SANDS Newsletter

AUGUST 15, 2012

3:30 - 4:45
WORKSHOP
Testing Treatment Effects with a Covariate and a Covariate-Dependent Stratification Factor
Joey Lin
SDSU

5:00 - 5:45
LIGHT DINNER / NETWORKING
5:45 - 6:00
WELCOME / SANDS BUSINESS

6:00 - 6:50
PRESENTATION #1
Sample Size Calculation and Power Analysis for Design of Experiments Using Proc Glmpower
Joey Lin
SDSU

7:00 - 7:50
PRESENTATION #2
Developing and Delivering an Effective Online SAS® Programming Course
Justina Flavin
UCSD Extension

8:00 - 8:15
STUMP THE PROGRAMMER AND CODERS' CORNER
8:15 - 8:30
DOOR PRIZES AND GIVEAWAYS

Pfizer Global R&D, La Jolla Labs
10770 Science Center Drive
Building CB2, Room 1110*
San Diego, CA 92121

From the President

While you are watching the London 2012 Summer Olympics, please do not forget our summer SANDS meeting. In the upcoming August meeting, Dr. Joey Lin from San Diego State University will give a workshop on testing treatment effects for clinical trials in the afternoon. He will give another talk on sample size calculation and power analysis using PROC GLMPOWER in the evening. The other speaker for the evening meeting is Justina Flavin. She will show you how to develop and deliver an online SAS programming course.

The SANDS Executive Committee is working on our third one-day meeting to be held on December 7, 2012. We have had two successful one-day meetings in 2005 and 2008, so we feel that it is time to have another one-day meeting this year. We plan to have six sections: (1) Programming: Foundations and Fundamentals, (2) Programming: Beyond the Basics, (3) Pharma and Health Care Providers, (4) Reporting and Information Visualization, (5) Statistics and Data Analysis, and (6) Predictive Analytics and Data Mining. We are looking for volunteers from our members to support this event.

The WUSS (Western Users of SAS Software) annual conference is around the corner! It will be held in Long Beach, California, on September 5-7 at the Hyatt Regency. Check it out at http://www.wuss.org/. Our EC member Scott Leslie will be the academic program chair for this conference.

PharmaSUG (Pharmaceutical SAS Users Group) is planning a one-day event to take place in the San Diego area around September 2013. This is to take place prior to the San Diego PharmaSUG conference scheduled to take place in June 2014. The Executive Board of PharmaSUG is looking for a venue in the San Diego area that can accommodate about 100 people. Also, they would like a venue that would charge little or nothing to use. If you know of any such venues, please let me know, and we would really appreciate your help.

SANDS is run by volunteers from the local SAS community. Last but not least, I would like to take this chance to thank all of the SANDS EC members for their hard work!
— **Wei Cheng**

Please fill out the RSVP form for the workshop and/or meeting by Monday, Aug 13th, at http://www.sandsug.org/meetings.html.

Workshop:
Testing Treatment Effects with a Covariate and a Covariate-Dependent Stratification Factor
Joey Lin, SDSU

Abstract

In the analysis of clinical trial data, it is well accepted that including baseline measurements of the outcome variables may improve the detection of treatment effect. Models using outcome variable or its change from baseline (or change score model) version as a response variable and the baseline measurements as a covariate for testing treatment effect are usually employed. To ensure that similarly represented samples regarding baseline values in both treatment groups are selected for a clinical trial, stratification based on a preselected cut point(s) of the baseline is used to define strata and randomization is then implemented within each stratum. The ICH and CHMP guidance documents have stated that the stratification factor shall be included in the analysis models.

In this workshop, we will review models such as analysis of covariance (ANCOVA) and change score models. We will also discuss issues that may cause the decisions made by these two approaches to be different. This is so-called Lord's paradox. We will also compare the efficiency of the ANCOVA and change score models under a variety of situations. This workshop is organized from an applied approach. We will not cover technical details but focus on how we can use these models in clinical trials and issues of which we have to be aware.

