

MWSUG 24 Program Guide
Sheraton Columbus Hotel, Columbus, OH



Operations Chair: George Hurley
Academic Chair: Matt Karafa

Thank you to the following:

The Trainers:

Art Carpenter	Ben Cochran
Russ Lavery	Kirk Lafler
Taylor Lewis	Arthur Li
Charlie Shipp	

SAS Institute, Inc.:

Multiple Papers Presented by SAS Staff
SAS Super Demos

Conference Committee:

The Operations Team:

Publications Coordinator – Mindy Kiss
Asst. Pubs Coordinator - Mary Wendt
Training Coordinator - Adrian Katschke
Asst. Training Coordinator - Richann Watson
Corporate Sponsorship Coordinator - Nate Derby
Volunteer Coordinator - David Corliss
Registration Coordinator - David Bruckner and Craig Wildeman

Webmaster - Josh Horstman
Asst. Webmaster - Joe Novotny
First Timer Coordinator - Andy Kuligowski
Paper Mentorship Coordinator - Michael Wilson
Constant Contact Coordinator - Deb Early

The Program Team:

Advanced Analytics - Jeff Ahnrsen and Bruce Lund
Banking & Finance – Dave Foster and Luke Castellanos
Beyond the Basics - Swati Agarwal and Brian Varney
BI Applications & Architecture - Mary MacDougal and Paul Genovesi
Black Belt SAS - Swati Agarwal and Brian Varney
Data Visualization and Graphics - Chuck Kincaid and Sue Douglass

Customer Intelligence - Steve Raimi and Laine Suppes
Hands-On Workshops - Arthur Li
In-Conference Training - Steve Popernack
JMP - Charlie Shipp and Dachao Liu

Conference Board

George Hurley, MWSUG President
D.J. Penix, MWSUG Vice President
Rex Pruitt, MWSUG Secretary

Cindy Lee, MWSUG Senior Vice President
Ken Schmidt, MWSUG Treasurer
Nancy Moser, SAS®

Welcome

Dear Fellow SAS® and JMP® users:

Thank you for making the decision to attend the 2013 MidWest SAS® Users Group (MWSUG) Conference this year.

Our 24th MWSUG Conference consists of two full days of presentations, workshops, tutorials, and other education to enhance your SAS skills. In addition, staff from SAS Institute will be available to provide their unique expertise and insight. You will also have the opportunity to meet some of the industry's best and most popular SAS educators, who have been active in providing pre- and post-conference training options as well as in-conference workshops and presentations.

The MWSUG conference is the premier SAS educational forum in the twelve-state region. MWSUG is a non-profit, all-volunteer user group and is officially recognized by SAS Institute.

We are working hard to make this conference a great opportunity for you to network, and learn more ways to make better use of SAS while getting more value from your SAS software investment. We are confident that your expectations will be exceeded in all areas with this conference experience!

Your 2013 MWSUG Conference Co-Chairs,

George Hurley
Operations Chair
Plymouth Rock Assurance
OperationsChair@mwsug.org

Matt Karafa
Program Chair
Cleveland Clinic
ProgramChair@mwsug.org

Contents

SUNDAY MEETINGS AND PRESENTATIONS	2
MONDAY MORNING PRESENTATIONS	3
MONDAY AFTERNOON PRESENTATIONS	15
MONDAY NIGHT EVENTS	22
TUESDAY MORNING PRESENTATIONS	23
TUESDAY AFTERNOON PRESENTATIONS	33
POSTERS	43
INNOVATION AND NETWORKING AREA	46

2013 Conference Schedule at a Glance

Sunday, September 22

2:30 p.m. – 6:30 p.m.	Registration and Information Desk Open – Outside Governor's Ballroom
4:00 p.m. – 4:30 p.m.	Speakers and Volunteers Meeting – House A
4:30 p.m. – 5:00 p.m.	First Timers Meeting – House A
6:30 p.m. – 8:30 p.m.	Opening Session and Keynote Address – Governor's Ballroom
8:30 p.m. – 11:00 p.m.	Opening Reception - Governor's Ballroom Foyer

Monday, September 23

7:00 a.m. – 8:30 a.m.	Continental Breakfast – Legislative Foyer
8:00 a.m. – 11:30 a.m.	Registration and Information Desk Open
8:00 a.m. – 11:30 a.m.	Speaker Presentations and Hands-on Workshops
8:00 a.m. – 11:30 a.m.	Innovation and Networking Area- Congressional Room
11:30a.m. – 12:30 p.m.	Lunch – Governor's Ballroom
12:30 p.m. – 5:00 p.m.	Registration and Information Desk Open
12:30 p.m. – 3:00 p.m.	Speaker Presentations and Hands-on Workshops
12:30 p.m. – 5:00 p.m.	Innovations and Networking Area- Congressional Room
2:30 p.m. – 5:00p.m.	SAS Certification Testing – Judicial Room
2:30 p.m. – 5:00p.m	Experts Panel Discussion – Governor's Ballroom
2:30 p.m. – 3:30 p.m.	Afternoon Refreshments – Innovation and Networking Area
5:30 p.m. – 11:00 p.m.	Monday Night Networking Event – Franklin Park Conservatory

Tuesday, September 24

7:00 a.m. – 8:30 a.m.	Continental Breakfast – Legislative Foyer
8:00 a.m. – 11:30 a.m.	Registration and Information Desk Open
8:00 a.m. – 11:30 a.m.	Speaker Presentations and Hands-on Workshops
8:00 a.m. – 11:30 a.m.	Innovations and Networking Area- Congressional Room
11:30a.m. – 1:00 p.m.	Lunch – on your own
1:00 p.m. – 4:00 p.m.	Speaker Presentations and Hands-on Workshops
4:00 p.m. – 4:45 p.m.	Closing Ceremony – Governor's Ballroom

Sunday Meetings and Presentations

First-Timers Meeting

Room House A—4:30-5:00 p.m.

For first-time attendees, this session is designed to help you get the most out of your conference experience. We'll highlight the opportunities that are available to you, give you some tips for planning your conference days, and answer questions you may have.

Volunteers and Speakers Meeting

Room House A—4:00-4:30 p.m.

This session is open to all conference speakers and volunteers. We'll review the volunteer schedule, discuss the audio-visual setup in the meeting rooms, review volunteer duties and presentation procedures, and answer any questions you may have. You'll also get to meet the chair of the section in which you will be speaking or volunteering, and speakers and session coordinators will get to meet each other.

Sunday Dinner Keynote Presentation:

Thornton May, IT Leadership Academy

Thornton May is a futurist, educator and author. His extensive experience researching and consulting on the role and behaviors of “C” level executives in creating value with information technology has won him an unquestioned place on the short list of serious thinkers on this topic. Thornton combines a scholar’s patience for empirical research, a stand-up comic’s capacity for pattern recognition and a second-to-none gift for storytelling to address the information technology management problems facing executives. The editors at *eWeek* honored Thornton, including him on their list of Top 100 Most Influential People in IT. The editors at *Fast Company* labeled him ‘one of the top 50 brains in business.’

Please join us after dinner in the Governor’s Ballroom Foyer for a mixer sponsored by SAS.

Monday Morning Presentations

Advanced Analytics

Paper AA-02

Monday:8:30 AM - 9:20 AM

Room: Executive ABC

PROC SURVEYSELECT as a Tool for Drawing Random Samples

Taylor Lewis

This paper illustrates many of the sampling algorithms built into PROC SURVEYSELECT, particularly those pertinent to complex surveys, such as systematic, probability proportional to size (PPS), stratified, and cluster sampling. The primary objectives of the paper are to provide background on why these techniques are used in practice and to demonstrate their application via syntax examples. Hence, this is not a how-to paper on designing a statistically efficient sample there are entire textbooks devoted to that subject. One exception is that the paper will discuss a few recently incorporated sample allocation strategies specifically, proportional, Neyman, and optimal allocation. The paper concludes with a few examples demonstrating how one can use PROC SURVEYSELECT to handle certain frequently-encountered sample design issues such as alternative sampling methods across strata and multi-stage cluster sampling.

Paper AA-03

Monday:9:30 AM - 10:20 AM

Room: Executive ABC

Finding the Gold in Your Data: An Overview of Data Mining

David Dickey

The term "data mining" has appeared often recently in analytic literature and even in popular literature, so what exactly is data mining and what does SAS* provide in terms of data mining capabilities? The answer is that data mining is a collection of tools designed to discover useful structure in large data sets. With an emphasis on examples, this talk gives an overview of methods available in SAS* Enterprise Miner and should be accessible to a general audience. Topics include predictive modeling, decision trees, association analysis, incorporation of profits and neural networks. We'll see that some of the basic ideas underlying these techniques are closely related to standard statistical techniques that have been around for some time but now have new more appealing names than their statistical ancestors and have been automated to become more user friendly. The talk is for a general audience and will use SAS Enterprise Miner, and SAS/STAT procedures. * SAS is the registered trademark of SAS Institute, Cary, NC.

Paper AA-04

Monday:10:30 AM - 10:50 AM

Room: Executive ABC

Logistics Performance Metrics using SAS® Macros

Michael Frick

Performance metrics for time-related processes, such as logistics delivery times, often follow an all too familiar pattern that is highly skewed to the right. In such cases, traditional central tendency statistics are of little use as they are dominated by the out-of-process events in the tail of the distribution and do not adequately describe the distribution of delivery times for in-process events. In this work, I propose a methodology that first splits the distribution into its two main components using a SAS® macro that employs piece-wise linear regression to automatically determine the

break point between the main body of the distribution and the tail. Using this automated splitting mechanism, we are able to classify the performance of a large number of individual shipping lanes across three categories: (1) percent of out-of-process events, (2) average performance for in-process events against an expected standard, and (3) variation in performance for in-process events. Note: Included code uses base SAS, PROC REG, and SAS Macros. An appendix is included as a macro tutorial for the macro language constructs used in the text.

Paper AA-16

Monday, 11:00 AM - 11:20 AM

Room: Executive ABC

Leading & Lagging Indicators in SAS®

David Corliss

Leading indicators are familiar to us from economics as factors whose changes are correlated with future events, while lagging indicators are correlated with previous activity. However, leading and lagging indicators appear in many areas on interest: changes in a person's diet, such as gaining or losing weight, correspond to risk group changes at a later date. An increase in poverty over time corresponds to an increasing prison population later on. Lagging indicators may be used to infer the past, unobserved history of a physical systems from automobile repairs to stellar explosions. This paper demonstrates code in Base SAS for identifying leading and lagging indicators and measuring the difference in time between two linked behaviors. A facility is included for addressing the presence of missing data by suppressing points in time where the data are insufficient to support accurate results. Examples are given in biostatistics, social sciences and astrophysics as well as econometrics to demonstrate how leading and lagging indicators may be used to better understand correlated past and future behaviors.

Beyond the Basics

Paper BB-01

Monday:8:00 AM - 8:20 AM

Room: Legislative B

Add a Little Magic to Your Joins

Kirk Paul Lafler

To achieve the best possible performance when joining two or more tables in the SQL procedure, a few considerations should be kept in mind. This presentation explores options that can be used to influence the type of join algorithm selected (i.e., step-loop, sort-merge, index, and hash) by the optimizer. Attendees learn how to add a little magic with MAGIC=101, MAGIC=102, MAGIC=103, IDXWHERE=Yes, and BUFFERSIZE= options to influence the SQL optimizer to achieve the best possible performance when joining tables.

Paper BB-02

Monday:8:30 AM - 8:50 AM

Room: Legislative B

Same Data Different Attributes: Cloning Issues with Data Sets

Brian Varney

When dealing with data from multiple or unstructured data sources, there can be data set variable attribute conflicts. These conflicts can be cumbersome to deal with when developing code. This paper intends to discuss issues and table driven strategies for dealing with data sets with inconsistent variable attributes.

Paper BB-03

Monday:9:00 AM - 9:20 AM

Room: Legislative B

The Secret Life of Data Step

Swati Agarwal

Each SAS DATA step functions as a self-contained mini program that is compiled and run within your overall SAS program. Much of DATA step processing is implicit. For example, DATA step statements are executed within an implied loop even though you do not loop control statements. This paper discusses several of the implied DATA step statements that control how your DATA step really works. Concepts you will be introduced to: " the Logical Program Data Vector " automatic SAS variables and how are they used " understanding the internals of DATA step processing " what happens at program compile time " what's actually happening at execution time " how variable attributes are captured and stored. There is something in this paper for all levels of programmers from the very beginner to the most advanced

Paper BB-04

Monday:9:30 AM - 10:20 AM

Room: Legislative B

How Do I . . . ? There is more than one way to solve that problem; Why continuing to learn is so important

Art Carpenter

In the SAS® forums questions are often posted that start with How do I . . . ? . Generally there are multiple solutions to the posted problem, and often these solutions vary from simple to complex. In many of the responses the simple solution is often inefficient and also reflects a somewhat naïve understanding of the SAS language. This would not be so very bad except sometimes the responder thinks that their response is the best solution, or perhaps worse, the only solution. Worse yet, when there is a range of solutions, the right answer that the original poster selects often reflects the simplest solution that the original poster understands. In both cases these folks have stopped learning and have stopped expanding their understanding of the language. The examples in this presentation will illustrate the progression of solutions from the simple (simplistic) to the sophisticated for a number of How do I . . . ? questions, and through the discussion of the individual techniques, we will learn how and why it is so very important to continue to learn.

Paper BB-05

Monday:10:30 AM - 11:20 AM

Room: Legislative B

A First Look at the ODS Destination for PowerPoint

Tim Hunter

This paper introduces the ODS destination for PowerPoint, part of the next generation of ODS destinations. Using our new PowerPoint destination you can send proc output directly into native PowerPoint format. See examples of slides created by ODS. Learn how to create presentations using ODS, how to use ODS style templates to customize the look of your presentations and how to use pre-defined layouts to make title slides and two-column slides. Find out how the PowerPoint destination is like other ODS destinations and how it's different. Stop cutting and pasting and let the ODS destination for PowerPoint do the work for you!

