

SANDS Newsletter

AUGUST 21, 2013

3:00 - 4:00

WORKSHOP

A Tour of New Features in
SAS® Enterprise Guide®
4.3, 5.1 and 6.1
Shannon Moore

4:00 - 5:00

PRESENTATION #1

Categorical Data Analysis:
Because There Really Is No
Fate Worse Than Death
AnnMaria De Mars

5:15 - 6:00

**LIGHT DINNER /
NETWORKING**

6:00 - 6:15

**WELCOME /
SANDS BUSINESS**

6:15 - 7:00

PRESENTATION #2

Factor Analysis Is Your
Friend
AnnMaria De Mars

7:00 - 7:30

PRESENTATION #3

Strategies and Techniques
for Getting the Most Out of
Your Antivirus Software for
SAS® Users
Ryan Paul Lafler

7:30 - 7:45

**STUMP THE
PROGRAMMER AND
CODERS' CORNER**

7:45 - 8:00

**DOOR PRIZES AND
GIVEAWAYS**

From the President

I hope you are all enjoying your summer with family and friends! There are lots of exciting events in San Diego during this time of year: [San Diego County Fair](#), [Del Mar Horse Racing](#), [Comic-Con International](#), [Mainly Mozart](#) concerts, [WTA Southern California Open](#), etc. I think some of you must have enjoyed at least one of these events.

We will have quite a few SAS-related events in the coming months too. PharmaSUG will have their first single day event in La Jolla in September, SDASA is planning a one-day mini-conference in October to celebrate the international year of statistics, and WUSS will have their 2013 annual conference in Las Vegas in November. And, of course, SANDS will have our third quarterly meeting of the year on August 21st! Hopefully, you can attend most, if not all, of these events.

In the upcoming August meeting, Dr. AnnMaria DeMars from the Julia Group will give a presentation on categorical data analysis in the afternoon and another presentation on factor analysis in the evening meeting. Shannon Moore from SAS Institute will give you a tour of new features in SAS Enterprise Guide as the afternoon workshop. We will have another special guest speaker for the evening meeting; Ryan Paul Lafler from Valhalla High School will give a short presentation on getting most out of your antivirus software while you are doing your SAS programming.

I would like to thank our sponsors for this meeting: WUSS Executive Committee, SAS Institute, and Pfizer. SANDS meetings are held with no cost to our members because of their generous sponsorships. If your company or other companies can be sponsors for our SANDS meetings or would like to advertise in our quarterly newsletter, please let us know. If you are interested in serving your local user group, please contact me. And, as always, I would like to thank all of the EC members for their hard work!

See you at the meeting!

— **Wei Cheng**

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

Please fill out the RSVP form for the workshop and/or meeting by Friday, Aug 16th, at
<http://www.sandsug.org/meetings.html>.

Workshop: A Tour of New Features in SAS® Enterprise Guide® 4.3, 5.1 and 6.1 Shannon Moore, SAS Institute

Abstract

As a SAS Enterprise Guide user or administrator, you have probably recently upgraded to 4.3 or 5.1 or are looking forward to going to these versions or the newest version, 6.1. In this paper, we will take you on a tour of new features available to you in each of these releases so that you can be more productive with your current or upcoming version, as well as obtain a preview of what's coming in the future. We will cover the differences between the various versions of SAS Enterprise Guide clients and their compatibility with the corresponding SAS server versions, and explain how to upgrade from your current version with the goal of giving you better decision points for planning your client/server upgrade strategy.

Biography

Shannon Moore has been with SAS since 1996 and is a charter member of the SAS' Customer Loyalty organization. His areas of expertise include Enterprise Business Intelligence, Enterprise Guide, SAS/GRAPH, SAS' mapping capabilities, Data Integration and ODS/ODS Statistical Graphics. He has experience in many industries including Health Insurance, Pharmaceuticals, Oil and Gas, Retail, Manufacturing, Financial Services, Telecommunications, State, Local and Federal Government. He is the recipient of a SAS Americas Sales Innovation Award and a multi-year member of the Century Club. Shannon graduated from the University of Colorado, Boulder, with a B.A. in Geology and Russian Language with an emphasis in Computer Science.