Presentation #1:
Sample Size Calculation and Power Analysis for Design of Experiments Using Proc Glmpower
Joey Lin, SDSU

Abstract

"How many samples do I need?" When a statistician or a quantitative analyst gets involved in a project, this is always one of the first questions to be asked. SAS has developed several tools to address this need. For example, PSS (power and sample size), Proc Power, and Proc Glmpower are all available for users to use in various designs. Introductions in PSS and Proc Power have been presented in many SAS related conferences. However, Proc Glmpower is relatively unknown compared to the other two tools. In this paper, we introduce how one can easily understand the syntax of Proc Glmpower for sample size calculation or power analysis related to the design of experiments. In conducting an experimental design, Proc Glm allows one to test for pre-planned contrasts. Similarly, Proc Glmpower can be used to calculate either the sample size or the power for any contrast one may have. The related syntax is discussed in this paper. In addition, how to take advantage of the powerful ODS features in Proc Glmpower is also addressed.

Biography

Dr. Joey Lin is a professor in statistics at San Diego State University. He is also a co-director of the Statistical Consulting Center at SDSU. He has been using SAS for more than 20 years. He teaches an introductory SAS course at SDSU regularly and has inspired students to become SAS programmers. He also uses SAS as his primary computational tool for his research.

Presentation #2:
Developing and Delivering an Effective Online SAS® Programming Course
Justina Flavin, UCSD Extension

Abstract

Teaching online courses is very different from teaching in the classroom setting. Developing and delivering an effective online class entails more than just transferring traditional course materials into written documents and posting them in a course shell. The licensing issue presents an additional challenge in teaching a course that uses SAS® software. I will discuss how I converted a traditional hands-on introductory SAS programming class into an online course that uses SAS OnDemand for Academics and present some ideas to promote successful learning and knowledge transfer when teaching online.

Biography

Justina Flavin is a consultant providing statistical programming services to clients in the biotechnology, pharmaceutical, and medical device industry. She is an instructor for courses in SAS programming and statistics at the University of California, San Diego - Extension Division. At Extension, she is also a curriculum developer and advisor for the Biostatistics certificate and advisor for the Data Mining certificate. Justina served as Conference Chair of PharmaSUG '99 and has taught seminars at PharmaSUG and WUSS.

SANDS Membership

Membership in the San Diego SAS Users Group, Inc. (SANDS) provides an opportunity for SAS users to meet and keep updated on SAS issues. Seminars and presentations allow SAS users to gain knowledge with little or no cost. It is also an opportunity for consultants and prospective employees to meet possible employers, and companies to show their SAS-related products and host a meeting. We welcome those interested in giving presentations and writing articles for the newsletter. Presently, there is no charge to attend our meetings or to receive our newsletter; small fees may be asked for some special functions or workshops.

The SANDS membership list is not given out to others and is used only to disseminate SANDS related business, such as newsletters or special announcements. Those wishing to sell their products or services have the opportunity to advertise in the newsletter at reasonable rates.

To become a member, please email the following to Cathy Liu: Name, Title, Organization, Telephone, E-mail, Fax, Address, Work or Home indicator, Computer Platform, SAS version, and SAS related areas of interest.

Tips for Using a Data Warehouse
Curtis Smith, Defense Contract Audit Agency

In our last episode, we examined methods for reducing disk space used when processing a SAS job. This time, we will explore a way to save disk by reducing the size of the SAS data sets. There are many ways to reduce the size of SAS data sets that you place in your temporary work space and in permanent storage. One way is to let SAS use compression algorithms to compress SAS data sets. This is a process of reducing the amount of space needed to store a SAS data set - it does not affect the data stored within that SAS data set.

Using the COMPRESS= system, any SAS data set created on disk will be compressed. Using the COMPRESS= data set option, the specific SAS data set will be compressed. SAS data set compression can greatly reduce the size of SAS data sets. To use the COMPRESS= system or data set option, set the option to either "YES" or "BINARY." (In newer versions of SAS, "CHAR" can be used as an alternative to "YES" with the same result.) The COMPRESS=YES value uses an algorithm that works better with SAS data sets that are primarily comprised of character variables. On the other hand, COMPRESS=BINARY uses a different algorithm that works better with SAS data sets that are primarily comprised of many variables including many numeric variables.