Customers Intelligence

Paper CI-01

Monday:8:00 AM - 8:50 AM

Room: House A

Few SAS PROCs that Help Improving College Recruitment and Enrollment Strategies

Harjanto Djunaidi and Monica Djunaidi

Maintaining or increasing the rate of college student enrollment growth practically is the biggest dream that most US college administrators have. This task never been so important today compare to what it was in the past few years. There are several reasons why keeping student enrollment growth rate positively. First, colleges need to balance their budget. With ever increasing operational cost, reduction on contributions, and less availability of tax payers money, therefore, the only open option that colleges have is to raise the tuition. Second, often enrollment growth is used as an important by the school Board to judge the administrators success. Third, increasing competition from for-profit universities has made the market share of traditional institution to shrink. Fourth, freshmen students preference may have changed due to increasing college tuition. This leads of taking the so-called general education courses to much cheaper institution such as Community College or online classes. Due to high unemployment rate, there is clear evidence that demand for education services may have declined even further and course substitutions are in their highest level which never been seen before. Because of these dynamic changes in the industry, among others, the US colleges have struggled in the last a couple of years to get their enrollment numbers grow. However, most of them face tough time, especially to those colleges which are managed under the BAU (business as usual) mindset. New education analytics such as Institutional Research Intelligence (IRI) approaches is most needed which will help colleges to deal with enrollment problem. The IRI concepts are based on new paradigms to approach the problem from different and more efficient angles. IRI as one of the new concepts introduced within the education analytics discipline suggests that higher education institutions in the US are able to improve the retention or persistent rate before the students transfer or quit school. This paper demonstrates clearly how few SAS PROCs combined with market intelligence can get the job done nicely, efficiently and painless. This procedure helps identifying the group of students which can be recruited. Therefore it provides recruitment strategies to the college decision makers such as Office of Enrollment Management the recruitment and direction or student recruitment. Applying the IRI mindsets are even more urgent if the US Senate also approves the H.R. 1911 Bill which makes the student loans interest rate to be variable. The implication of passing the Bill is obvious in that it will make financial planning more difficult for the students and their family which may push demand for education further down.

Paper CI-02

Monday:9:00 AM - 9:20 AM

Room: House A

Introducing PROC PSYCHIC

David Corliss

Here at MWSUG 2063, we are pleased to present some of the functionality of the soon-to-be released procedure PSYCHIC. While release of this long-awaited procedure has been somewhat delayed, much of the functionality of PROC PSYCHIC is available in the most recent release of SAS; this introductory-level paper will focus on currently available capabilities, processes and techniques. Features of this data integration master procedure include interrogation of data to determine errors and the design of data models best suited to the data itself despite the best intentions of helpful managers. While the brain wave text mining feature remains in beta in the upcoming release, recommendations are made to replicate this functionality using voice-based communication to better identify customer needs and wants in the usual absence of clear direction.

Paper CI-03

Monday:9:30 AM - 10:20 AM

Room: House A

Marrying Customer Intelligence with Customer Segmentations to Drive Improvements to Your CRM Strategy

Phil Krauskopf

There are a plethora of segmentation schema available and that can be used to group your customer base. However, choosing which one is best for your customer base, is often a point of differing opinion within a marketing department. And once a particular schema is chosen, selecting which intelligence-gathering activities should be applied in order to learn about the customers within each segment presents its own set of issues. Finally, integrating that intelligence into your current marketing strategy to create a structured and comprehensive, even stronger strategy muddies things even further. This presentation will review some of the most common segmentation schema which the author has encountered in Industry, and some ideas on how to select among various options. He will also discuss some of the intelligence-gathering tools available to the data scientist, and how this intelligence might be integrated into a marketing strategy.

Paper CI-07

Monday:10:30 AM - 10:50 AM

Room: House A

SAS PROC REPORT: How It Helps Improving College Student Retention Rate

Harjanto Djunaide and Monica Djunaide

Abstract Maintaining college freshmen students never been so important today compare to what it was in the past few years. There are several reasons why keeping retention rate high for first-year college student will positively affect the schools report card as well as their financial resources. First, increasing college competition to enroll the best students cannot be taking lightly. The probability for stronger and well performed freshmen students to transfer school is much higher compared to that of sophomore, junior or senior-year students. Second, maintaining those students who are in the college is much easier than to recruit the new ones. Third, high retention rate will ensure a better financial resources to the institution as well as improve their institution perceived value. Because of these reasons, among others, the US colleges have struggled over a number of years to get their retention rate increased. However, most of their efforts have been focused on after the fact. Meaning these colleges identify the problems after the students are no longer around or left the school. This is a typical situation facing by colleges who are still operating under the BAU (business as usual) mindset. The new paradigms discussed in Institutional Research Intelligence (IRI) approaches the problem from different and more efficient angles. IRI as one of the new concepts introduced within the education analytics discipline suggests that higher education institutions in the US are able to improve the retention or persistent rate before the students transfer or quit school. This paper demonstrates clearly how SAS PROC REPORT can get the job done nicely, efficiently and painless. This procedure help flagging the group of risky students ahead of time or before they even attend their first class. Therefore it gives ample opportunities to the college decision makers such as Student Success Office or the Provost Office the ability to convert the unpleasant outcome to otherwise. This new approach is much needed by US colleges under the present unfavorable competitive environments which have drastically changed. The example the 2011 Satisfactory Academic Progress (SAP) regulation and potentially the 2013 H.R. 1911 Bill might have direct effects on the entire operation of every single US higher learning institution in the future.

Banking and Financial Services

Paper FS-01

Monday:10:00 AM - 10:20 AM

Room: Senate A

PROC NL MIXED and PROC IML Mixture distribution application in Operational Risk

Sabri Uner and Yunyun Pei

The main purpose of this study is to provide guidelines on how to consider mixture distributions for operational risk severity distribution modeling, with an emphasis on truncated loss data. Mixture model probability distribution function for truncated operational loss data is introduced and we presented our findings for empirical tests to estimate distribution parameters. However, this study does not intend to advocate or to propose adopting mixture forms without exploring other alternatives, but rather highlights the flexibility of the mixture models and present examples where it can serve better for some specific cases.

Paper FS-03

Monday:11:00 AM - 11:20 AM

Room: Senate A

Use SAS/ETS to Forecast Credit Loss for CCAR

Xinpeng "Tom" Wang

After the 2007-2009 financial crises, the Federal Reserve begins to conduct annual stress tests of Bank Holding Companies (BHCs) with total consolidated assets of \$50 billion or more (Covered Company). To estimate credit losses for BHCs loan portfolio, one solution is to employ SAS logistic and regression procedures to predict probability of default (PD) and loss given default (LGD). This paper presents an alternative utilizing SAS/ETS package to forecast net charge offs (NCOs) on the aggregated data. Then the paper discusses general limitations of models in financial risk management and specific shortages for both approaches.

Hands-On Workshops

Paper HW-01

Monday:8:00 AM - 9:20 AM

Room: Judicial

Know Thy Data: Techniques for Data Exploration

Andrew Kuligowski and Charu Shankar

Get to know the #1 rule for data specialists: Know thy data. Is it clean? What are the keys? Is it indexed? What about missing data, outliers, and so on? Failure to understand these aspects of your data will result in a flawed report, forecast, or model. In this hands-on workshop, You will learn multiple ways of looking at data and its characteristics. You will learn to leverage PROC MEANS and PROC FREQ to explore your data. You will learn how to use PROC CONTENTS and PROC DATASETS to explore attributes and determine whether indexing is a good idea. You will learn to employ powerful PROC SQL s dictionary tables to easily explore aspects of your data.

Paper HW-02

Monday:9:30 AM - 10:50 AM

Room: Judicial

Fast Access Tricks for Large Sorted SAS Files

Russell Lavery

This HOW will provide an explanation of, and sample code for, several ways that you can, very quickly, find rows in sorted SAS files. The techniques are not generally known but can be useful when you need to retrieve specified rows

from a large sorted file. This cartoon format of this presentation makes it a good review of the internals of the SAS data step and the program data vector. Even if you do not need these techniques, if you have never read these sections in the manuals this might be just the thing for you.

JMP

Paper JM-01

Monday:8:00 AM - 8:50 AM

Room: Senate B

Dealing with Data below the Detection Limit: Limit Estimation and Data Modeling

Diane Michelson and Mark Bailey

Measuring trace levels of contaminants in chemicals or gases can be difficult. When the signal is very small, it can be lost in the noise. Chemical analyses are characterized by their accuracy, precision, and linear range. A detection limit is the smallest amount that can be reliably detected by the procedure. The procedure can be used on a blank, with no amount of the substance, or on a sample containing the substance to be measured. Various methods of estimating a detection limit are compared. We will examine the differences between using tolerance intervals or confidence intervals on measurements on blanks, as well as regression approaches, including the use of linear regression by ordinary and weighted least squares. A contamination example will be demonstrated, using the Distribution platform, Fit Y by X, and Fit Model in JMP. Responses below the detection limit can be included in your data analysis. Ad hoc approaches produce biased estimates and should be avoided. Such responses are censored data, and likelihood methods exist to handle censoring. An example of a designed experiment is handled using the Parametric Survival personality in Fit Model in JMP.

Paper JM-03

Monday:9:00 AM - 9:20 AM

Room: Senate B

The JMP Journal: An Analyst's Best Friend

Nate Derby

The JMP Journal is an incredibly useful tool for consultants and analysts, yet it's not commonly used. We first explain what the JMP Journal is, then describe how it can be effectively used to keep track of an analysis (and its underlying code), to present results to a boss or client, or to use as a collaboration tool.

Paper JM-02

Monday:9:30 AM - 9:50 AM

Room: Senate B

2x10-Minute JMP®

George Hurley

Heard of JMP®, but haven't had time to try it? Don't want to devote 50 minutes to a talk about software that you might not want to use? This is the talk to you. In the first 10 minutes, you will learn some of the amazing visualization and modeling features in JMP and how to use them. In the second 10 minutes, we'll go through a live example! This talk will JMP-start your JMP usage. When it's complete, we suspect you will want to attend some of the longer talks, too.

Paper JM-07 Monday:10:00 AM - 10:20 AM Room: Senate B
JMP, and Teaching JMP: a panel discussion
Shipp, Derby, Gundlach, Bailey, and Di

A panel of the invited speakers will [each] present a viewpoint on JMP software, JMP usage, and JMP user and management training. We will discuss the importance of management visibility, user group possibilities, and JMP training within your enterprise and in the community. You, the audience, will be an important part of the lively discussion that follows.

Paper JM-04 Monday:10:30 AM - 11:20 AM Room: Senate B
Google® Search Tips and Techniques for SAS® and JMP® Users
Charles Edwin Shipp and Kirk Paul Lafler

Google® (www.google.com) is the world's most popular and widely-used search engine. As the premier search tool on the Internet today, SAS® and JMP® users frequently need to identify and locate SAS and JMP content wherever and in whatever form it resides. This paper provides insights into how Google works and illustrates numerous search tips and techniques for finding articles of interest, reference works, information tools, directories, PDFs, images, current news stories, user groups, and more to get search results quickly and easily.

Pharmaceutical Applications

Paper RX-01 Monday:8:00 AM - 8:50 AM Room: House B
Introduction to REDCap for Clinical Data Collection
Shannon Morrison

REDCap (Research Electronic Data Capture) is a free web-based tool that attempts to make collecting clinical data easy. The application is used primarily for creating databases and/or surveys by using the "Online Designer" in a web browser or constructing a "Data Dictionary" in Excel, which can be uploaded to REDCap. Multiple users can be added to a project and changes can be made quickly and easily online. Users have the ability to add validations and/or branching logic to variables, and once data collection is complete data can be automatically exported to common statistical packages...including SAS! This talk will focus on using the Online Designer to create databases in REDCap.

Paper RX-02 Monday:9:00 AM - 9:20 AM Room: House B
SAS and REDCap API: Efficient and Reproducible Data Import and Export
Sarah Worley and Dongsheng Yang

REDCap (Research Electronic Data Capture), a web application for the development and use of online research databases, allows SAS users to download automatically-generated SAS code for importing, labeling, and formatting data. REDCap Application Programming Interface (REDCap API) provides SAS users with the option to import and export data, files, and database settings through SAS programs. The use of REDCap API with both PC and server installations of SAS 9.3 will be demonstrated.

Paper RX-03

Monday:9:30 AM - 9:50 AM

Room: House B

Let SAS® Do Your DIRty Work

Richann Watson

Making sure you have all the necessary information to replicate a deliverable saved can be a cumbersome task. You want to make sure that all the raw data sets are saved, all the derived data sets, whether they are SDTM or ADaM data sets, are saved and you prefer that the date/time stamps are preserved. Not only do you need the data sets, you also need to keep a copy of all programs that were used to produce the deliverable as well as the corresponding logs when the programs were executed. Any other information that was needed to produce the necessary outputs needs to be saved. All of these needs to be done for each deliverable and it can be easy to overlook a step or some key information. Most people do this process manually and it can be a time-consuming process, so why not let SAS do the work for you?

Paper RX-04

Monday:10:00 AM – 10:20 AM

Room: House B

An Analysis of Risk Behavior Trends and Mental Health in American Youth Using PROC SURVEYLOGISTIC

Deanna Schreiber-Gregory

The current study looks at recent mental health and risk behavior trends of youth in America. Data used in this analysis was provided by the Center for Disease Control and Prevention and gathered using the Youth Risk Behavior Surveillance System (YRBSS). Interactions between risk behaviors and reported mental states - such as depression, suicidal ideation, and disordered eating - are the main subject of this analysis. This study also outlines demographic differences in risk behaviors and mental health issues as well as correlations between the different mental health issues. A final regression model including the most significant contributing factors to suicidal ideation, depression, and disordered eating is provided and discussed. Results included reporting differences between the years of 1991 and 2011. All results are discussed in relation to current youth health trend issues. Data was analyzed using SAS® 9.3 and JMP® 10. This presentation is meant for any level of SAS User.

Paper RX-06

Monday:10:30 AM - 10:50 AM

Room: House B

Creation and Implementation of an Analytic Dataset for a Multisite Surveillance Study using Electronic Health Records (EHR) and Medical Claims Data

Renuka Adibhatla, Gabriela VazquezBenitez, Mary Becker and Amy Butani

SAS code is utilized for research projects of the Health Maintenance Organization Research Network (HMORN). The HMORN is a consortium of 18 health care delivery organizations with integrated research divisions with over 15 million patients. Multisite research is conducted by utilizing the HMORN Virtual Data Warehouse (HMORN VDW), a distributed data network where each site locally stores their data in standardized data structures, in this case as SAS datasets. HMORN VDW is an excellent source of data for surveillance and observational studies of chronic conditions such as diabetes and cardiovascular disease. SAS programs are shared among HMORN project sites in order to extract site-specific data from each site which is then combined to represent the overall HMORN patient population and results. This paper focuses on the various programming techniques, including some borrowed from previous SAS conference papers and lessons learned from an ongoing HMORN multi-site project involving 11 HMORN research centers. These methods were applied to produce a clean analytic dataset with extensive use of PROC SQL, arrays, and macros. The

paper clearly demonstrates the advantages of utilizing distributed SAS code for extracting data, the benefits from a HMORN VDW for multisite research, and the required organization of the SAS datasets and programs. The paper also discusses the challenges presented with data volumes and storage. SAS 9.2 was used on both Windows and UNIX environments. The paper is intended for beginner to intermediate level SAS programmers preferably with minimal knowledge of EHR and Claims data.