Presentation #1: Categorical Data Analysis: Because There Really Is No Fate Worse Than Death AnnMaria De Mars, The Julia Group

Abstract

What do birth, death, high school dropout, failing a course, losing a sale, engineering majors and being fired all have in common? If you answered, "Things my mother warned me about", you need therapy. The correct answer is all of these are categorical variables, where data pretty much DO fit in neat little boxes. You either graduated high school or you didn't. The customer either bought insurance from you or opened the door and let his pit bull chase you down the street. The voter checked the box for the Democratic, Republican or Independent candidate. You get the idea. This presentation begins with a brief discussion of simple statistics that are part of the FREQ procedure and SAS/GRAPH. The remainder of the session is a straightforward approach to logistic regression using PROC LOGISTIC. Syntax will be provided to produce SAS analyses of categorical data and provide clear interpretations of the size of relationships, identify which variables matter in predicting outcomes and whip out impressive graphic displays of your data and relationships.

Biography

Dr. De Mars is president of The Julia Group, a statistical consulting company, and CEO of 7 Generation Games, a start-up producing educational games for elementary and middle school mathematics. She has been programming with SAS for 30 years, an amazing feat considering that she claims to only be 32 years old. Her experience with SAS categorical data analysis actually began with PROC CATMOD. Although that procedure is now obsolete, she is not, having progressed to PROC LOGISTIC, PROC SURVEYLOGISTIC and more. In a long and strange career, she has been senior statistical consultant at the University of Southern California, statistician on national surveys, industrial engineer monitoring machine loading and production targets and professor of statistics and research methods.

**Presentation #2:
Factor Analysis Is Your Friend
AnnMaria De Mars, The Julia Group**

Abstract

Why should YOU know about factor analysis? Imagine yourself in this scenario - someone, maybe you, has collected survey data at great expense. Maybe you paid subjects to answer questions about themselves, gave students credit to participate in a study, and now you have dozens, perhaps hundreds, of variables on each person. How on earth do you analyze these data? You could just go through and start putting questions together to form sub-scales, but that is pretty arbitrary. Enter factor analysis to help you make sense of your data.

Factor analysis is extremely useful. Conceptually, it is relatively easy to understand - mathematically, um, not so much so. Fortunately, SAS Enterprise Guide will have you analyzing with ease (really). You take a large number of questions and find what few, underlying traits they represent, such as supervision, collaborative decision making and ambition. How do you know the number of factors? How do you decide which survey item goes with which factor? Why would you rotate and which rotation would you use?

This presentation uses examples from the 500 Family Survey to demonstrate how to: generate screen plots, factor patterns and commonality estimates with SAS Enterprise Guide. Decision-making rules for number of factors, rotation and commonality are discussed. The need for iteration is explained and the ease of iteration with SAS Enterprise Guide is illustrated. Don't be scared - if you even half-way understood correlation in your college statistics course, you can master factor analysis. You'll be glad you did. (Code will also be provided for the Enterprise-Guide-phobic.)

**Presentation #3:
Strategies and Techniques for Getting the Most Out of Your Antivirus
Software for SAS® Users
Ryan Paul Lafler, Valhalla High School**

Abstract

Malware, sometimes referred to as malicious software, represents software threats engineered to damage computer systems without the knowledge of the owner using the system. SAS® users are increasingly becoming more prone to malware attacks and need to have a set of guidelines to help them get the most out of their antivirus software. This presentation highlights the many different types of computer threats, classification approaches, detection strategies, and removal methods. Attendees learn what malware is; the types of malware including viruses, Trojans, rootkits, zombies, worms, spyware, adware, scareware, spam email, and denial of service (DOS) attacks; password protection and management strategies; software to detect and protect computer systems; techniques for the removal of malicious software; and strategies and techniques for protecting your computer and data assets.