My experience has been that COMPRESS=YES reduces the size of the SAS data set by about 50 percent. An option to use with COMPRESS= is REUSE=. Specifying this option allows SAS to reuse space within the compressed SAS data set that has been freed by deleted observations. Otherwise, SAS cannot reclaim the space made available by deleted observations. Consider the following examples.

```
/*use the data step compress option*/
data temp.file2(compress=binary reuse=yes);
  set work.file1;
  where rec_type='1';
run;

/*use the system compress option*/
options compress=yes reuse=yes;
data work.file3;
  set work.file1;
  where rec_type='1';
run;
```

After running the DATA step with the COMPRESS= set to "YES" or "BINARY", you will see a message in your SAS log that looks something like this:

```
NOTE: The data set WORK.FILE3 has 117658 observations and 20 variables.
NOTE: Compressing data set WORK.FILE3 decreased size by 31.70 percent. Compressed
is 935 pages; un-compressed would require 1369 pages.
```

I thought you might want to see some benchmark results. I took four SAS data sets with differing numbers of observations and variables. I started with uncompressed SAS data sets, then compressed with normal compression, then compressed with binary compression. File1 and File3 have 20 variables, 5 of which are numeric; and File2 and File4 have 29 variables, 10 of which are numeric.

File	Obs	Char	Num	Compress YES	Reduction BINARY
File1	120,641	15	5	31.91%	23.72%
File2	183,476	19	10	52.10%	49.22%
File3	1,542,257	15	5	31.38%	21.02%
File4	6,976,838	19	10	48.81%	46.61%

In my examples, COMPRESS=YES always produced better results. However, the files with more variables and more numeric variables (File2 and File4) got almost the same benefit using COMPRESS=BINARY. I recommend you try both methods with your files to determine which provides the better performance. In some cases, compressing a SAS data set will result in a file that is larger in storage size than the uncompressed original. When this happens, SAS will warn you with a message. Since version 8, SAS will not compress the SAS data set when the result would be a larger file size.

Thanks for reading.

Curtis Smith, IT Technical Specialist

Stump the Programmer #61

Date Comparisons – When the Same Dates are Unequal

Art Carpenter, CA Occidental Consultants

SAS dates are stored as numeric values that measure the number of days that have elapsed since the beginning of time (January 1, 1960). The following IF statement compares two variables (TODAY and ALSOTODAY) which contain valid SAS date values.

```
if today ne alsotoday then
  put 'Today ' today date9. ' is not
equal to ' alsotoday date9.;
```

The LOG shows:

```
Today 11JUL2012 is not equal to 11JUL2012
```

Can you think of two scenarios that would cause these two dates to display as equal and yet be found to be unequal in the comparison?

Kirk's Korner

Kirk Paul Lafler, Software Intelligence Corporation

Changing the SAS 9.3 Output Destination with ODSDEST=

As was presented in an earlier SAS Tip I wrote, "New Output Default in SAS 9.3 – HTML", beginning in SAS 9.3, the new HTMLBlue style output will be the default destination in the SAS windowing environment for the Microsoft Windows and UNIX operating systems. For users running SAS in batch mode, versions earlier than SAS 9.3, and/or operating systems other than Windows and UNIX, the Listing destination will continue to be the default destination. For SAS 9.3 users who prefer, or who need, to continue using the 'Listing' destination for your output, you're in luck. SAS Institute provides a system option, ODSDEST=, to change this default. For Microsoft Windows and UNIX operating system users, the ODSDEST= system option can be specified at SAS invocation or in the SAS configuration file. The available ODSDEST= options are:

1. **HTML** – Specifies that TAGSET.HTML4 be used as the output destination.

2. **LISTING** – Specifies that monospace output be used as the output destination.

3. **AUTO** – Specifies that SAS registry settings be used to determine the default output destination.

The following example illustrates the statement that is used to invoke SAS specifying the Listing destination should be the default output destination.

Example:
"C:\Program Files\SASHome\SASFoundation\9.3\sas.exe" **-ODSDEST Listing**

About The Author

Kirk is senior consultant and founder of Software Intelligence Corporation and has been using SAS since 1979. He is a SAS Certified Professional, provider of IT consulting services, trainer to SAS users around the world, and sasCommunity.org Advisory Board member emeritus. As the author of four books including PROC SQL: Beyond the Basics Using SAS, Kirk has written more than five hundred papers and articles, been an Invited speaker and trainer at three hundred-plus SAS International, regional, local, and special-interest user group conferences and meetings, and is the recipient of 19 "Best" contributed paper awards. His popular SAS Tips column, "Kirk's Korner of Quick and Simple Tips" and SASword Puzzles, appears regularly in several SAS User Group newsletters and Web sites, and in SAScommunity.org.