Paper RX-07

Monday:11:00 AM - 11:20 AM

Room: House B

Examining Risk Factors of Referred and Substantiated Child Maltreatment in California Latino Infants Using PROC GENMOD

Kechen Zhao

Count data are frequently encountered in social science, epidemiological research, and biomedical field. The GENMOD procedure is commonly used to build a Poisson regression model for count data. The use of PROC GENMOD is demonstrated in our study of California Latino children maltreatment. In this study, number of referred or substantiated child maltreatment incidences over a particular period is modeled by using Poisson regression. This study aimed to examine to what extent to which variations in referred and substantiated child maltreatment are attributable to the variations in child and maternal demographic and characteristics of California Latino population. Model building process by using PROC GENMOD alone with various features which were newly implemented in SAS® 9.2 will also be introduced in this paper.

Support You Can Count On

The SAS Customer Loyalty team is your liaison to a wealth of support services and resources.

How can we assist you?

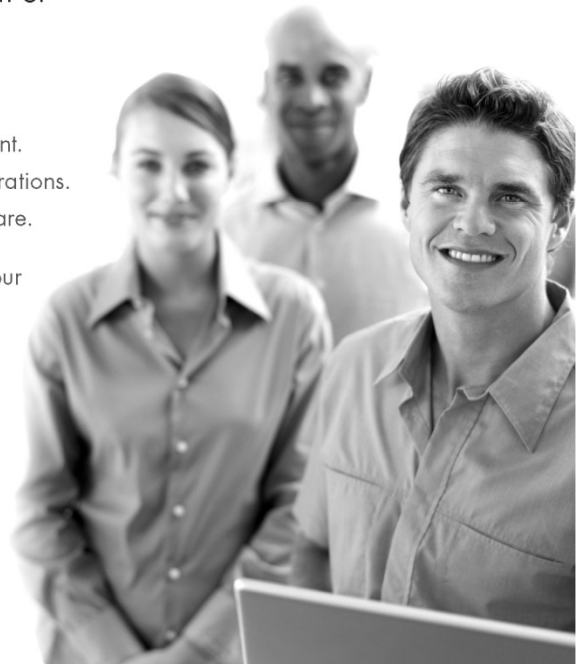
- Determine which resources can increase the value of your investment.
- Keep you informed about new features, software releases and migrations.
- Make recommendations for getting the most value from your software.

These are just a few examples of what we can do for you. Let us be your guide to all that SAS has to offer.

Learn more at support.sas.com/promise and contact us at support.sas.com/customerloyalty.



support.sas.com/promise



Paper S1-01

Monday:8:00 AM - 8:50 AM

Room: Legislative A

The Essence of DATA Step Programming

Arthur Li

The fundamental of SAS® programming is DATA step programming. The essence of DATA step programming is to understand how SAS processes the data during the compilation and execution phases. In this paper, you will be exposed to what happens behind the scenes while creating a SAS dataset. You will learn how a new dataset is created, one observation at a time, from either a raw text file or an existing SAS dataset, to the program data vector (PDV) and from the PDV to the newly-created SAS dataset. Once you fully understand DATA step processing, learning the SUM and RETAIN statements will become easier to grasp. Relating to this topic, this paper will also cover BY-group processing.

Paper S1-02

Monday:9:00 AM - 9:20 AM

Room: Legislative A

The Power of Macros

Audrey Yeo

The SAS® Macro facility is an extremely powerful tool that should be in the toolbox of every SAS programmer. However, without some proper training it is difficult to implement, or when it is implemented it often results in hard to understand code. On the other hand once you have mastered the macro facility, it opens up a whole new world. This paper will show some examples of creating using SAS macros to help simplify coding and reduce repetition or duplication of code.

Paper S1-05

Monday:9:30 AM - 9:50 AM

Room: Legislative A

Strategies and Techniques for Debugging SAS® Program Errors and Warnings

Kirk Paul Lafler

As a SAS® user, you've probably experienced first-hand more than your share of program code bugs, and realize that debugging SAS program errors and warnings can, at times, be a daunting task. This presentation explores the world of SAS errors and warnings, provides important information about syntax errors, input and output data sources, system-related default specifications, and logic scenarios specified in program code. Attendees learn how to apply effective techniques to better understand, identify, and fix errors and warnings, enabling program code to work as intended.

Paper S1-04

Monday:10:00 AM - 10:20 AM

Room: Legislative A

Using SAS ® to Analyze Data Submitted to the National Healthcare Safety Network (NHSN)

Michelle Hopkins

Hospitals across the nation are required by the Centers for Medicare & Medicaid Services (CMS) to submit data on a variety of Healthcare-associated infections (HAIs). These are infections that are acquired in the course of receiving care for other health conditions. These HAIs include Central Line-associated Bloodstream Infection, Catheter-associated Urinary Tract Infections, Surgical Site Infections, Clostridium difficile Infection, etc. The number of HAIs in the United States is alarming, costly, and can be deadly. Hospitals report this data to the Center for Disease Control's National Healthcare Safety Network (NHSN). Using SAS ® 9.2, this paper will show how you can use the data found on NHSN for analysis. This analysis includes importing data, summarizing infection rates, and using the data to drive much needed improvement.

Paper S1-09

Monday:10:30 AM - 11:20 AM

Room: Legislative A

Introducing a Colorful Proc Tabulate

Ben Cochran

Several years ago, one of my clients was in the business of selling reports to hospitals. He used PROC TABULATE to generate part of these reports. He loved the way this procedure 'crunched the numbers', but not the way the final reports looked. He said he would go broke if he had to sell naked PROC TABULATE output. So, he wrote his own routine to take TABULATE output and render it through Crystal Reports. That was before SAS came out with the Output Delivery System (ODS). Once he got his hands on SAS ODS, he kissed Crystal Reports good-bye. This paper is all about using PROC TABULATE to generate fantastic reports. If you want to generate BIG money reports with PROC TABULATE, this presentation is for you.

Monday Afternoon Presentations

Advanced Analytics

Paper AA-11

Monday:12:30 PM - 1:20 PM

Room: Executive ABC

Ideas and Examples in Generalized Linear Mixed Models

David Dickey

SAS® PROC GLIMMIX fits generalized linear mixed models for nonnormal data with random effects, thus combining features of both PROC GENMOD and PROC MIXED. I will review the ideas behind PROC GLIMMIX and offer examples of Poisson and binary data. PROC NLMIXED also has the capacity to fit these kinds of models. After a brief introduction to that procedure, I will show an example of a zero-inflated Poisson model, which is a model that is Poisson for counts 1,2,3,&, but has more 0s than is consistent with the Poisson. This paper was first delivered at the 2010 SAS Global Forum and again that year at MWSUG. The paper is intended for a somewhat statistically sophisticated audience. In particular, familiarity with the idea of random effects is helpful, though a less technical user might find the examples of interest anyway. ® SAS is the registered trademark of SAS Institute, Cary, NC

Paper AA-12

Monday:1:30 PM - 2:20 PM

Room: Executive ABC

Uncovering Patterns in Textual Data for SAS Visual Analytics and SAS Text Analytics

Meera Venkataramani

SAS Visual Analytics is a powerful tool for exploring big data to uncover patterns and opportunities hidden with your data. The challenge with big data is that the majority is unstructured data, in the form of customer feedback, survey responses, social media conversation, blogs and news articles. By integrating SAS Visual Analytics with SAS Text Analytics, customers can uncover patterns in big data, while enriching and visualizing your data with customer sentiment, categorical flags, and uncovering root causes that primarily exist within unstructured data. This paper highlights a case study that provides greater insight into big data, demonstrates advanced visualization, while enhancing time to value by leveraging SAS Visual Analytics high-performance, in-memory technology, Hadoop, and SAS advanced Text Analytics capabilities.

Banking and Financial Services

Paper FS-04

Monday:12:30 PM - 12:50 PM

Room: Senate A

SAS Automation - From Password Protected Excel Raw Data to Professional-Looking PowerPoint Report

Jeff Hao

Microsoft® Excel is an excellent tool for formatting and presenting data. For security sake, especially in banking and finance sectors, Excel files are typically password protected to restrict access to the information. SAS has become a universal business solution tool for clinical trial reporting, data management, statistical analysis, econometrics, data mining, business planning, forecasting, and decision supports, etc., especially its automation capability. The first step of SAS automation process is to directly read password protected Excel raw data into SAS datasets. The final step deals with presenting the tables and figures into customized profession-looking PowerPoint report. Actually, the two steps are

the most difficult parts of the whole automation process. This paper presents a detailed step by step approach to combine SAS with DDE, Proc Template, ODS Printer, and VBScript to realize the seamless automation process from password protected Excel raw data to PowerPoint by single SAS run click.

Paper FS-05

Monday:1:00 PM - 1:50 PM

Room: Senate A

Modeling Proportions in SAS

Wensui Liu

In practice, OLS (Ordinary Least Square) regression has been widely used to model rates and proportions bounded between 0 and 1 due to its simplicity. However, the conditional distribution of an OLS regression model is assumed Gaussian $N(X'B, \sigma^2)$, which is questionable for a variate in the open interval (0, 1). In this paper, we surveyed six alternatives modeling methods for such outcomes, including OLS regression with the LOGIT transformation, NLS (Nonlinear Least Square) regression, Tobit model, Beta model, Simplex model, and Fractional LOGIT model, and demonstrated their implementations in SAS through a data analysis exercise. The purpose of my study is to provide a comprehensive survey in SAS user community on how to model percentage and proportion outcomes in SAS.

KEYWORDS Rate and Proportion outcomes, OLS, NLS, Tobit, Beta, Simplex, Fractional LOGIT, PROC NL MIXED

Beyond the Basics

Paper BB-06

Monday:12:30 PM - 12:50 PM

Room: Legislative B

The Joinless Join; Expand the Power of SAS® Enterprise Guide® in a New Way

Kent Phelps, Ronda Phelps and Kirk Lafler

SAS® Enterprise Guide® (SAS EG) can easily combine data from tables or data sets by using a Graphical User Interface (GUI) PROC SQL Join to match on like columns or by using a Base SAS® Program Node DATA Step Merge to match on the same variable name. However, what do you do when tables or data sets do not contain like columns or the same variable name and a Join or Merge cannot be used? Well, we have the answer for you! We invite you to attend our presentation on the Joinless Join where we expand the power of SAS EG in a new way. You will learn how to design and utilize a Joinless Join to perform Join and Merge processing using tables or data sets which do not contain like columns or the same variable name. We will briefly review the various types of Joins and Merges and then quickly delve into the detailed aspects of how the Joinless Join can advance and enhance your data manipulation and analysis. We look forward to introducing you to the surprising paradox of the Joinless Join.

Paper BB-07

Monday:1:00 PM - 1:50 PM

Room: Legislative B

A Better Way to Flip (Transpose) a SAS® Dataset

Arthur Tabachneck, Xia Ke Shan, Robert Virgile and Joe Whitehurst

Many SAS® programmers have flipped out when confronted with having to flip (transpose) a SAS dataset, especially if they had to transpose multiple variables, needed transposed variables to be in a specific order, had a mixture of character and numeric variables to transpose, or if they needed to retain a number of non-transposed variables. Wouldn't it be nice to have a way to accomplish such tasks that was easier to understand and modify than PROC

TRANSPOSE, was less system resource intensive, required fewer steps and could accomplish the task as much as fifty times or more faster?

Paper BB-08

Monday:2:00 PM - 2:20 PM

Room: Legislative B

Anatomy of a Merge Gone Wrong

Joshua Horstman, James Lew

The merge is one of the SAS programmer's most commonly used tools. However, it can be fraught with pitfalls to the unwary user. In this paper, we look under the hood of the data step and examine how the program data vector works. We see what's really happening when datasets are merged and how to avoid subtle problems.

Customer Intelligence

Paper CI-08

Monday:12:30 PM - 1:20 PM

Room: House A

Big Data, Fast Processing Speeds

Gary Ciampa

As data sets continue to grow, it is important for programs to be written very efficiently to make sure no time is wasted processing data. This paper covers various techniques to speed up data processing time for very large data sets or databases, including PROC SQL, data step, indexes and SAS® macros. Some of these procedures may result in just a slight speed increase, but when you process 500 million records per day, even a 10 percent increase is very good. The paper includes actual time comparisons to demonstrate the speed increases using the new techniques.

Paper CI-06

Monday:1:30 PM - 2:20 PM

Room: House A

SAS® Treatments: One to One Marketing with a Customized Treatment Process

Dave Gribbin and Amy Glassman

The importance of sending the right message to the right person at the right time has never been more relevant than in today's cluttered marketing environment. SAS® Marketing Automation easily handles segment level messaging with out-of-the-box functionality. But how do you send the right message and the appropriately valued offers to the right person? And, how can an organization efficiently manage many treatment versions across distinct campaigns? This paper presents case studies of companies across different industries that send highly personalized communications and offers to their clientele using SAS® Marketing Automation. It describes how treatments are applied at a segment and a one-to-one level. It outlines a simple custom process that streamlines versioning treatments for reuse in multiple campaigns.

Hands-On Workshops

Paper HW-03

Monday:12:30 PM - 1:50 PM

Room: Judicial

Getting Excel-lent Data Generating SAS Datasets from a Directory of Spreadsheets

Ben Cochran

Sometimes a SAS programmer has to convert many spreadsheets into many SAS datasets. This presentation starts with a directory full of spreadsheets. Next, a basic program is written to read the spreadsheets one at a time into a series of

SAS datasets. Finally, the basic program is converted into a macro program that can dynamically read any number of spreadsheets into SAS datasets.

JMP

Paper JM-06

Monday:12:30 PM - 1:20 PM

Room: Senate B

When will it break? Exploring Product Reliability Using JMP.

Erich Gundlach

The reliability of your product strongly influences business success whether you're making semiconductors, light bulbs, automobiles, shoes, medical devices or jet engines. Reliability analysis can help ensure that a product functions as intended throughout its life, guaranteeing happy customers who provide repeat business as well as referrals. Dependable products also reduce warranty costs, which can otherwise quickly eat away at your profit margins. Join Erich Gundlach and learn more about JMP's Reliability platform. Find out how you can model system reliability and increase mean time between failures.