Biography

Ryan Paul Lafler is a software and security enthusiast, and student at Valhalla High School in El Cajon, California. Ryan has many interests in the world of computing including statistical, database and open source software; network security, including the threats that Malicious Software (Malware) poses to users and computer systems; and the removal and eradication of malware using Antivirus Software solutions.




The mission of Synchrony Solutions is simple:

To provide your company with timely and effective search solutions that exceed expectations. Our business is about relationships. We build our network, and our reputation, one person at a time, every time. When you have a position that needs to be filled, or you are ready to find your next opportunity, let us do the work for you!

molly@synchrony-solutions.com
Phone: 1-866-272-1228

www.synchrony-solutions.com

Biostatisticians & Statisticians
 SAS Programmers
 Clinical Data Managers/Coordinators
 Clinical Research Associates
 Regional CRA's
 Clinical Study Managers
 Director Clinical Operations
 Medical Writers
 Regulatory Affairs Specialists
 Senior Management
 Project Management
 Analysis
 Programming & Development
 Production Support
 LAN/WAN engineering
 Desktop network engineering
 Help Desk Support
 Web Developers
 Software Developers

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](mailto:Cathy.Liu): Name, Title, Organization, Telephone, E-mail, Fax, Address, Work or Home indicator, Computer Platform, SAS version, and SAS related areas of interest.

Stump the Programmer #65

Double SET Statement (Part 4)

Art Carpenter, CA Occidental Consultants

In the previous three Stump the Programmers, we examined the behavior of a DATA step with two SET statements. Use what you learned in those problems to write a many-to-many merge using the DATA step and two SET statements.

* Write a DATA step that matches each female with each male (many-to-many merge) - think dating service.

* This is trivial in a SQL step, but will test your understanding of a DATA step with two SET statements.

* The final data set should have 90 (10 X 9) observations and four columns (MNAME, MAGE, FNAME, and FAGE).

* [name and age for each gender]

```
*****;
```

```
data Males;
  set sashelp.class (where=(sex='M'));
run;
data Females;
  set sashelp.class (where=(sex='F'));
run;
```

```
data ManyMany;
  ..... your code here .....
  set females;
  ..... your code here .....
  set males;
  ..... your code here .....
run;
```

Data Explorations

Data Management & Analyses for Biological, Ecological & Health Sciences.

SAS® Programming Services: Extensive experience with SAS macros, SAS/STAT®, SAS/GRAPH®, SAS/AF®, SAS/FSP®, & SAS Component Language.

Statistical and data analysis support for clinical trial studies. Data entry and reporting systems. Application Development. Database conversion services. SAS Training services. Experimental Design and Data Management.

Data Explorations is a SAS Alliance Affiliate Member™ with SAS Advanced Certified Professionals.



Art Carpenter & Richard Smith - Partners

(907) 865-9167 Art
(760) 613-5373 Richard

Email: Art: Art@DataExplorations.com
Richard: RSmith@DataExplorations.com

Web: <http://www.DataExplorations.com>

Tips for Using a Data Warehouse

Curtis Smith, Defense Contract Audit Agency

In our last episode, we explored three of my favorite data set options that help resolve common issues within a DATA step when using a MERGE statement. This time, we will consider a handy way to create new rows in a data set. The key features we will explore are an iterative do loop and the OUTPUT statement.

In the example code below, we want to create a new data set with the name *work.expected* that will contain two variables, *firstdgt* and *expected*. We want a row for nine *firstdgt* values, 1 through 9, and, for each *firstdgt*, we want a calculated *expected* value. The *expected* value is calculated using a log function. So, the task is to put the DATA step through a loop of nine iterations, store the value of the iteration as the *firstdgt* variable and store the calculated value from the log function in the *expected* variable.

```
data work.expected(index=(firstdgt) drop=i);
  format expected 8.3;
  do i = 1 to 9;
    firstdgt=i;
    expected=(log10(1+(1/i))*100);
    output;
  end;
run;
```

The real work is being done in the DO loop. Before we explain the DO loop code, let's get some help on the subject from the SAS Institute. (Sas(r) 9.2 language, 2013)

The DO statement is the simplest form of DO group processing. The statements between the DO and END statements are called a DO group. The iterative DO statement executes statements between DO and END statements repetitively based on the value of an index variable. The index-variable names a variable whose value governs execution of the DO group.