Contact Information
Kirk Paul Lafler,
Senior Consultant, Application Developer, Trainer and Author
Software Intelligence Corporation
E-mail: KirkLafler@cs.com
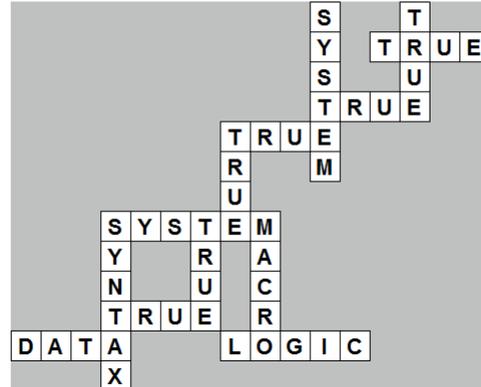LinkedIn: http://www.linkedin.com/in/KirkPaulLafler
Twitter: @sasNerd

SANDS Sponsorship

We would like to thank Pfizer for having once again generously provided our meeting room. This month, SANDS is sponsoring our meeting dinner and activities. It is our sponsors who make it possible for us to hold our meetings free of charge, so at the event, please do let sponsor representatives know how much you appreciate their support. If your company would like to sponsor SANDS activities, please contact the SANDS sponsorship coordinator, Tyler Smith.

Exploring Errors in the SAS® Software

Copyright © 2012 by Kirk Paul Lafler

Solution:

Upcoming Conferences
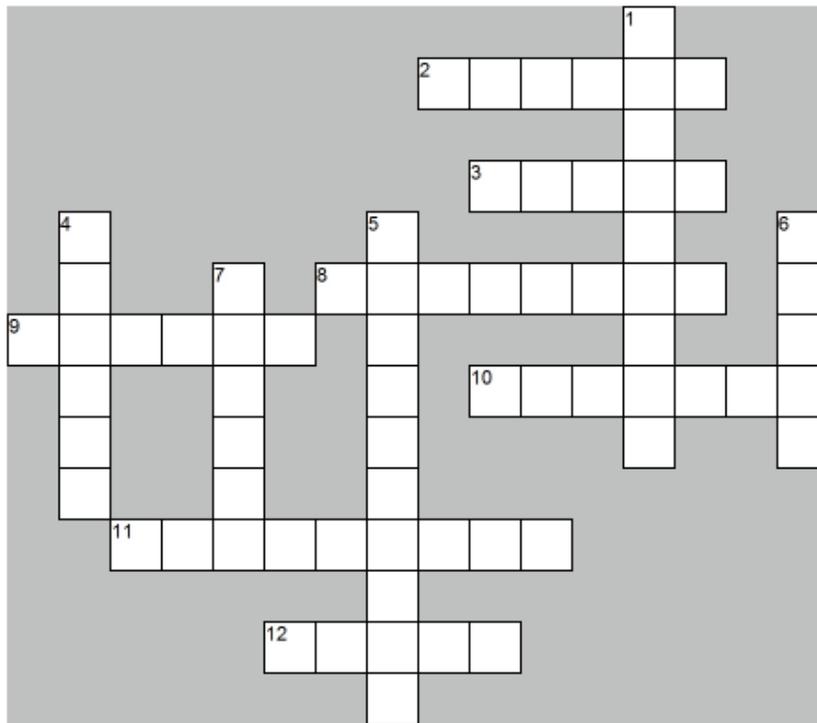
September 5-7, 2012
WUSS
Long Beach, CA

April 28-May 1, 2013
SAS Global Forum
San Francisco, CA

Optimizing SAS® System Performance (Part 1 of 2)
Copyright © 2012 by Kirk Paul Lafler

This puzzle was made by Kirk Paul Lafler, Software Intelligence Corporation using Crossword Weaver(TM)

ACROSS

2 By increasing this SAS system option's value, an application's performance can be improved by allowing SAS to read more data with fewer passes.
3 This data set option is used to remove unwanted variables.
8 This system option can help improve I/O by reducing the number of disk accesses by setting the permanent page size for the data set.
9 This statement is used to reduce the size of a variable.
10 By increasing this system option, the total amount of memory available to a SAS session can be increased preventing out of memory conditions.
11 This system option controls the amount of memory available for sorting data.
12 This data set option is used to retain desired variables.