Paper JM-05

Monday:1:30 PM - 2:20 PM

Room: Senate B

Design of Experiments (DOE) using JMP

Charles Edwin Shipp

JMP has provided some of the best design of experiment software for years. The JMP team continues the tradition of providing state-of-the-art DOE support. In addition to the full range of classical and modern design of experiment approaches, JMP provides a template for Custom Design for specific requirements. The other choices include: Screening Design; Response Surface Design; Choice Design; Accelerated Life Test Design; Nonlinear Design; Space Filling Design; Full Factorial Design; Taguchi Arrays; Mixture Design; and Augmented Design. Further, sample size and power plots are available. We give an introduction to these methods followed by two examples with data. A lively discussion will follow.

Pharmaceutical Applications

Paper RX-08

Monday:12:30 PM - 1:20 PM

Room: House B

A Tutorial on PROC LOGISTIC

Arthur Li

In the pharmaceutical or health care industries, we often encounter data with dichotomous outcomes, such as having (or not having) a certain disease. This type of data can be analyzed by building a logistic regression model via the LOGISTIC procedure. In this paper, we will address some of the model building issues that are related to logistic regression. In addition, some statements in PROC LOGISTIC that are new to SAS® 9.2 and ODS statistical graphics relating to logistic regression will also be introduced in this paper.

Paper RX-09

Monday:1:30 PM – 1:50 PM

Room: House B

Using SAS in Clinical Programming Project Management from Excel Spreadsheets

Jean Crain

Tracking Clinical Programming Project Management as we do from Microsoft Excel© Spreadsheets can give you a wealth of information about your projects if you have standardized format and method of entry by programming and statistics. From our Excel© spreadsheets, using Base SAS©, SAS Macros and Proc Report we can retrieve the following data and present valuable information on how long a time period project spans, number of resources, time spent, productivity, and audit checking. The data obtained can provide multiple reports for analysis: Number of Programs Number of Outputs Number of Programs by Type (Analysis Dataset, Graph, Listing, Summary) Number of Outputs by Type (Analysis Dataset, Graph, Listing, Summary) Number of Programs by Employee (Programming, Validation and Statistical Validation and Review) Number of Outputs/Reviews by Employee Percent completion on projects by type: Original Programming, Validation Programming, Initial Statistical Review, Senior Statistical Review (number of outputs completed / number of outputs). First program completion date Last program completion date Number of calendar days (last date - first date), Overall and by Type. Number of work days (count unique dates), Overall and by Type. Audit check: Do Initials for Programming and Statistical Activities have a matching Name on Signature Page? Comment Data to analyze for trends: When programs/outputs need correction, does that indicate there are there areas for improvement in Requirements? Do the same issues come up from project to project? We can use this because we do have a standard for the spreadsheet tracking and entries by programming and statistics staff. Our hours are tracked elsewhere, but this will give you a good accounting of number of days for projections and it is an informative method of conducting process improvement. This is not complex programming but it demonstrates many functions in Base SAS, Proc Import of Excel spreadsheet data, and Proc Report to create multiple reports.

Paper RX-05

Monday:2:00 PM – 2:20 PM

Room: House B

Reading and Writing RTF Documents as Data: Automatic Completion of CONSORT Flow Diagrams

Arthur Carpenter

Whenever the results of a randomized clinical trial are reported in scientific journals, the published paper must adhere to the CONSORT (CONsolidated Standards Of Reporting Trials) statement. The statement includes a flow diagram, and the generation of these CONSORT flow diagrams is always problematic, especially when the trial is not the typical two-arm parallel design. Templates of the typical two-arm design flow diagram are generally available as RTF documents, however the completion of the individual fields within the diagram is both time consuming and prone to error. The SAS Macro language was used to read a RTF template file for the CONSORT flow diagram of choice, fill in the fields using information available to the SAS program, and then rewrite the table as a completed RTF CONSORT flow diagram. This paper describes the process of reading and writing RTF files.

Paper S1-07

Monday:12:30 PM - 12:50 PM

Room: Legislative A

Insurance Designation Randomization Application or How to Automate Your Bragging

Irvin Snider

Having taken over sixty-five insurance courses over the course of a twenty-five year home office career, I have acquired fifteen insurance designations to list behind my name. This causes problems because they usually never fit neatly on one line even if only the initials are used and to some, it appears ostentatious. To solve this problem, I have developed SAS code to randomly select only five of the designations at any one time to place behind my name in my emails. Thus the list appears to be fresh with each email and I do not appear to be bragging. This presentation will include the topics of random selection without replacement, placing this output in a macro variable via CALL SYMPUT and calling the macro variable in the email message.

Paper S1-08

Monday:1:00 PM - 1:20 PM

Room: Legislative A

Let the CAT Out of the Bag: String Concatenation in SAS 9

Joshua Horstman

Are you still using TRIM, LEFT, and vertical bar operators to concatenate strings? It's time to modernize and streamline that clumsy code by using the string concatenation functions introduced in SAS 9. This paper is an overview of the CAT, CATS, CATT, and CATX functions introduced in SAS 9.0, and the new CATQ function added in version 9.2. In addition to making your code more compact and readable, this family of functions also offers some new tricks for accomplishing previously cumbersome tasks.

Paper S1-03

Monday:1:30 PM - 1:50 PM

Room: Legislative A

Data Cleaning 101: An Analyst's Perspective

Anca Tilea and Deanna Chyn

On a daily basis, we are faced with data, both clean and dirty. SAS[®] offers a multitude of ways to clean and maintain data. It is up to us, the analysts, to choose the best way. Often times the choice we make depends on the analysis needed further down the road. When you start as a novice analyst you are familiar with the basics of data steps and procedures: SET statement, PROC MEANS, PROC FREQ, PROC SORT. We hope this paper will provide insight to a SAS[®] user that has this basic knowledge but wants to code more efficiently. We will introduce you to some basic SAS[®] procedures (PROC TRANSPOSE, PROC SQL) and SAS[®] tricks (using CALL SYMPUT, DO-LOOP, IF-ELSE, and PRX-functions) that are beyond a simple data step, but not too complicated to be understood by someone with basic SAS[®] skills. This paper is intended for the novice SAS user, with basic to intermediate skills and SAS 9.1 or above.

Power Trip: A Road Map of PROC POWER

Melissa Plets and Julie Strominger

Sample size and power analyses are extremely important components to consider when designing, planning and recruiting for prospective clinical research projects. Unfortunately, these are often considered to be scary and complicated calculations for research clinicians. PROC POWER offers a systematic solution to finding the balance between efficiency and conclusive results, while remaining simple enough for non-statisticians to use. More power is universally considered to be advantageous, even outside of the domain of statistics. It is especially important in statistics, as more power results in a higher probability that the null hypothesis will be rejected when it is false. Additionally, sample size is directly related to power. In general, a larger sample size will result in more accuracy, precision, and higher power. An important aspect of study design is maximizing power while remaining within the bounds of financial feasibility. This is where PROC POWER comes in, providing methods of determining sample sizes and power calculations. Throughout this paper, we will construct a road map of the POWER procedure to assist both statisticians and non-statisticians with the implementation of this POWERful SAS tool.

STATISTICAL DISCOVERY

Explore data visually and interactively.
And then share your discoveries with others.

JMP® software covers the most widely used statistical techniques in a most visual way. Its comprehensive menu of analytics includes:

- Univariate Analysis
- Bivariate Analysis
- Matched Pairs
- Regression, GLM and ANOVA
- Screening
- Nonlinear Modeling
- Neural Networks
- Gaussian Process Modeling
- Decision Trees
- Time Series Analysis
- Categorical Data Analysis
- Conjoint Analysis
- Cluster Analysis
- Multivariate Analysis
- Principal Components Analysis
- Discriminant Analysis
- Partial Least Squares
- Item Response Modeling
- Reliability and Survival Analysis
- Design of Experiments
- R and SAS® Integration



jmp.com/statistics



SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. © 2012 SAS Institute Inc. All rights reserved. 583571US.0212



Monday Night Event

Sponsored by SAS Institute, Inc.

This event is included in your conference registration. Complimentary transportation between the hotel and the conservatory will be available at your convenience. Shuttles running from 5:30 to 11:00 every half hour.

An Evening at the Conservatory

Franklin Park Conservatory and Botanical Gardens

Join us Monday evening for a lavish networking event at the stunning Franklin Park Conservatory and Botanical Gardens. Enjoy drinks, a sumptuous dinner, and plenty of time to network with your fellow SAS users while surrounded by the breathtaking natural beauty of over 400 species of plants on display indoors. Wander with your colleagues through various biomes such as desert, rainforest, the Himalayan mountains, and the Pacific Islands. Or step outside and stroll through 88 acres of abundant outdoor gardens.



Tuesday Morning Presentations

Advanced Analytics

Paper AA-08

Tuesday:8:00 AM - 8:20 AM

Room: Executive ABC

Improved Interaction Interpretation: Application of the EFFECTPLOT Statement and Other Useful Features in PROC LOGISTIC.

Robert Downer

The interpretation of fitted logistic regression models for students, collaborators or clients can often present challenges. Explanation of significant interactions among continuous predictors can be particularly awkward. The EFFECTPLOT statement and other features in PROC LOGISTIC of SAS/STAT can be useful aids in meeting these challenges. The CONTOUR and SLICEFIT options of this statement are particularly advantageous for more effective displays. Through logistic modeling of Titanic survival data, this paper also illustrates other ODS graphics and output from models with categorical and continuous predictors. Some basic familiarity with logistic regression is assumed.

Paper AA-09

Tuesday:8:30 AM - 8:50 AM

Room: Executive ABC

Analyzing Continuous Variables from Complex Survey Data Using PROC SURVEYMEANS

Taylor Lewis

This paper explores features available in PROC SURVEYMEANS to analyze continuous variables in a complex survey data set, where `complex` denotes a data set characterized by one or more of the following features: unequal weights, stratification, clustering, and finite population corrections. Using a real-world complex survey data set, this paper demonstrates the necessary syntax to have PROC SURVEYMEANS properly estimate totals, means, ratios, and quantiles, as well as their corresponding design-based measures of variability.

Paper AA-10

Tuesday:9:00 AM - 9:20 AM

Room: Executive ABC

Analyzing Categorical Variables from Complex Survey Data Using PROC SURVEYFREQ

Taylor Lewis

This paper explores features available in PROC SURVEYFREQ to analyze categorical variables in a complex survey data set, where `complex` denotes a data set characterized by one or more of the following features: unequal weights, stratification, clustering, and finite population corrections. Using a real-world complex survey data set, this paper illustrates the necessary syntax to calculate descriptive statistics and conduct select bivariate analyses, such as tests of association and the computation of odds ratios and relative risk statistics. Given alongside the syntax examples is some discussion of the theoretical reasons certain standard statistical techniques like the chi-square test of association require modification(s) when applied to complex survey data.

Paper AA-05

Tuesday:9:30 AM - 10:20 AM

Room: Executive ABC

Adventures in Path Analysis and Preparatory Analysis

Brandy Sinco and Phillip Chapman

This presentation will focus on the basics of path analysis, how to run path models in Proc CALIS, and how to use SAS to test for multivariate normality. Two estimation methods for path analysis: ML (Maximum Likelihood) and FIML (Full Information Maximum Likelihood) will be explained and compared. The results from Proc CALIS will also be compared with the SPSS AMOS module and the Mplus software. Further, SAS can be used to check data for the MAR (Missing At Random) assumption and to estimate a path model after imputing missing data. Power calculations for structural equation models can be done with Proc IML. The concepts behind SEM power calculation will be explained and an IML program to perform SEM power calculations will be presented. The data used for the analysis in this presentation is from the REACH-Detroit project, a culturally tailored Diabetes intervention for African American and Latino/a persons in inner city Detroit. This presentation is based on my masters research project at Colorado State University, with the title, USING PATH ANALYSIS TO TEST A HYPOTHESIS ON THE THEORY OF CHANGE IN HEMOGLOBIN A1C (HBA1C) AMONG CLIENTS IN A CULTURALLY TAILORED DIABETES INTERVENTION FOR AFRICAN AMERICANS AND LATINOS .

Paper AA-06

Tuesday:10:30 AM - 10:50 AM

Room: Executive ABC

Managing and Monitoring Statistical Models

Nate Derby

Managing and monitoring statistical models can present formidable challenges when you have multiple models used by a team of analysts over time. How can you efficiently ensure that you're always getting the best results from your models? In this paper, we'll first examine these challenges and how they can affect your results. We'll then look into solutions to those challenges, including lifecycle management and performance monitoring. Finally, we'll look into implementing these solutions both with an in-house approach and with SAS Model Manager.

Paper AA-07

Tuesday:11:00 AM - 11:20 AM

Room: Executive ABC

Introduction to Market Basket Analysis

Bill Qualls

Market Basket Analysis (MBA) is a data mining technique which is widely used in the consumer package goods (CPG) industry to identify which items are purchased together and, more importantly, how the purchase of one item affects the likelihood of another item being purchased. This paper will first discuss this traditional use of MBA, as well as introduce the concepts of support, confidence, and lift. It will then show how one company used MBA to analyze safety data in an attempt to identify factors contributing to injuries. Finally, a Base SAS macro which performs MBA will be provided and its usage demonstrated. Intended audience is anyone interested in data mining techniques in general, and in market basket analysis in particular, and while a Base SAS macro will be provided, no programming knowledge is required, and non-programmers will benefit from this paper.

Beyond the Basics

Paper BB-09

Tuesday:8:00 AM - 8:20 AM

Room: Governors B

Optimize Your Delete

Brad Richardson

ABSTRACT: Have you deleted a data set or two from a library that contains thousands of members using PROC DATASETS? If so, you probably have witnessed some wait time. To maximize performance, we have reinstated PROC DELETE as a SAS-supported procedure. One of the main differences between PROC DELETE deletion methods versus PROC DATASETS DELETE is that PROC DELETE does not need an in-memory directory to delete a dataset. So what does this mean exactly? This paper will explain all.

Paper BB-10

Tuesday:8:30 AM - 9:20 AM

Room: Governors B

Not All Equals are Created Equal: Nonstandard Statement Structures in the DATA Step

Arthur Carpenter

The expression is a standard building block of logical comparisons and assignment statements. Most of us use them so commonly that we do not give them a second thought. But in fact they definitely do deserve that second thought. A more complete understanding of their construction and execution can greatly expand our ability to more fully take advantage of this fundamental component of the SAS® Language. Once we understand the basic form of the expression and how it is used in various statements, we can use this understanding to create statement forms that would otherwise appear to be illegal or just plain wrong. Further and perhaps even more importantly this deeper understanding can help to prevent us from committing errors in logic.

