



PhilaSUG newsletter



Spring 2019 Meeting Announcement

PhilaSUG Spring 2019 Meeting Wednesday, June 5, 2019

The Philadelphia Area SAS Users Group Spring Meeting will be on Wednesday, June 5, 2019 at 1:00 PM, and will be hosted by Cytel, 2200 Renaissance Boulevard, King of Prussia, PA 19406. A map, is available later in the newsletter.

PhilaSUG meeting and membership registration will begin at 12:15 P.M. The meeting will commence at 1 PM. Membership for the year is \$30. All meetings are free for members, full time students, and employees of our host. We accept cash, check (made out to "PhilaSUG"), or credit/debit card (this charge will show up on your credit/debit card statement as paid to "Michael Davis / PhilaSUG"). Receipts will be available at registration time.

Presenter's abstracts and bios can be found later in the newsletter.



Important – In order to expedite entrance and to help us obtain a better food count we are asking all attendees to please complete the required [Meeting Attendee Sign-up Form](#) by **June 2nd**. It is not necessary to bring proof of

attendance sign up (i.e., printed ticket or via cell phone app) as we will have a list containing the names of everyone that RSVP'd, at the check-in table.

Agenda

We thank our host for providing lunch and break refreshments

12:15-1:00	Registration and lunch
1:00-1:10	Opening Remarks
1:10-1:30	Jonas Bilenas – Formats for Lags and Leads
1:30-2:00	Vincent Rabatin – Enforcing Data Integrity with SAS Audit Trails
2:00-2:30	Mark Keintz – Why and How to Use Condensed Indexes for Sorted SAS® Datasets
2:30-2:50	Break, posters, and refreshments
2:50-3:40	Frank DiIorio – Dictionary Tables and Views
3:40-4:10	Sujaykumar Bothisathuvan – SDTM Metadata and Define.xml – An Overview
4:10-4:30	Kathy Zhai – Throw Away the Key: Blockchain-ed Healthcare Data
4:30-4:45	Open Forum, and Raffle Prizes

Posters

Santosh Kothamasu – SAS Add-in for Excel: Overview

The presenters and the PhilaSUG Executive Committee will adjourn for dinner at a nearby restaurant when the meeting concludes. You are invited to join us. The location will be announced at the meeting.

About Our Host



2200 Renaissance Boulevard,
King of Prussia, PA 19406

Cytel is shaping the future of drug development. As the world's largest independent clinical biostatistics research and development organization, Cytel helps leading pharmaceutical, biotech and medical device (continued on next column)

companies improve clinical success rates via optimal study design, effective data management, and accurate statistical analysis. Cytel provides both software solutions for the design and analysis of clinical trials, including industry standards East®, StatXact® and LogXact®, as well as data focused clinical research services. With operations across North America, Europe, and India, Cytel employs 900 professionals, with strong talent in biostatistics, programming, and data management.

<http://www.cytel.com>

Click on map below for interactive directions by Google.



PhilaSUG Executive Committee

Michael Davis, President
Kajal Tahiliani, Secretary, and Treasurer
George Laskaris, Membership Chairman
Robert Schechter, Newsletter Editor, and Web Master
Jonas V. Bilenas Max Cherny
Barry Cohen Diane Foose
David Horvath Mark Keintz
Amos Odeleye Steve Rhoades
Haibin Shu

Host Sites Wanted

We continuously seek host sites for future PhilaSUG meetings. There is not a lot of work involved, and it is a great way to put your company on the local SAS map. We need your help with this. If your company would like to host a meeting, within reasonable geographic proximity to Philadelphia, PhilaSUG would be grateful if you would contact Michael Davis at michael.davis@alumni.duke.edu.

E-mail Announcements



PhilaSUG-L is a low volume, announcement-only e-mail notification service provided free of charge to all members who wish to subscribe. In order to sign up for this service, you need only send a blank e-mail message to PhilaSUG-L-subscribe@onelist.com. Note that you can subscribe as many times with as many different e-mail addresses as you wish to have the e-mail sent to; e.g., home and office.