DOWN

1 This data set option can help improve I/O by storing a data set in compressed format.
4 This data set option is used to subset observations.
5 When specified, this SAS system option writes all performance statistics for your computer to the SAS Log.
6 This optional file can be created for a SAS data set to provide direct access to desired observations.
7 When specified, this SAS system option writes to the SAS Log the amount of time it took for SAS to complete a DATA step or procedure task.

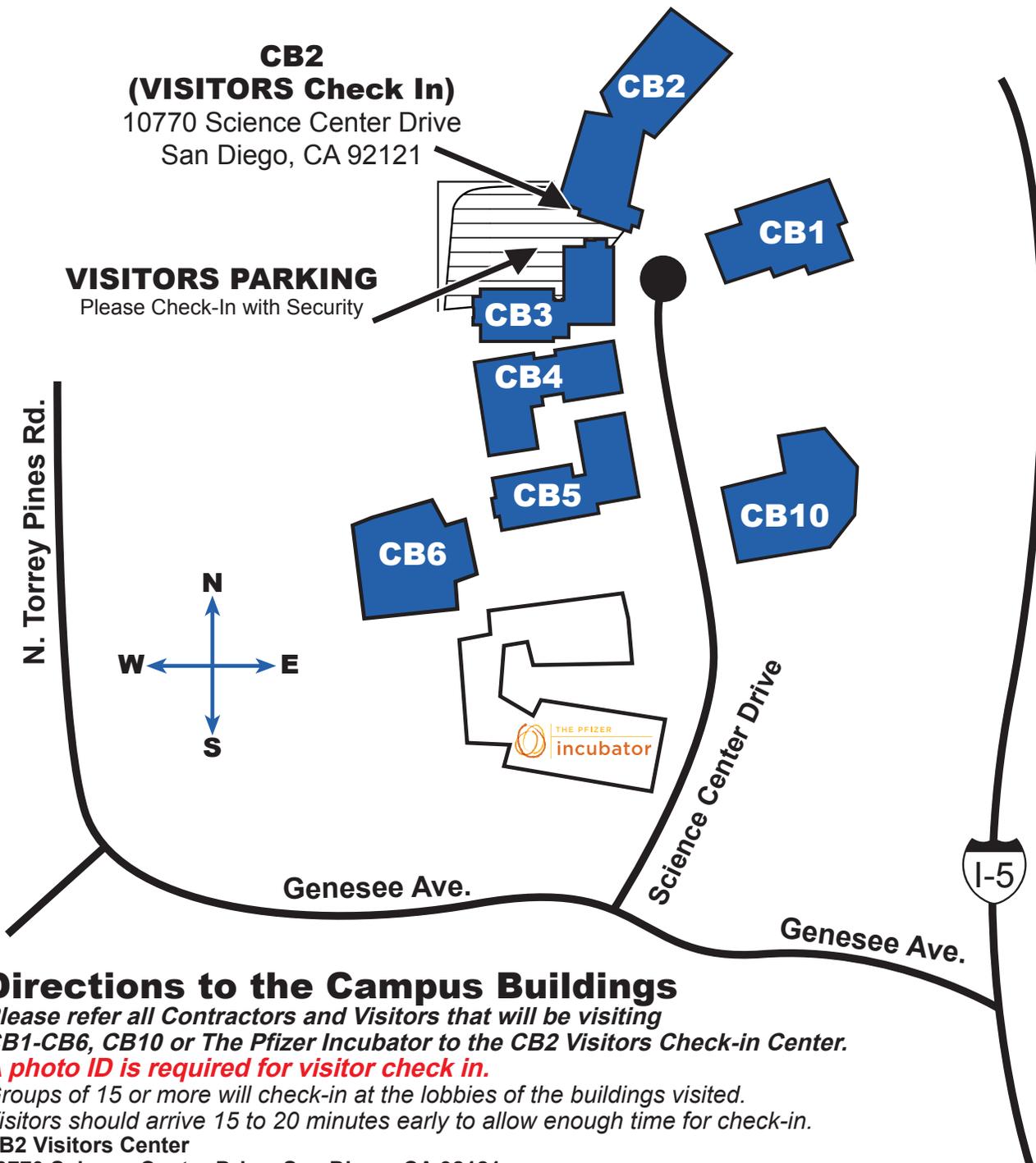Comments and suggestions can be sent to:

Kirk Paul Lafler
Software Intelligence Corporation
E-mail: KirkLafler@cs.com
LinkedIn: http://www.linkedin.com/in/KirkPaulLafler
Twitter: @sasNerd

Map to Pfizer La Jolla Campus

CB2
(VISITORS Check In)
10770 Science Center Drive
San Diego, CA 92121

CB2

CB1

VISITORS PARKING
Please Check-In with Security

CB3

CB4

CB5

CB6

CB10

N
W   E
S

THE PFIZER
incubator

N. Torrey Pines Rd.

Science Center Drive

I-5

Genesee Ave.

Genesee Ave.

Directions to the Campus Buildings

Please refer all Contractors and Visitors that will be visiting
CB1-CB6, CB10 or The Pfizer Incubator to the CB2 Visitors Check-in Center.
A photo ID is required for visitor check in.
Groups of 15 or more will check-in at the lobbies of the buildings visited.
Visitors should arrive 15 to 20 minutes early to allow enough time for check-in.
CB2 Visitors Center
10770 Science Center Drive, San Diego, CA 92121
From San Diego International Airport.
Follow signs to I-5 North.
Proceed on I-5 North for approx. 10 miles.
Exit I-5 at Genesee Ave. and make a left at the stop light.
Make a right at the second stop light which will be Science Center Drive.
CB2 Visitors Center is located at the end of the cul-de-sac.
Visitor Parking is to the left. **Please follow Visitors Signs.**

Pfizer

CSLJ 194423 3/08

SANDS EXECUTIVE COMMITTEE

ELECTED OFFICERS:

President	Wei Cheng	760-603-3807	wcheng@isisph.com
Vice-President	Archie Medrano	858-550-0466	amedrano@ucsd.edu
Secretary/Treasurer	Song Lin	858-677-9637	Song.Lin@nielsen.com

Advertisements	Tyler Smith	619-553-7593	tsmith@nu.edu
	Vivian Huang	858-587-5873	Vivian.Huang@biotheranostics.com

Type	Dimensions (w x h)	Cost
Business Card	3.5" x 2"	$25
Quarter Page	3.5" x 4.5"	$50
Half Page	7" x 4.5"	$125
Full Page	7" x 9"	$225

Note: Prices may change without notice. Additional fees may be charged if modifications are needed. For verification of current prices and/or to send ads in PDF format, please contact Tyler Smith and/or Vivian Huang.

Catering	Ning Ding		ndingbox@yahoo.com
Facilities Coordinator	Ren-Yu Tzeng	858-638-3979	ren-yu.tzeng@pfizer.com
Job Bank	Robert Hull	760-268-8003	RHull@synteract.com
Membership Coordinator	Cathy Liu	760-476-8800	cliu@synteract.com
SAS Institute Liaison	Andrea Littleton		Andy.Littleton@sas.com
Speakers/Program Coordinator	Scott Leslie	858-790-6685	scott.leslie@medimpact.com
Sponsorship	Tyler Smith	619-553-7593	tsmith@nu.edu
	Vivian Huang	858-587-5873	Vivian.Huang@biotheranostics.com

Newsletter Editor/Webmaster	Archie Medrano	858-550-0466	amedrano@ucsd.edu

The SANDS newsletter is published quarterly and distributed to members by e-mail. The SANDS EC team appreciates the efforts of all contributors, without whom this newsletter would not be possible. We are always looking for volunteers to contribute original content to the quarterly newsletter. If you have ideas or suggestions for the newsletter, please share them with us. We would love to hear them!

Newsletter text not otherwise indicated was written or edited by the Editor. Please send any questions or comments to the Editor.

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.

Websites:

www.sandsug.org/index.html

sasCommunity.org

LinkedIn