Paper BB-11

Tuesday:9:30 AM - 9:50 AM

Room: Governors B

Efficient and Smart Ways to Manage Datasets for Clinical Data- Let SAS Do the Dirty Laundry!

Gowri Madhavan and Alan Leach

The ASQ (ages and stages) is an important questionnaire that is delivered to parents to assess any developmental delay in children 9-24 months. If the child fails an ASQ, he/she is referred to developmental agency at the initial stages for treatment. A child can have several well child care visits and can either pass or fail an ASQ; a child can pass once and fail subsequently or vice versa. It is important to capture the number of tests administered and the most recent test results for failures. However, dealing with these test results on a weekly basis for reporting public health information can be daunting! A patient's clinical history can vary and we don't always know what datasets to expect. Fortunately SAS provides us ways to dynamically intake and process this information with a minimum of effort to maintain. This paper discusses time saving techniques used to manage the intake and processing of numerous clinical data sources. Using macros we will read in number of clinical datasets which can vary in number and name from week to week. Through the use of PROC TRANPOSE we will reshape the data for further processing. Using DICTIONARY.TABLES to capture information about our data we will initialize MACRO variables dynamically build ARRAYS and DO LOOPS to process each patient's ASQ test results history and categorize them per requirements. Then by summarizing these results and exporting into Excel spreadsheets we can automatically email results to stakeholders.

Paper BB-12

Tuesday:10:00 AM - 10:50 AM

Room: Governors B

Life Imitates Art: ODS Output Data Sets that Look like Listing Output

Dylan Ellis

Confusingly, the output data set from a summary procedure often looks nothing like the display in our output destination. Have you ever wished for an output data set that looks like the cross-tabulation you see in the listing window? Have you ever run a series of one-way frequencies on a list of variables and wished for a more compact tabular display? This presentation will show how we can leverage string functions and the automatic `_TYPE_` variable from PROC TABULATE to reshape the output data set into a more intelligible and practical table visualization. Readers should be familiar with the basics of how to specify a categorical frequency table using Proc Freq or Proc Tabulate.

Paper BB-13

Tuesday:11:00 AM - 11:20 AM

Room: Governors B

Creating Formats on the Fly

Suzanne Dorinski

The Census Bureau conducts the Common Core of Data surveys for the National Center for Education Statistics annually. We have written SAS programs to automate the database documentation. We try to avoid including hard-coded values in the programs. Thanks to a record layout spreadsheet, the analysts can quickly update the survey metadata outside the SAS programs. This paper explains how SAS can read the record layout spreadsheet to create formats on the fly. The analysts can update the values as changes occur over time without having to worry about writing correct SAS syntax. Behind the scenes, SAS is using dictionary views, macros, ODS output, PROC TEMPLATE, PROC FORMAT, the ODS Report Writing Interface, and RTF to create the desired results. This paper uses syntax for SAS 9.2, written for programmers at the intermediate level.

BI Applications, Systems Architecture and Administration

Paper BI-01

Tuesday:8:00 AM - 8:50 AM

Room: House A

Using the SAS Projman Application for Scheduling Projects

Jon Patton

Abstract: SAS/OR software has four major procedures that can be used to manage projects. The CPM and PM procedures are used to schedule tasks for a project subject to precedence, time, and resource constraints. The PM procedure is the interactive version of the CPM procedure. The Gantt procedure displays this schedule, and the Netdraw procedure displays the project network consisting of these tasks. These four procedures are integrated into the Projman application which is a friendly graphical user interface included as part of the SAS/OR software. This tutorial will cover the usage of these four procedures and the Projman application. The early part of the tutorial will cover the definition of terminology that is critical for understanding the output results. Then example projects containing resource, time, and precedence constraints will be scheduled using the Projman application. Finally two SAS macros that allows the conversion back and forth between data of Microsoft Project and the PM procedure will be discussed. This presentation is for SAS users of all skill levels.

Seamless Dynamic Web (and Smart Device!) Reporting with SAS®

DJ Penix

The SAS® Business Intelligence platform provides a wide variety of reporting interfaces and capabilities through a suite of bundled components. SAS® Enterprise Guide®, SAS® Web Report Studio, SAS® Add-In for Microsoft Office, and SAS® Information Delivery Portal all provide a means to help organizations create and deliver sophisticated analysis to their information consumers. However businesses often struggle with the ability to easily and efficiently create and deploy these reports to the web and smart devices. If it is done, it is usually at the expense of giving up dynamic ad-hoc reporting capabilities in return for static output or possibly limited parameter-driven customization. The obstacles facing organizations that prevent them from delivering robust ad-hoc reporting capabilities on the web are numerous. More often than not, it is due to the lack of IT resources and/or project budget. Other failures may be attributed to breakdowns during the reporting requirements development process. If the business unit(s) and the developers cannot come to a consensus on report layout, critical calculations, or even what specific data points should make up the report, projects will often come to a grinding halt. This paper will discuss a solution that enables organizations to quickly and efficiently produce SAS reports on the web and your mobile device - in less than 10 minutes! It will also show that by providing self-service functionality to the end users, most of the reporting requirement development process can be eliminated, thus accelerating production-ready reports and reducing overall maintenance costs of the application. Finally, this paper will also explore how the other tools on the SAS Business Intelligence platform can be leveraged within an organization.

Beating Gridlock: Parallel Programming with SAS® Grid Computing and SAS/CONNECT®

Jack Fuller

Long running SAS jobs often include sets of independent subtasks that can be split up and distributed across a SAS Grid. When these subtasks are then run in parallel the total run time will usually increase; however, total elapsed time can often be made to decrease. This paper will present an introduction to parallel processing using SAS Grid and SAS Connect with an emphasis on the following: when to use parallel processing, how to use parallel processing and points to consider when implementing parallel processing.

Building a dynamic SAS HTML report with JavaScript and SAS Tagsets

Mike Libassi

The ability to generate dynamic reports using HTML tagsets allows the report end user to manipulate the output. When we add a dynamic process to generate that output (such as running the report for a selected date, or date range) we add even greater flexibility to the report. This paper and presentation covers the process used in building this framework.

Going to SAS EG from SAS PC ?

Ira Shapiro

The purpose of this presentation is to provide a step by step process for the user to follow to successfully migrate their work from SAS/PC to SAS Enterprise Guide. It also points out some of the very basic features of Enterprise Guide such as the definition of a project, how to run a project, how to run a portion of code and how to run a program within a project.

Customer Intelligence

Residential Energy Efficiency and the Principal-Agent Problem

Ryan Anderson

Investments in residential energy-efficiency are associated with both positive externalities (reduced greenhouse gas emissions, reduced need for new power-generating capacity) and private cost savings. However, these investments lag far behind those levels predicted by conventional cost-benefit analysis. This paper explores one potential explanation for this efficiency gap, the principal-agent (PA) problem as it applies to rental housing. I employ SAS 9.3 to apply a logit model to data from the 2009 Residential Energy Consumption Survey and find that the PA problem is particularly pronounced in regards to weatherization improvements, a result that has implications for public policy. Somewhat advanced SAS techniques were used to report the results of this regression in terms of marginal effects. A basic familiarity with statistical processes and microeconomic theory is useful in reviewing this paper.

sasNerd®: Better Searches = Better Results

Kirk Paul Lafler, Richard W. La Valley, Lex Jansen and Charles Edwin Shipp

As SAS®- and JMP®-related content continues to grow to new levels the world's leading search engines (Google®, Bing®, and Yahoo®) and their proprietary software, organizes this information and makes it useful and accessible to everyone. Growing numbers of users benefit from the speed, accuracy, organization, and reliability of these powerful search engines, as well as the capabilities that LexJansen.com provides as a web portal for finding relevant SAS and JMP content. Due to the importance of finding desired content whenever it's needed, users turn to their favorite search engine, or LexJansen.com with its repository of 25,000-plus published papers, for their search needs. This paper introduces the user community to a new application called, sasNerd® that is designed to help find published papers and other searchable content from SAS Global Forum (SGF) and SAS Users Group International (SUGI) conferences; MWSUG, NESUG, PNWSUG, SCSUG, SESUG, and WUSS regional conferences; and PharmaSUG, PhUSE, and CDISC special-interest conferences.

Paper CI-09

Tuesday:9:30 AM - 10:20 AM

Room: Senate B

Applying Customer Analytics to Promotion Decisions

Wanda Shive

For years, retailers have struggled to measure the effectiveness of their promotional advertising efforts. Harnessing the big data within their customer and transaction files continues to be a major challenge. Approaches for gleaning true customer insights from that data are becoming more common. Measuring total shopping behavior in conjunction with specific promotions provides a better understanding of the overall impact on profitability. This paper describes how retailers are utilizing customer analytics to measure the effect that mass promotions have on the total basket spend of customers and to identify the most relevant offers for each individual customer.

In Conference Training

Paper CT-01

Tuesday:8:00 AM - 11:20 AM

Room: Legislative AB

Top 10 SAS Best Programming Practices They Didn't Teach You in School

Charu Shankar

How can programming time, I/O, CPU and memory be reduced? What is the data analysts' #1 rule? What are three questions you need to answer? What new features will help improve performance? These are age-old questions that have had data analysts thinking since the dawn of the first SAS program. Get all the answers in this new and informative seminar.

Data Visualization and Graphics

Paper DV-01

Tuesday:8:00 AM - 8:20 AM

Room: Senate A

Using PROC SGPLOT Over PROC GPLOT

Shruthi Amruthnath

SAS® offers different statistical graphic procedures for data visualization and presentation. SAS® 9.2 brought out new family of template-based graphical procedures to create high-quality graphics called Statistical Graphics (SG) procedures. These procedures are so powerful that, more complex information can be presented effectively with minimal coding. This paper will focus on employing SGPLOT versus GPLOT; applying SGPLOT to financial data to create reports for trending data, consolidated reports and yearly reports; and managing, displaying and styling procedural output using ODS PDF. The SGPLOT procedure has different types of graphical figures like bar charts, line graphs and scatter plots. This paper explains how to produce line graphs like line plot and area plot by setting different options. In this presentation, each of these topics is illustrated using different techniques with an example.

Paper DV-03

Tuesday:8:30 AM - 9:20 AM

Room: Senate A

The Graph Template Language: Beyond the SAS/GRAPH® Procedures

Jesse Pratt

The SGPLOT and SGPANEL procedures are powerful tools that are capable of producing many types of high quality graphs; however, these procedures have some limitations. What happens when one is asked to specifically produce a graph that these procedures cannot create? The Graph Template Language (GTL) is much more flexible when it comes to creating customized displays. This paper presents situations where the SGPLOT and SGPANEL procedures break down, then briefly introduces GTL, and finally uses GTL to generate the displays not possible in PROC SGPLOT and PROC SGPANEL.

Paper DV-04

Tuesday:9:30 AM - 11:20 AM

Room: Senate A

SAS® Tile Charts: Thousands of Business Tips with One Click

Martha Hays

The ever growing complexity, increasing size and availability of business information makes getting to the critical issues a major challenge. A well organized SAS Tile Chart will grab and focus your attention on those critical issues that require you to take action. This seminar outlines the capabilities and business uses of the SAS Tile Chart. The Tile Chart makes very effective use of the color, size and data tips associated with multiple (drillable) tiles presented in a hierarchical grid to communicate business intelligence information for volumes of relevant data that is available for analysis.

Hands On Workshops

Paper HW-04

Tuesday:8:00 AM - 9:20 AM

Room: Judicial

Using SAS® ODS Graphics

Chuck Kincaid

This presentation will teach the audience how to use SAS® ODS Graphics. Now part of Base SAS, ODS Graphics are a great way to easily create clear graphics that allow any user to tell their story well. SGPLOT and SGPANEL are two of the procedures that can be used to produce powerful graphics that used to require a lot of work. The core of the procedures are explained, as well as the options available. Furthermore, we explore the ways to combine the individual statements to make more complex graphics that tell the story better. Any user of Base SAS on any platform will find great value from the SAS ODS Graphics procedures.

Paper HW-05

Tuesday:9:30 AM - 10:50 AM

Room: Judicial

Building a Better Bar Chart with SAS® Graph Template Language

Perry Watts

This workshop combines instructions for chart building with principles defined by the statistical graphics experts: Edward Tufte, William Cleveland, and Naomi Robbins. Step-by-step instructions are provided for building basic and group charts that display frequencies, sums, percents, and means with associated confidence intervals. Group charts

are further subdivided into repeating and non-repeating categories. For repeating group charts varying bar displays such as stacked, cluster, and nested will also be reviewed. Exercises make use of the BARCHART statement included in the SAS® 9.3 Graph Template Language Reference manual. Along the way, comparisons are made between GTL's BARCHART statement and the GCHART procedure in SAS/GRAPH software. The data used in the examples come either from the SASHELP library or from pre-summarized data read into WORK data sets with a DATALINES statement. While there may not be time in the workshop to cover advanced topics, two enhanced bar charts will be covered in the paper. Intermediate to advanced SAS programmers with experience using any graphics software package will get the most out of this presentation.

SAS 101

Paper S1-10

Tuesday:8:00 AM - 8:50 AM

Room: Governors CDE

Effectively Utilizing Loops and Arrays in the DATA Step

Arthur Li

The implicit loop refers to the DATA step repetitively reading data and creating observations, one at a time. The explicit loop, which utilizes the iterative DO, DO WHILE, or DO UNTIL statements, is used to repetitively execute certain SAS® statements within each iteration of the DATA step execution. Utilizing explicit loops is often used to simulate data and to perform a certain computation repetitively. However, when an explicit loop is used along with array processing, the applications are extended widely, which includes transposing data, performing computations across variables, etc. Being able to write a successful program that uses loops and arrays, one needs to know the contents in the program data vector (PDV) during the DATA step execution, which is the fundamental concept of DATA step programming. This paper will cover the basic concepts of the PDV, which is often ignored by novice programmers, and then will illustrate how utilizing loops and arrays to transform lengthy code into more efficient programs.

Paper S1-11

Tuesday:9:00 AM - 9:20 AM

Room: Governors CDE

SAS: Tips and Tricks

Audrey Yeo

Using SAS® in the work environment is different from using SAS in the educational environment. Data in the work environment is neither as perfect nor as simple as data in the educational environment. This paper will highlight some tips and tricks that will help a new SAS user, whether a recent graduate or just new to SAS, deal with some common but challenging problems.