PhilaSUG Web Site

Our site on the World Wide Web always contains the latest information concerning upcoming meetings, SAS training and seminars, links to SAS related hot topics, and local SAS job opportunities.



Visit us regularly at: <http://www.PhilaSUG.org>

Presenters Wanted



The PhilaSUG Executive Committee invites presentation abstracts. If you have thought presented before, PhilaSUG is an excellent venue for first-timers. Presentations can range from ten minute coders corner tips to 50 minutes talks. Both new and previously presented topics are welcome. If this is of interest to you, PhilaSUG invites you to submit electronic abstracts and bio for either a poster or oral presentation by using our online submission form. Check our home page for the Call for Papers.

A Thank You



The PhilaSUG Executive Committee wishes to thank Nand Kishore Rawat, Ph.D. at Cytel for arranging the hosting of this meeting. In addition, we wish to thank two members of our Executive Committee, Haibin Shu for speaker/program coordination and Kajal Tahiliani for site coordination.

Paper Abstracts

Enforcing Data Integrity with SAS Audit Trails

Vincent Rabatin, Rabatin Software Services

SAS Audit Trails are a type of SAS dataset that can be applied to a primary dataset in order to track additions, deletions, and updates as well as failed attempts to perform these actions on the primary dataset. This paper demonstrates how to apply audit trails to clinical trial datasets in order to ensure that the dataset meets its expected requirements. Macros supporting the process are driven by metadata derived from the project enhanced with constraints applied to the variables which may include ranges, missing data checks, adherence to lists of values, correspondence to data in another dataset, consistency of variables within the same record, and primary key constraints. Alternative ways of implementing the audit trails using Data Step programming or SQL are presented to conform to the preference of the programmer..

Vincent Rabatin, the founder and CEO of Rabatin Software Services LLC, has worked in the pharmaceutical industry since 1998 in clinical trials, clinical systems development, outcomes research and epidemiology. Prior to moving to the Pharmaceutical industry Vincent had a 24-year career at the Pennsylvania State University involving health related issues including the physiology of human performance, Drug Utilization Review, HIV/AIDS, Elderly use of health care, Pharmacoeconomics, and Health Insurance Coverage. His experience with the SAS programming language spans 35+ years. Recently Vincent was the Head of Statistical Programming at Inference Inc.

Using User Defined FORMATS and the INTNX Date Function to Extract LAGS and LEADS Using SAS Date Variables as VALUES for the Look-Up

Jonas V. Bilenas, Banking Industry Consultant

In many time series models one often uses independent variables that are linked to date and time keys. One can develop user defined table look-ups using PROC FORMAT to generate a table look-up to capture the LAGS and/or LEADS of various time series variables. We will look at an example of generating permanent user FORMAT CATALOGS that can be used to extract macroeconomic variables based on DATE and using the INTNX function to quickly calculate LAGS and or LEADS for model development and model forecasting into the future.

Jonas Bilenas has been using SAS for more than 34 years in the field of Consumer Risk Management. He has presented at many Local, Regional, and International User Groups and is currently active in the PhilaSUG Executive Committee. He has also authored the SAS Press Book titled, "The POWER of PROC FORMAT."



Dictionary Tables and Views: Essential Tools for Serious Applications

Frank DiIorio, CodeCrafters, Inc., Chapel Hill NC

Dictionary tables were introduced to the SAS System in the early 1990's, in Version 6.07. Laden with information that is often difficult, and sometimes impossible, to get through other means, they still appear to be on the outside of many programmers' Bag of Tricks. This is both perplexing and unfortunate for as we will see in this paper, once their content and organization is understood, they are readily adapted for a range of applications that, to use an old saw, "are only limited by your imagination."

This presentation describes dictionary tables and their associated SASHELP library views. It:

- presents **scenarios** that show how they can be used
- gives high-level **descriptions** of some of the more important (a relative term, to be sure) tables
- identifies features of **SQL** and the **macro language** that are commonly used when writing programs that effectively use the tables
- shows **examples** of the tables' use, emphasizing the use of SQL and the macro language interface

The reader should come away from the discussion with an understanding of the tables as well as with a checklist of SQL skills that are required to use the tables most effectively.