So, in the example code, we are naming the index variable *i* and setting it to "1 to 9" so we will have nine iterations. So, every statement between DO and END will execute 9 times. The first iteration will have *i* equal to 1. Thus, *firstdgt* will be set to 1 and the log function will use "1" in the formula to calculate the value of *expected*. Without any help, SAS won't output a data row until the end of the DATA step. But, within the DO group processing, the value of *i* will change with each iteration and replace the values stored in *firstdgt* and *expected*. So, we need a way to force SAS to write out a data row at the end of each DO loop iteration. For this, we need a little help from the SAS Institute on the OUTPUT statement. (Sas(r) 9.2 language, 2013b)

The OUTPUT statement tells SAS to write the current observation to a SAS data set immediately, not at the end of the DATA step. By default, every DATA step contains an implicit OUTPUT statement at the end of each iteration that tells SAS to write observations to the data set being created. Placing an explicit OUTPUT statement in a DATA step overrides the automatic output, and SAS adds an observation to a data set only when an explicit OUTPUT statement is executed.

So, placing an OUTPUT statement at the end of the DO group statements will force SAS to write out a data row (observation) at the end of each iteration. That'll do it. One last thing, we didn't want the index variable to be retained in the output data set, so we used the drop= option.

Thanks for reading.

Curtis Smith, IT Technical Specialist

(2013). Sas(r) 9.2 language reference: Dictionary, fourth edition, do statement. Cary, NC: SAS Institute, Inc. Retrieved from <http://support.sas.com/documentation/cdl/en/lrdict/64316/HTML/default/viewer.htm#a000201951.htm>

(2013). Sas(r) 9.2 language reference: Dictionary, fourth edition, output statement. Cary, NC: SAS Institute, Inc. Retrieved from <http://support.sas.com/documentation/cdl/en/lrdict/64316/HTML/default/viewer.htm#a000194540.htm>

Kirk's Korner

Kirk Paul Lafler, Software Intelligence Corporation



Selecting the First 10 Results in a Query Request

An OUTOBS= option in an SQL procedure along with a SELECT clause can be used to control the number of rows to return. This approach is very useful on large tables with thousands or more rows because returning a large number of rows can adversely impact performance and/or provide too much information. **Note:** Not all database systems support an OUTOBS=, or equivalent, option as the SAS software does.

To demonstrate the effect of specifying an **OUTOBS=n** option, the following example illustrates a simple SQL query that selects the ten longest movies from the MOVIES table, after a descending order sort is performed.

SQL Code

```
PROC SQL OUTOBS=10;
  SELECT TITLE, RATING, LENGTH
  FROM MOVIES
  ORDER BY LENGTH DESCENDING;
QUIT;
```

SAS Log Results

```
PROC SQL OUTOBS=10;
  SELECT TITLE, RATING, LENGTH
  FROM MOVIES
  ORDER BY LENGTH DESCENDING;
WARNING: Statement terminated early due to OUTOBS=10 option.
QUIT;
```

SAS Results

Title	Length	Rating
Titanic	194	PG-13
Brave Heart	177	R
Scarface	170	R
Forrest Gump	142	PG-13
The Hunt for Red October	135	PG
Dracula	130	R
Ghost	127	PG-13
Jurassic Park	127	PG-13
Jaws	125	PG
Star Wars	124	PG