Paper S1-12

Tuesday:9:30 AM - 9:50 AM

Room: Governors CDE

Data Presentation 101: An Analyst's Perspective

Anca Tilea and Deanna Chyn

You are done with the tedious task of data cleaning, and now the fun begins. You have several results and must present them in a useful manner to a statistician (if you are lucky) or a non-statistician (as is usually the case). SAS® provides a

multitude of resources to do this very thing: HISTOGRAM statement in PROC UNIVARIATE and/or PROC SGPLOT, PROC REPORT, Output Delivery System statements for various PROCs, and so on. This paper will describe a few of these methods that we, as data analysts, have found to be most helpful when presenting results. This paper is intended for the novice SAS user, with basic to intermediate skills and SAS" 9.2 or above.

Paper S1-13

Tuesday:10:00 AM - 10:20 AM

Room: Governors CDE

I Heart SAS Users

Joanne Ellwood

In my 20 some years experience, I have found SAS users to be smart, helpful and highly motivated. I would like to share some of the experiences I have had in supporting SAS users I am very humbled whenever I find a SAS user snagged in a simple error and am able to be of assistance in helping to resolve the error. It is an honor and a privilege to work with such magnificent people as the much heart d SAS users. This paper is targeted for anyone who has ever used SAS.

Paper S1-14

Tuesday:10:30 AM - 10:50 AM

Room: Governors CDE

A Poor/Rich SAS® User s Proc Export

Arthur Tabachneck, Tom Abernathy, Randy Herbison and Matthew Kastin

Have you ever wished that with one click you could copy any SAS® dataset, including variable names, so that you could paste the text into a Word file, powerpoint or spreadsheet? You can and, with just base SAS, there are some little known but easy to use methods that are available for automating many of your (or your users) common tasks.

Paper S1-15

Tuesday:11:00 AM - 11:20 AM

Room: Governors CDE

Writing Macro Do Loops with Dates from Then to When

Ronald Fehd

Dates are handled as numbers with formats in SAS(R) software. The SAS macro language is a text-handling language. Macro \%do statements require integers for their start and stop values. This article examines the issues of converting dates into integers for use in macro \%do loops. Two macros are provided: a template to modify for reports and a generic calling macro function which contains a macro \%do loop that can return items in the associative array of dates. Example programs are provided which illustrate unit testing and calculations to produce reports for simple and complex date intervals.

Tuesday Afternoon Presentations

Advanced Analytics

Paper AA-14

Tuesday:1:00 PM - 1:50 PM

Room: Executive ABC

Information Value Statistic

Bruce Lund and David Brotherton

The Information Value (IV) statistic is a popular screener for selecting predictor variables for binary logistic regression. Familiar, but perhaps mysterious, guidelines for deciding if the IV of a predictor X is high enough to use in modeling are given in many textbooks on credit scoring. For example, these texts say that $IV > 0.3$ shows X to be a strong predictor. These guidelines must be considered in the context of binning. A common practice in preparing a predictor X is to bin the levels of X to remove outliers and reveal a trend. But IV decreases as the levels of X are collapsed. This paper has two goals: (1) Provide a method for collapsing the levels of X which maximizes IV at each iteration and (2) show how the guidelines (e.g. $IV > 0.3$) relate to other measures of predictive power. All data processing was performed using Base SAS®. The presentation will be appropriate for predictive modeling practitioners who use PROC LOGISTIC.

Paper AA-15

Tuesday:2:00 PM - 3:50 PM

Room: Executive ABC

Structural Equation Modeling Using the CALIS Procedure in SAS/STAT® Software

Yiu-Fai Yung

The CALIS procedure in SAS/STAT software is a general structural equation modeling (SEM) tool. This workshop introduces the general methodology of SEM and applications of PROC CALIS. Background topics such as path analysis, confirmatory factor analysis, measurement error models, and linear structural relations (LISREL) are reviewed. Applications are demonstrated with examples in social, educational, behavioral, and marketing research. More advanced SEM techniques such as the analysis of total and indirect effects and full information maximum likelihood (FIML) method for treating incomplete observations are also covered. This workshop is designed for statisticians and data analysts who want an overview of SEM applications using the CALIS procedure in SAS/STAT 9.22 and later releases. Attendees should have a basic understanding of regression analysis and experience using the SAS language. Previous exposure to SEM is useful but not required. Attendees will learn how to use PROC CALIS for (1) specifying structural equation models with latent variables, (2) interpreting model fit statistics and estimation results, (3) computing and testing total and indirect effects (4) using the FIML method for treating incomplete observations.

Blackbelt SAS

Paper 00-01

Tuesday:1:00 PM - 1:20 PM

Room: Governors C

Using Microsoft® Windows® DLLs within SAS® Programs

Rajesh Lal

SAS has a wide variety of functions and call routines available. More and more Operating System level functionality has become available as part of SAS language and functions over the versions of SAS. However, there is a wealth of other Operating System functionality that can be accessed from within SAS with some preparation on the part of the SAS

programmer. Much of the Microsoft Windows functionality is stored in easily re-usage system DLL (Dynamic Link Library) files. This paper describes some of the Windows functionality which may not be available directly as part of SAS language and methods of accessing that functionality from within SAS code. Using the methods described here, practically any Windows API should become accessible. User created DLL functionality should also be accessible to SAS programs.

Paper 00-02

Tuesday:1:30 PM - 2:20 PM

Room: Governors C

Macro Design Ideas, Theory or Template

Ronald Fehd

This paper provides a set of ideas about design elements of SAS(R) macros. This article is a checklist for programmers who write or test macros.

Paper 00-03

Tuesday:2:30 PM - 2:50 PM

Room: Governors C

SAS® Commands PIPE and CALL EXECUTE; Dynamically Advancing from Strangers to Best Friends

Kent Phelps, Ronda Phelps and Kirk Lafler

Communication is the foundation of all relationships, including your relationship with SAS® and the Server/PC/Mainframe (S/P/M). There are times when you need to communicate with the S/P/M through the UNIX, Windows, or z/OS Operating System (OS) to obtain important data to use in your various projects. To communicate with the S/P/M you will ideally design your SAS program to request, receive, and utilize data to automatically create and execute Dynamic Code. Our presentation highlights the powerful SAS partnership which occurs when the PIPE and CALL EXECUTE commands are surprisingly and creatively used together within SAS Enterprise Guide® (EG) Base SAS® Program Nodes. You will have the opportunity to learn how 1,259 time-consuming Manual Steps are amazingly replaced with only 3 time-saving Dynamic Automated Steps. We look forward to introducing you to the powerful PIPE and CALL EXECUTE partnership Your newest BFF (Best Friends Forever) in SAS.

Paper 00-04

Tuesday:3:00 PM - 3:50 PM

Room: Governors C

Top 10 SAS Best Programming Practices They Didn't Teach You in School

Charu Shankar

How can programming time, I/O, CPU and memory be reduced? What is the data analysts' #1 rule? What are three questions you need to answer? What new features will help improve performance? These are age-old questions that have had data analysts thinking since the dawn of the first SAS program. Get all the answers in this new and informative seminar.

Banking and Financial Services

Paper FS-07

Tuesday:1:00 PM - 1:50 PM

Room: Senate A

Advanced Multithreading Techniques for Performance Improvement of SAS® Processes

Viraj Kumbhakarna

Paper discusses a host of new SAS 9.2 functionality related to parallel processing. Parallel processing refers to processing that is handled by multiple CPUs simultaneously. This technology takes advantage of hardware that has multiple CPUs, called SMP computers, and provides performance gains for two types of SAS processes: " threaded I/O " threaded application processing. Paper explores use of newer, faster, practically applicable parallel processing techniques supported by SAS 9.2 and later versions that can be used for processing large volume of data in parallel on AIX UNIX as well as windows SAS environments. It further dwells on remediating some practical limitations such as: a. identifying efficient ways to implement parallel processing, b. determining optimum number of threads to process in parallel, c. using newer SAS 9.2 procedures (such as SCAPROC) for multithreading SAS, d. analyzing advanced SAS 9 support for spawning and manage multiple threads. Paper proposes multiple techniques to execute SAS processes in parallel on AIX UNIX platform. It also explores use of Piping which is an extension of the MP connect functionality to address pipeline parallelism. Piping enables user s to overlap the execution of SAS data steps and/or certain SAS procedures. This is accomplished by spawning one SAS session to run one data step or proc and pipes its output through a TCP/IP socket as input into another SAS session running another data step or proc. This pipeline can be extended to include any number of steps and can even extend between different physical machines. Paper also discuss various techniques used to analyze and monitors execution of parallel threads on AIX UNIX servers, to analyze the current server performance and use it to determine optimal number of threads to be submitted in parallel thereby taking into consideration real time load balancing and process execution. In conclusion we compare and contrast benefits, cost overhead and return over investment (ROI) of implementing parallel processing for practical complex SAS processes processing huge data volumes (in the author s experience, over 5 million observations and up to 2200 variables) and argue benefits of parallel processing in terms of improved performance, reduced processing times and reduced I/O.

Paper FS-08

Tuesday:2:00 PM - 2:20 PM

Room: Senate A

Changing a Static Condition to a Dynamic Data-Driven Field with SAS®

Misty Johnson

SAS® programs that are data driven are efficient and reduce the possibility of error. The design of SAS programs should be carefully considered to incorporate the ease of use and ability to accommodate future change. This paper describes the edit of a SAS program to change a static condition to a dynamic, data-driven variable. Also discussed is the effect of the edit upon the business process and editing SAS programs to invoke minimum process change. The SAS program demonstrates the use of macro variables assigned with the %LET statement, the LIBNAME statement with an Excel engine and the mixed option, literals when referring to Excel sheet names and the creation of output text files with the FILE statement. Methods described in this paper use base SAS and the Access to PC Files module and is geared toward beginning to intermediate SAS users.

Paper FS-09

Tuesday:2:30 PM - 3:20 PM

Room: Senate A

Addressing Fraudulent Payment Activity with Advanced Decision Management Analytics

Rex Pruitt

Advanced Decision Management Analytics is used in many areas of business today resulting in millions of dollars in new revenue and/or expense reduction including fraud loss mitigation. The successful implementation of these analytic strategies is a significant problem in many businesses. With Big Data challenges or lack of proficiency with sophisticated software solutions, projects tend to stall and never get implemented. This presentation includes examples of the successful implementation of decision management analytic software solutions specific to fraudulent payment activity:

1. Validation of 3rd party NSF models 2. Determining check float criteria based on consumer behavior models 3. Fraud Loss mitigation modeling that includes payment patterns and attributes 4. Collections call routing based on propensity to pay 5. Forecasting detailed profitability components including fraud loss This presentation will reference Base SAS, Enterprise Miner, and Forecast Server solutions for users at any skill level as the focus is on methodology and best practices. SAS software usage will be discussed and not demonstrated.

Paper FS-10

Tuesday:3:30 PM - 3:50 PM

Room: Senate A

Data Quality Governance for Data Sourcing and Analytics Team

Anand Kumar

Having data that are consistent, reliable and well linked is one of the biggest challenges facing by financial institutions. The paper describes how SAS data management platform helps to connect people, process and technology to deliver consistent results for the data sourcing and analytics team and minimize the cost and time in the development life cycle. The paper concludes with the best practices learned from various enterprise data initiatives.

Business Intelligence

Paper BI-05

Tuesday:1:00 PM - 1:50 PM

Room: House A

Transitioning from Batch and Interactive SAS to SAS Enterprise Guide

Brian Varney

Although the need for access to data and analytical answers remains the same, the way we get from here to there is changing. Change is not always adopted nor welcomed and it is not always voluntary. This paper intends to discuss the details and strategies for making the transition from traditional SAS software usage to SAS Enterprise Guide. These details and strategies will be discussed from the company as well as the individual SAS developer level. The audience for this paper should be traditional SAS coders who are just getting exposed to SAS Enterprise Guide but still want to write code.

Paper BI-06

Tuesday:2:00 PM - 2:50 PM

Room: House A

Improving your Relationship with SAS EG: Tips from SAS Tech Support

Jennifer Bjurstrom

SAS® Enterprise Guide has proven to be a very beneficial tool for both novice and experienced SAS® users. Because it is such a powerful tool, SAS Enterprise Guide has risen in popularity over the years. As a result, SAS Technical Support

consultants field many calls from users who want to know the best way to use the application to accomplish a task or to obtain the results they want. This paper encompasses many of those tips that SAS Technical Support has provided to customers over the years. These tips are designed to improve your proficiency with SAS Enterprise Guide in the areas of workflow preferences, data manipulation, scheduling projects, logging, layout, and more.

Paper BI-04

Tuesday:3:00 PM - 3:50 PM

Room: House A

SAS Stored Processes on the Web Building Blocks

Mark Roberts

This paper explains, in a step-by-step format, the concepts, issues, techniques, and code needed to develop a web-based application from stored processes. Running an application on the web allows users to execute powerful SAS procedures without knowing SAS, without having SAS on their desktop, and without a SAS license. It also allows companies to distribute the application to many locations. If the application needs a change, it is fixed in one place, loaded to the SAS portal, and all locations see the corrected code at the same time. A SAS developer needs only an understanding of stored processes, prompts, macros, a little HTML, and a little knowledge of Proc Reports to achieve this. The paper explains the concepts and techniques clearly via an example that runs consistently throughout. The code used for each concept is also included. The author put everything he learned developing his own application together in one place, something he couldn't find when he first started. One of his objectives for this paper is for the audience to be able to apply the code and successfully develop their own web-based applications with a minimum of difficulty and without spending hours searching the web and SAS documentation.

Data Visualization and Graphics

Paper DV-05

Tuesday:1:00 PM - 1:50 PM

Room: Senate B

Using ODS PDF, Style Templates, Inline Styles, and PROC REPORT with SAS® Macro Programs

Patrick Thornton

A production system of SAS macro programs is described that modularize the generation of syntax to produce client-quality reports of descriptive and inferential results in a PDF document. The reusable system of macros include programs that save all current titles, footnotes, option settings, establish standard titles, footnotes and option settings, and initially create the PDF document. Macro programs may be called to generate PROC REPORT syntax to produce various tables of descriptive and inferential results with customized bookmarks, table numbers and footnotes. A custom style template is used to determine the look of the whole document, and a macro program of inline style definitions is used to define the defaults for each type of table. A version of the style macro program may be customized to accommodate the needs of each project, and inline style parameters may be modified for any given macro call. A macro program to end the PDF creates a standard data documentation page, and restores all original titles, footnotes and option settings. This paper is designed for the intermediate to advanced SAS programmer using Foundation SAS Software on a Windows operating system.

