come away with an extended toolkit for various situations that require combining data from multiple sources.
Contact	If you have questions about the course contact: Bridget Neville (Bridget_Neville@dfci.harvard.edu) or Stephen Freitas (sfreitas76@yahoo.com).
Registration	Pre-registration is required. Please see detailed information below.

Combining Data: From SET to HASH

Course Topics

- Concatenation (SET with more than one data set)
- Interleave (SET + BY)
- Merge (MERGE)
- Match-Merge (One-to-One, One-to-Many, Many-to-Many)
- PROC SQL for alternatives to data step matches and for fuzzy matching
- Using a self-interleave
- Using multiple SET statement and DOW loops in a single data step
- Look-up tables using arrays, formats, and a data step hash object.

Instructor Bio

Mike Zdeb is an Assistant Professor in the Epidemiology Department of the University of Albany, School of Public Health. He has been a SAS user for over 20 years and is the author of “Maps Made Easy Using SAS”, a book in the SAS books-by-users series. Mike has presented papers and conducted workshops at SUGI (aka SAS Global Forum) and NESUG, and also at meetings of various local SAS User Groups in the United States and Canada. Among courses taught by Mike at the School of Public Health are a three credit introductory SAS course and a one credit course in advanced topics in SAS. Mike (along with his students) has been the subject of several SAS e-newsletter articles and links to those articles are available on Mike’s school-related web site: <http://www.albany.edu/~msz03>.

Registration and Payment Instructions

Please read this entire section carefully!

1. Pre-register

You ***must pre-register*** for this training by going to this site:

<http://www.basug.org/register.php3>

**INDIVIDUAL, ON-LINE PRE-REGISTRATION IS REQUIRED.
NO EMAIL, PLEASE!**

All registrants will receive an immediate email confirmation of their pre-registration. Within 72 hours you will then receive another e-mail, with a customized invoice for the training, based on your membership and registration status.

Seats are limited, so register early! Seating is guaranteed only for pre-paid registrants.

2. Mail your check

Complete the invoice form and mail it along with your check, postmarked by November 25th, 2008, to our Treasurer at:

**BASUG
PO Box 253
Boston, MA 02117**

3. Members only!

Note that this class is open to BASUG members only. This means that your 2008 \$30 individual membership dues must be paid up in order to attend! If you are not currently a member, your invoice will include the \$30 fee for 2008 BASUG membership dues.

If you have any questions about your BASUG membership status, please

contact our Membership Coordinator (Robert Rosofsky) at:
Robert.Rosofsky@Verizon.net

4. Cutoff date

Please do not mail any checks after November 25th, 2008.

If you miss the mail-in deadline, bring your completed invoice form and a check made payable to BASUG to cover the at-the-door tuition rate of \$150, plus any \$30 membership fees due, as specified in your invoice form. For walk-ins, we will only be able to accommodate people on a first-come, first-served basis until seats are filled.

Full payment is due by the day of the class. There will be NO EXCEPTIONS. As a small voluntary organization, we cannot accept credit cards at this time. We also will not handle cash, so checks only please.

5. Refund Policy

To receive a full refund, notify either of the meeting coordinators (see "Contact" above) by 5PM on November 25th. **After November 25th, we will refund your payment only if we can fill your seat with attendees at the door.**

MEMBERSHIP

Keep your BASUG Membership up-to-date! Print out a membership form from our website, and bring it with you to the meeting (along with your check). Please note that we have discontinued Corporate Memberships - we now provide only Individual Membership at \$30 annually. For more information on our membership policy, or to print out a form, visit: <http://www.basug.org/basugj.shtml>

DIRECTIONS TO THE TRAINING

Directions to the Brookline Holiday Inn can be found at: <http://www.basug.org/directions.html>

BASUG CONTACTS

BASUG's Mailing Address: BASUG
 PO Box 253
 Boston, MA 02117

To email our Webmaster: basugwm@basug.org