A SAS programmer since 1975, Frank DiIorio is President of CodeCrafters, Inc. and the author of "SAS Applications Programming: A Gentle Introduction" and "Quick Start to Data Analysis with SAS." A frequent presenter at local and regional SAS user groups, he is past President of the SouthEast SAS Users Group, and co-chaired its 1994 and 1996 conferences. He is also active in several local SAS user groups and was a cofounder of the Research Triangle CDISC Users Group.

Why and How to Use Condensed Indexes for Sorted SAS® Datasets

Mark Keintz, Wharton Research Data Services

One of the primary benefits of the standard SAS index is improved speed when retrieving a subset from a SAS data set, especially if the data are not sorted. This arises from the use of an index file, sorted by the index variable (or multiple variables for compound indexes), with a record id pointing to each of the records in the data set. However, for data sets sorted or grouped by the index variable(s), it is frequently possible to dramatically improve performance by reducing the number entries in the index file from NOBS (the number of records in the original dataset) to NIX (the number of unique index values).

If the relative cardinality of the index (i.e. the ratio of NIX/NOBS) is sufficiently small, the result is an index file (in the form of a SAS dataset) that minimizes disk space requirements, reduces CPU time and elapsed time by an order of magnitude.

This presentation will demonstrate, with sample DATA steps, how to create and use such an index, along with performance results compared to SAS indices. Other advantages such as precise prediction of the number of retrieved records will be discussed.

Mark Keintz's current programming interests are in financial research applications and large data sets. He has previously served on NIH Small Business Innovation Research review panels and directed computing support for demography and sociology research.

Mark has been using SAS® since it was documented in a single book.



Throw Away the Key: Blockchain-ed Healthcare Data

Kathy Zhai, GlaxoSmithKline

After following SAS blogs and other social media outlets that correspond to the latest pharmaceutical trends, the words "blockchain" and "bitcoin" have been prevalent. How are any of these words associated with healthcare and patient data? Like other programmers working in the pharma industry, my curiosity grew beyond just procs and data steps. Many can agree that the credibility of clinical outputs can be undermined by a plethora of common issues including incomplete, missing, or inaccurate data. After multiple layers of data manipulation, how is it certain that what is being submitted to publications is an undistorted version of the benefits and risks of these drugs? This paper will give the audience a glimpse of how blockchain technology, whose implementation is cryptographically validated by a network, has enough potential and momentum to emerge into the healthcare industry and stick around for quite some time.

Kathy is a professional data wrangler, expert negotiator, avid problem solver -- or plainly, a SAS programmer for GlaxoSmithKline. She has been in the industry for 8 years in both CROs and Pharma. She is an active attendee of PharmaSUG, PhUSE and pilates.

SDTM Metadata and Define.xml – An Overview

Sujaykumar Bothisathuvan, Cytel

This topic speaks about SDTM metadata that is required to implement, which creates an SDTM database. We will look a little into how Microsoft Excel files are used to store the SDTM metadata that the SAS code will leverage to create the SDTM datasets. We will also look into how to build Define.xml in this presentation.

Sujaykumar has over 12 years' experience utilizing SAS to analyze clinical trial data in the CRO industries. Extensive experience implementing CDISC standards in creating SDTM/ADAM datasets and TLF's across several therapeutic areas, including Circulatory, Nervous System, Musculoskeletal, Oncology, Vaccines, Infectious Diseases, Diabetes, etc.



Poster Abstract

SAS Add-in for Excel: Overview

Santosh Kothamasu, TD Bank

SAS Add-in for MS Office is a powerful utility. Poster illustrates this functionality for Excel, Combine SAS and Excel for advanced report automation!

Santosh Kothamasu is an enthusiastic SAS user with 15 plus years experince in the banking industry.

Future Meetings and SAS Events



**PharmaSUG 2019:
June 16-19, Philadelphia, PA**



Next PhilaSUG Meeting

PhilaSUG Fall 2019 Meeting

Host: West Chester University
Department of
Mathematics and Applied
Statistics Program

Where: Graduate Center
West Chester University
1160 McDermott Drive
West Chester PA 19380

When: Tuesday October 22