Paper DV-06

Tuesday:2:00 PM - 2:20 PM

Room: Senate B

Bordering on Success with PROC GMAP in SAS®: Utilizing Annotate Datasets to Enhance Your Maps

Kathryn Schurr and Jonathan Wiseman

PROC GMAP in SAS® is a valuable tool in the visualization of data. GMAP allows users to geographically represent their data so that audiences can relate on multiple levels at once. One limitation of PROC GMAP that has been identified is its

inability to allow for multiple geographic regions (such as zip codes and counties) to be plotted at once, especially if they do not have the shape of a polygon. This paper will discuss how PROC GMAP in conjunction with the %MAPLABLE and %ANNOTATE macros can create a map in SAS® that annotates various locations on the map, creates an (x,y) coordinate grid corresponding to the shapefile being used, presents the capability to move singular labels to a more desirable location, and draws borders of differing regions regardless of shape.

Paper DV-07

Tuesday:2:30 PM - 3:20 PM

Room: Senate B

A Concise Display of Multiple Response Items

Patrick Thornton

Surveys often contain multiple response items, such as language where a respondent may indicate that she speaks more than one language. In this case, an indicator variable (1=Yes, 0=No) is often created for each language category. This paper shows how a concise tabulation of the count and percent of respondents with a Yes on one or more indicator variables may be obtained using PROC TABULATE and a MULTILABEL format. A series of indicator variables is used to create a binary variable and its base-10 equivalent, and a MULTILABEL format is created to properly aggregate observations with a Yes on two or more indicator variables.

Paper DV-06

Tuesday:3:30 PM - 3:50 PM

Room: Senate B

Time Contour Plots

David Corliss

A time contour plot is two-dimensional color chart for visualizing changes in a population or system over time. Data for one point in time appear as a thin horizontal band of color. Bands for successive periods are stacked up to make a two-dimensional, with the vertical direction showing changes over time. As a system evolves over time, different kinds of events have different characteristic patterns. Creation of Time Contour Plots is explained step by step. Examples are given in astrostatistics, biostatistics, econometrics and demographics.

Hands On Workshops

Paper HW-06

Tuesday:1:00 PM - 2:20 PM

Room: Judicial

Hands-on SAS® Macro Programming Tips and Techniques

Kirk Paul Lafler

The SAS® Macro Language is a powerful tool for extending the capabilities of the SAS System. This hands-on workshop presents numerous tips and tricks related to the construction of effective macros through the demonstration of a collection of proven Macro Language coding techniques. Attendees learn how to process statements containing macros; replace text strings with macro variables; generate SAS code using macros; manipulate macro variable values with macro functions; handle global and local variables; construct arithmetic and logical expressions; interface the macro language with the DATA step and SQL procedure; store and reuse macros; troubleshoot and debug macros; and develop efficient and portable macro language code.

Paper HW-07

Tuesday:2:30 PM - 3:50 PM

Room: Judicial

Using PROC FCMP to the Fullest: Getting Started and Doing More

Arthur Carpenter

The FCMP procedure is used to create user defined functions. Many users have yet to tackle this fairly new procedure, while others have only attempted to use only its simplest options. Like many tools within SAS®, the true value of this procedure is only appreciated after the user has started to learn and use it. The basics can quickly be mastered and this allows the user to move forward to explore some of the more interesting and powerful aspects of the FCMP procedure. Starting with the basics of the FCMP procedure, this paper also discusses how to store, retrieve, and use user defined compiled functions. Included is the use of these functions with the macro language as well as with user defined formats. The use of PROC FCMP should not be limited to the advanced SAS user; even those fairly new to SAS should be able to appreciate the value of user defined functions.

Rapid Fire

Paper RF-04

Tuesday:1:00 PM - 1:10 PM

Room: House B

Copy and Paste from Excel to SAS®

Arthur Tabachneck

Most, if not all, of us are far too familiar with the problems that one can confront when trying to import an Excel workbook into a SAS® dataset. Numeric fields might be imported as character fields, dates might be imported as either character fields or dates that are sixty (60) years earlier than what they actually represent, and fields containing time values might be imported as representing one-half of a second when they actually represent 43,200 seconds. While such discrepancies can be corrected by following the use of PROC IMPORT with a carefully written datastep, the present paper presents an alternative that only requires one step, uses less code, only requires base SAS and, on our test data, ran almost fifty (50) times faster than using PROC IMPORT.

Paper RF-05

Tuesday:1:15 PM - 1:25 PM

Room: House B

Increase Your Productivity by Doing Less

Arthur Tabachneck, Xia Ke Shan, Robert Virgile and Joe Whitehurst

Using a keep dataset option when declaring a data option has mixed results with various SAS® procedures. It might have no observable effect when running PROC MEANS or PROC FREQ but, if your datasets have many variables, it could drastically reduce the time required to run some procs like PROC SORT and PROC TRANSPOSE. This paper describes a fairly simple macro that could easily be modified to use with any proc that defines which variables should be kept and, as a result, make your programs run 12 to 15 times faster.

Paper RF-01

Tuesday:1:30 PM - 1:40 PM

Room: House B

Getting Wild with Imports

Kathryn Schurr and Jonathan Wiseman

ABSTRACT Ever have many similar datasets to import and not enough time to come up with a good macro to import and merge them all together? Using a wildcard import in SAS® v9.3 will solve all of your woes. This paper presents an efficient technique for importing multiple data files under one folder at the same time. Then it will further address the wildcard import as a SAS® Macro that can be used time and again for various folders containing differing data. This is a great tool for records kept electronically per facility, per student, or per patient; and especially when an extravagant query based data warehouse is unavailable.

Paper RF-02

Tuesday:1:45 PM - 1:55 PM

Room: House B

Does foo Pass-Through? SQL Coding Methods and Examples using SAS® software

Stephen Crosbie

Rapid Fire: Processing data between SAS® software and a DBMS, such as Netezza® or DB2®. Application of the SQL Procedure with the CONNECT Statement (requires SAS/CONNECT® software) and use of the PROC SQL Pass-Through Facility. While a SAS library reference gives PROC SQL access to DBMS tables (requires SAS/ACCESS® software), using Pass-Through SQL coding instead can often provide faster results. This paper provides useful examples of queries that were passed-through to a DBMS to take advantage of greater processing resources or better query optimization.

Paper RF-12

Tuesday:2:00 PM – 2:10 PM

Room: House B

The Utility of the DATA Step Debugger in Logic Errors

Darryl Nousome

There are a few strategies in debugging errors that arise during SAS® programming. Errors in programming can manifest as either syntax or logic errors. Syntax errors tend to be easily identified as they stop the program and generate error messages in the log window. While logic errors will not halt SAS during the DATA step compilation phase, this may result in data that is unintended. A useful tool in examining logic errors is invoking the DATA Step Debugger, which allows viewing of the Program Data Vector (PDV) during the execution of the DATA step. When the DATA Step Debugger is running, there are many options that can examine values of selected variables, suspend statements, display the values of any variables in the PDV, or other commands. This process can aid in identifying segments of code that may contain logic errors. This paper will highlight some techniques that are used by the DATA Step Debugger.

Paper RF-06

Tuesday:2:15 PM - 2:25 PM

Room: House B

Reuse, Don't Reinvent: Extending Model Selection Using Recursive Macro

Anca Tilea and Philip Francis III

The existing SAS® macro: %model_select - used to recursively select the best model based on the R2 selection method - is performing great for specific data and specific variables. The steps of the macro %model_select are as follows: perform a BEST SUBSETS selection based on R2 to get the list of candidate models, calculate the estimates and associated p-values, eliminate the covariates that are never significant, and repeat. This macro does not allow for multilevel variables to be considered for model selection, it does not allow for hierarchical modeling, and it only

considers one model selection method R2. This paper aims to enhance existing SAS® methods, specifically, by allowing various model selection methods (e.g., adjusted-R2, Mallow's Cp, etc.) and the inclusion of categorical variables with multiple levels. The added capabilities will allow the macro to still be user friendly, yet be more robust. Several PROCEDURES, CALL SYMPUT, SCAN, DO-LOOP, IF-THEN-ELSE statements and functions, and various Output Delivery System (ODS) statements are used in expanding the %model_select macro. This paper is intended for the intermediate SAS® user, with intermediate statistical skills and SAS® 9.1 or above.

Paper RF-07

Tuesday:2:30 PM - 2:40 PM

Room: House B

Data Review Information: N-Levels or Cardinality Ratio

Ronald Fehd

This paper reviews the database concept: Cardinality Ratio. The SAS(R) frequency procedure can produce an output data set with a list of the values of a variable. The number of observations of that data set is called N-Levels. The quotient of N-Levels divided by the number-of-observations of the data is the variable's Cardinality Ratio (CR). Its range is in (0--1] Cardinality Ratio provides an important value during data review. Four groups of values are examined.

Paper RF-08

Tuesday:2:45 PM - 2:55 PM

Room: House B

Improve Your ODS Experience with These Essential VBScript Tools

Rose Grandy

While so much is available in version 9.2 and 9.3 of SAS® to make practically perfect submission-ready output using the Output Delivery System (ODS), there are still a few important things that just cannot be done with tools available within SAS. This paper will present some very useful SAS macros that will write and execute VBScript code to: convert ODS RTF files to true DOC files; merge table cells in ODS RTF files; convert RTF or DOC files to PDF; and, convert XML files created using ODS TAGSETS.EXCELXP to true Excel files.

Paper RF-09

Tuesday:3:00 PM - 3:10 PM

Room: House B

An Efficient Approach to Automatically Convert Multiple Text Files (.TXT) to Rich Text Format Files (.RTF) Using SAS

Xingxing Wu and Jyoti Rayamajhi

A large number of legacy and Third-Party-Organization (TPO) provided clinical trial statistical analysis output files stored in SAS Drug Development (SDD) are in .txt (text) format. Because of many advantages of rich text format (RTF) files compared with text files, it's a common task to convert these files from text format to RTF to meet the submissions, regulatory responses, or other requirements. In addition, clinical reviewers often desire to have the RTF outputs since they can write comments on them. This paper provides an efficient and easy-to-use approach to automatically convert multiple text files to RTF files. This approach can directly run in SDD without relying on other tools, such as Microsoft Office VBA, and SDD desktop connection tool. Furthermore, it can also be directly used in other SAS development environment, such as PC SAS. In this paper, an innovative approach is also proposed to resolve the issue that the underline character _ in the text files cannot be displayed after converted to RTF files in some situations. Compared with other approaches, the proposed one has the advantage of robustness. It can be directly applied to any situation without requiring the users to adjust the code to fit into their own situations. This paper is intended for the audiences with some general knowledge about Base SAS and RTF.

Maintaining Formats when Exporting Data from SAS into Microsoft Excel

Nate Derby and Colleen McGahan

Data formats often get lost when exporting from SAS into Excel using common procedures such as PROC EXPORT or the ExcelXP tagset. In this paper we describe some tricks to retain those formats.

Don't Get Blindsided by PROC COMPARE

Joshua Horstman and Roger Muller

"NOTE: No unequal values were found. All values compared are exactly equal." That message is the holy grail for the programmer tasked with independently replicating a production dataset to ensure its correctness. Such a validation effort typically culminates in a call to PROC COMPARE to ascertain whether the production dataset matches the replicated one. It is often assumed that this message means the job is done. Unfortunately, it is not so simple. The unwary programmer may later discover that significant discrepancies slipped through. This paper surveys some common pitfalls in the use of PROC COMPARE and explains how to avoid them.



Discover all that you need ... On your journey to SAS® knowledge

Expert advice in SAS books, documentation, and eBooks.

Visit us at the conference ... We've got you covered.

Save 20% *plus* Free Shipping*

on all books while at the conference!

*Applies to orders placed at the conference only. U.S. addresses only. No minimum purchase required. UPS Ground only.

Have an idea for a SAS book? We want to talk with you.



support.sas.com/bookstore
for additional books and resources.

sas
THE POWER TO KNOW.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. © 2013 SAS Institute Inc. All rights reserved. S107069US.0613

Posters

(on display in the Legislative Foyer)

Presenters will be available to discuss the poster content on Monday at 2:30 and during hours listed on the poster.

PT-01 : Exploring the PROC SQL _METHOD Option

Kirk Paul Lafler, Software Intelligence Corporation

The SQL Procedure contains powerful options for users to take advantage of. This poster illustrates the fully supported _METHOD option as an applications development and tuning tool. Attendees learn how to use this powerful option to better understand and control how a query processes.

PT-02 : Macro to Compute Best Transform Variable for the Model

Nancy Hu, ASA

This study is intended to assist Analysts to generate the best of variables using simple arithmetic operators (square root, log, loglog, exp and rcp) and such as monthly amount paid, daily number of received customer service calls, weekly worked hours on a project, or annual number total sales for a specific product. During a statistical data modeling process, Analysts are often confronted with the task of computing derived variables using the existing available variables. The advantage of this methodology is that the new variables may be more significant than the original ones. This paper gives a new way to compute all the possible variables using a set of math transformation. The codes include many SAS features that are very useful tools for SAS programmers to incorporate in their future codes such as %SYSFUNC, SQL, %INCLUDE, CALL SYMPUT, %MACRO, DICTIONARY.XXXX (where XXXX can be TABLE, COLUMN), SORT, CONTENTS, MERGE, MACRO _NULL_, as well as %DO & %TO & and many more.

PT-03 : You Could Be a SAS® Nerd If . . .

Kirk Paul Lafler, Software Intelligence Corporation

Are you a SAS® nerd? The Wiktionary (a wiki-based Open Content dictionary) definition of nerd is a person who has good technical or scientific skills, but is generally introspective or introverted. Another definition is a person who is intelligent but socially and physically awkward. Obviously there are many other definitions for nerd, many of which are associated with derogatory terms or stereotypes. This presentation intentionally focuses not on the negative descriptions, but on the positive aspects and traits many SAS users possess. So lets see how nerdy you actually are using the mostly unscientific, but fun, Nerd detector.

PT-04 : Trend Reporting Using the MXG® Trend Performance Database

Neal Musitano Jr., U.S. Department of Veterans Affairs

This paper is about my user experience on trend reporting for z/OS server performance using the MXG Trend Performance Database (PDB). In order to implement trend reporting with graphs and charts, the trend pdb must be built first. Thus my user experience of building the trend pdb is included. MXG users build a daily pdb and they use it to produce daily reports for z/OS performance monitoring. However; because of daily tasks or firefighting duties, or deadlines in getting out daily reports or other reasons they sometimes overlook building the trend performance database. Additional reasons for the delaying the trend pdb implementation, include uncertainties of what is involved in building the trend pdb, what to include in it, unfamiliarity of the trend pdb structure and what information to collect for the trends. This paper will address those issues to help users get started with z/OS trend reporting and building the MXG trend performance database.

PT-05 : Customizing a Multi-Cell Graph Created with SAS ODS Graphics Designer

Yanhong Liu, Cincinnati Children's Hospital Medical Center

Combining multiple graphs and/or statistical data tables into one graph is an effective way to compare research data side by side or to present a summarized report of related information together. The SAS® Graph Template Language (GTL) provides the ability to create such multi-cell graphs by using its powerful building-block syntax. In SAS® 9.3, GTL provides a new feature, discrete attribute maps, which enables us to map visual attributes such as shape and color to input data values. You can add this statement block to the GTL code and customize the appearance of each cell of the multi-cell graph, thereby improving the overall graphical visualization, and allowing better interpretation of the graph. To avoid writing GTL code from scratch, we can use the ODS Graphics Designer which is based on Graph Template Language to create the graph, then copy the GTL code generated by the ODS Graphics Designer into the program editor for further customization.

PT-06 : Aligning Parallel Axes in SAS® GTL

Perry Watts, Stakana Analytics

Sometimes it is important to display data using parallel axes that need to be aligned. Unfortunately Graph Template Language (GTL) does not have an `ALIGN=TRUE|FALSE` axis option. Therefore to compensate, GTL options are manipulated so that weights can be displayed in pounds and kilograms, heights in inches and centimeters, frequencies in counts and percents, and temperatures in Celsius and Fahrenheit. The following GTL statements are used in version 9.3 SAS to create the graphs: `SCATTERPLOT`, `HISTOGRAM`, `BARCHART`, and `VECTORPLOT`. Intermediate to advanced SAS programmers with experience using any graphics software package will get the most out of this presentation.

PT-07 : Increase Pattern Detection in SAS® GTL with New Categorical Histograms and Color Coded Asymmetric Violin Plots

Perry Watts, Stakana Analytics

Detecting patterns in graphics output is much easier when numeric data can be grouped categorically. Such is the case with the Body Mass Index and its four classifications: underweight, normal weight, overweight and obese. This presentation goes from conventional histogram to asymmetric violin plot with coverage of the categorical histogram along the way. `HISTOGRAM`, `BANDPLOT` and `LATTICE` statements are described in context. Version 9.3 SAS must be used to replicate the graphs. Intermediate to advanced SAS programmers with experience using any graphics software will get the most out of this presentation.

PT-08 : The DOs and DONTs of PROC REPORT: Building From Introductory Ideas into Professional Results

Daniel Sturgeon, Priority Health

Erica Goodrich, Grand Valley State University/MPI Research

It is important to be able to effectively display results in a way that is both useful and aesthetically pleasing. `PROC REPORT`'s flexibility to procedure reports in SAS teamed with `PROC TEMPLATE` and ODS graphics and outputs can create professional looking reports. However for a user who is not well versed in using SAS or `PROC REPORT` this coding can seem daunting and downright foreign since the syntax does not always match other SAS procedures. In this paper we will address the do's and don'ts of `PROC REPORT` touching on methods, options, and tricks that can be used; while discussing some of the more common mistakes users make that prevent them from creating the best looking reports possible.

PT-09 : Twin Ports Area SAS Users Group: Doing great things on a great lake

Steve Waring, Essentia Institute of Rural Health (EIRH)

Ron Regal, University of Minnesota-Duluth/EIRH

Paul Hitz, EIRH

The Twin Ports Area SAS Users is a newly formed local user group comprised of SAS users from academia, health research, and business in and around the Duluth (MN) Superior (WI) area. Our mission is to promote all things SAS among current users and the SAS-curious, as well as foster teaching, research, and technical development collaborations that take advantage of the many strengths of SAS. We held our first meeting in November 2012, became an officially recognized SAS local user group and launched our website in January, and have already hosted a seminar led by a SAS speaker. We meet bimonthly in a casual forum to discuss and share topics ranging from specific programming issues to better utilization of SAS software. This presentation is for any skill level with the intent of providing the audience with an introduction of our group, highlighting some of the areas of teaching and research our various members are engaged in and our plans for the future.

PT-10 : SAS Enterprise Guide® Implementation Hints and Techniques for Insuring Success With Traditional SAS Programmers

Roger Muller, Data-to-Events.com

Traditional SAS programmers develop SAS code in files and submit it for processing regardless of the operating environment (PC, Unix, etc.). SAS Enterprise Guide (EG) follows this model, but adds some unique additional capabilities. This paper addresses setup, initialization and workflow ideas to smoothen and enhance the transition to the EG graphical user interface centered around a process flow window. Work flow will be addressed in either standalone PC mode, or in conjunction with other servers (local vs. remote processing). Starting with hardware and network capabilities, the paper then moves into a discussion on data location and how that affects work flow. Dual screen systems, split views and internal vs. external SAS code storage will be addressed. Certain features in SAS EG may either be hidden or exposed with advantages to doing either. The number of process flows in an Enterprise Guide Project offers flexibility in constructing the total programming effort. Techniques for submitting developing code for step-by-step processing vs. the submission of entire project files is discussed. The paper will address environmental settings for both EG itself and for the advanced code editor. And lastly, the handiest key in EG, the F4 key, which is used to toggle back-and-forth between the current Process Flow window and the most recent window (program, data set, log, output, etc.) will be repeatedly emphasized. The importance of understanding a strong file backup process relative to the work flow being followed will also be discussed.

Innovation and Networking Area (Congressional Room)

Monday 8:00-11:30 and 1:00-5:00, and Tuesday 8:00-11:30

SAS presenters and developers, SAS products, e-Learning, and SAS publications are just a few things that you will find in the Innovation and Networking room this year. This area is a central location for a wealth of exposure to SAS resources. Take this opportunity to meet people, get a sneak peek at future products, chat with vendors, or browse the books that you've been thinking about purchasing.

Code Doctors	Super Demos
<p>The Code Doctors are in! This unique section provides SAS users the opportunity to bring their problematic SAS code for a one-on-one consultation with experts from the SAS user community. Be sure to bring a hard copy and/or electronic file with your code, processes, and/or logs for the Code Doctor to examine. You are sure to benefit from this personalized learning experience.</p> <p>Doctors: Shrathi Amrathnath Art Carpenter Ben Cochran David Corliss Ron Fehd Peter Flom Jack Fuller Chuck Kincaid Kirk Lafler Rajesh Lal Wensui Liu Russ Lavery Steve Popernack Charlie Shipp Art Tabachneck Patrick Thornton</p>	<p>Short 15-20 minute demos of selected topics from SAS presenters <i>(located in Skyway A)</i></p> <p><i>Monday</i></p> <p>9:00 Introducing SAS® Visual Analytics: Mobile (Martha Hays) 10:00 Introducing SAS Visual Analytics Overview (Meera Venkataramani) 11:00 What's New in SAS® Customer Intelligence (Amy Glassman) 1:00 Executing SQL inside DS2 (Charu Shankar) 2:30 What's New in Foundation SAS for 9.4 (Amy Peters)</p> <p><i>Tuesday</i></p> <p>9:00 Publish with SAS (Shelley Sessoms) 10:30 What's New in 9.4 for SAS Intelligence Platform Administrators (Amy Peters)</p> <p style="text-align: center;">Meet the Presenters</p> <p>Monday 8:00 - 11:30 Gary Ciampa and Wanda Shive 12:30 - 5:00 Jennifer Bjurstrom and Brad Richardson</p> <p>Tuesday 8:00 - 10:30 Yiu-Fai Yung and Tim Hunter 10:30 – 1:00 Gary Ciampa and Amy Glassman</p>

SAS Technical Support helps you to use SAS software and solutions effectively. Our consultants are available to answer your questions on the spot or demonstrate the available Web resources, which give you full access to our knowledge bases, downloadable hot fixes, and online problem reporting.

Enjoy the conference!

Thank you to all of our sponsors for making this conference a huge success!

Platinum Sponsors



Gold Sponsors



Silver Sponsors



MWSUG Board of Directors

George Hurley – President

Plymouth Rock Assurance, NJ
gh@mwsug.org

D.J. Penix – Vice President

Pinnacle Solutions, Inc., Indianapolis, IN
dp@mwsug.org

Ken Schmidt – Treasurer

Truven Health Analytics, Ann Arbor, MI
ks@mwsug.org

Cindy Lee – Sr. Vice President

Eli Lilly & Company, Indianapolis, IN
cl@mwsug.org

Rex Pruitt – Secretary

PREMIER Bankcard LLC, Sioux Falls, SD
rp@mwsug.org

SAS Liaison

Nancy Moser
SAS Institute, Inc.
nancy.moser@sas.com

Local User Groups

SAS Institute maintains a list of local user groups in the Midwest region on the SAS Institute Web site:

<http://www.sas.com/usergroups/namerica/mwsug.hsqli> In-house groups are not included here. If we have missed an active local group, please send a message to an MWSUG board member with the appropriate contact information.

ILLINOIS

WISCONSIN ILLINOIS SAS USERS
Dr. LeRoy Bessler
Le_Roy_Bessler@wi.rr.com
webpages.charter.net/wiilsu

INDIANA

CENTRAL INDIANA SAS USER GROUP
D.J. Penix
dj.penix@psiconsultants.com
www.indysug.org

IOWA

IOWA SAS USERS GROUP
Zhuan (John) Xu
johnxu@1st-consulting.com
iowasasuser.wordpress.com

UNIV. OF IOWA SAS USERS GROUP

Bill Knabe
william-knabe@uiowa.edu

KANSAS

KANSAS CITY AREA SAS USERS GROUP
Bill Jones
kcasug@gmail.com
www.kcasug.org

MICHIGAN

MICHIGAN SAS USERS GROUP
Nancy Brucken
nancy.brucken@i3global.com
www.misug.org

MINNESOTA

TWIN CITIES SAS USER GROUP
Donalee Wanna
donalee.wanna@usbank.com
www.tcasug.org

(Also see North Dakota)

MISSOURI

GATEWAY AREA USERS OF SAS
Jeffrey Crabb
jeff.crabb@citi.com

KANSAS CITY AREA SAS USERS GROUP
(see Kansas)**MID-MISSOURI SAS USERS GROUP**

Dana Schmitz
573-817-8300 x171
dschmitz@primaris.org

NEBRASKA

NEBRASKA SAS USERS GROUP
or John Xu
johnxu@1st-consulting.com

NORTH DAKOTA

RED RIVER VALLEY USERS GROUP
Paul Fisk
Paul.Fisk@ndsu.edu
<http://www.und.nodak.edu/org/rrvsug/>

OHIO

CENTRAL OHIO SAS USER GROUP
Christopher Aultman
caultman@cheryls.com
www.cosug.net

CINCINNATI SAS USERS GROUP

Jeff Ahnrsen
jeff.ahnrsen@us.dunnhumby.com
www.cinsug.org

CLEVELAND SAS USER GROUP

Mary MacDougall
info@clevesug.org
www.clevesug.org

SOUTH DAKOTA

SOUTH DAKOTA SAS USERS GROUP
Toby Eich
teich@premierbankcard.com

(Also see North Dakota)

WISCONSIN

WISCONSIN ILLINOIS SAS USERS
Dr. LeRoy Bessler
Le_Roy_Bessler@wi.rr.com
webpages.charter.net/wiilsu

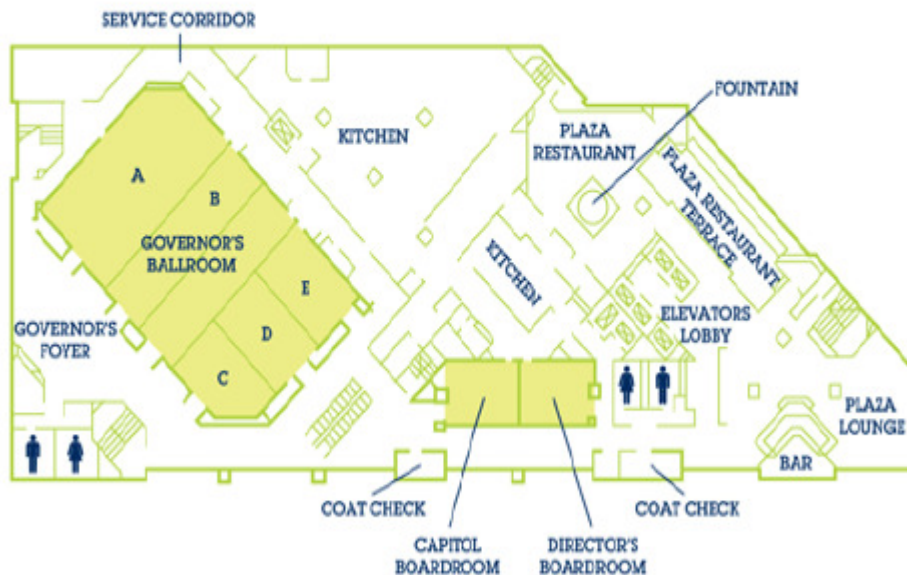
WISCONSIN SAS USERS GROUP

Jennifer First
jfirst@sys-seminar.com
www.sys-seminar.com/EE/wisug

FIRST FLOOR



SECOND FLOOR



MidWest SAS® Users Group (MWSUG) exists to serve users of SAS software in the Midwestern region by providing educational opportunities and resources and by facilitating communication among SAS users and SAS Institute, Inc.

MWSUG is open to all, but primarily serves users in the following states: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin.



We thank you for attending
MWSUG's 24th annual
conference here in Columbus.
We hope you can attend
MWSUG 25, being held October
5-7, 2014 in Chicago, IL



SAS EDUCATION
Grow with us

Learn SAS® from the most experienced training organization in the world.

SAS Education offers multiple learning formats to fit YOUR learning needs.
Choose the style that is right for you—SATISFACTION GUARANTEED.

- Face-to-face classroom training.
- Live Web, instructor-led training.
- Self-paced e-Learning.
- On-site training at your location.
- And more.

Visit the Education booth in the demo area to learn more about our training services,
programs and SAS Certification.



 support.sas.com/training
to learn more and register.


THE POWER TO KNOW.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. © 2013 SAS Institute Inc. All rights reserved. S107969US.0413