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



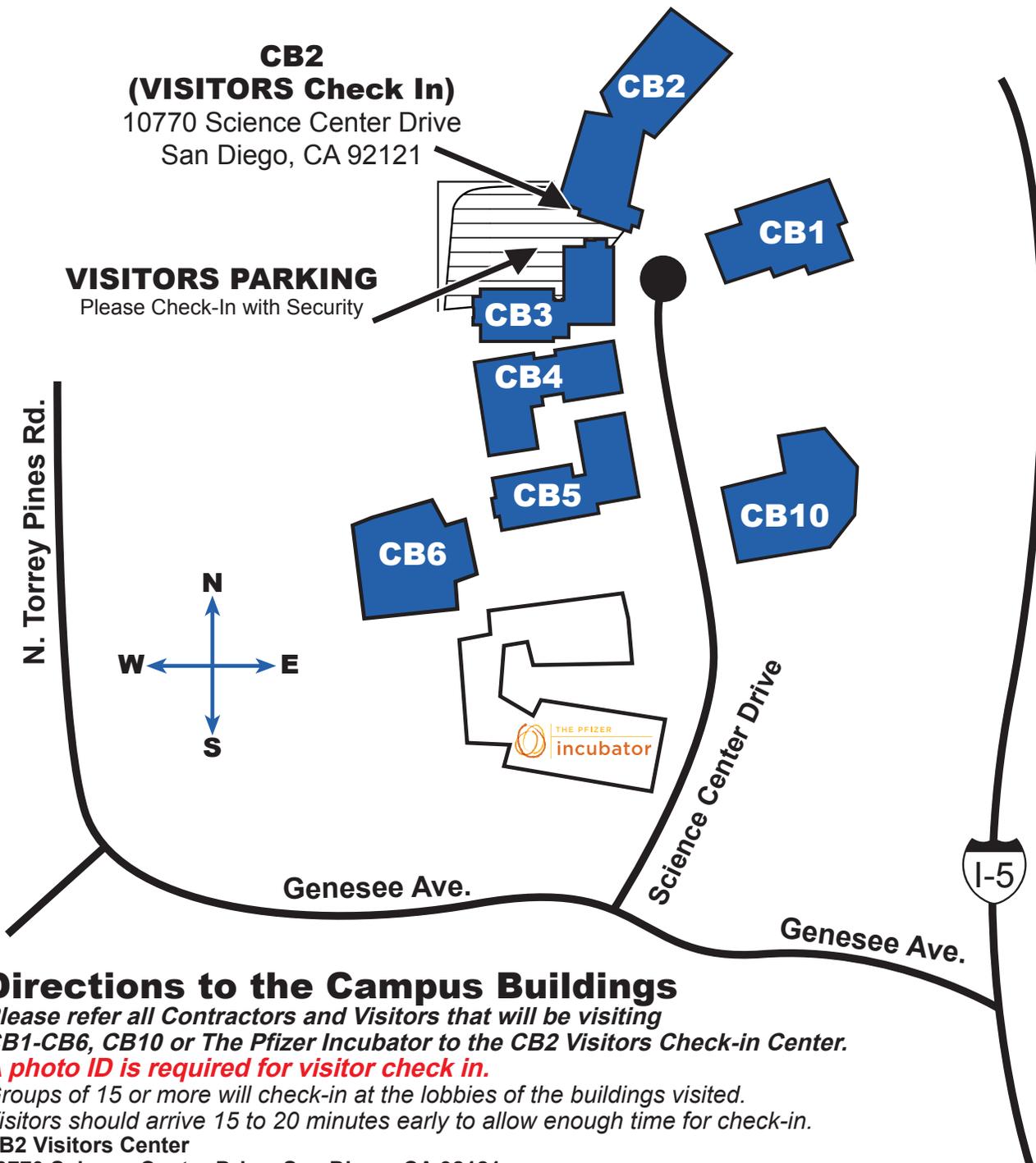
www.gerardgroupinc.com

650-360-5500

SANDS Sponsorship

We would like to thank [Pfizer](#) for having once again generously provided our meeting room. This month, the WUSS Executive Committee and SAS Institute are 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 coordinators, [Tyler Smith](#) and [Vivian Huang](#).

Map to Pfizer La Jolla Campus



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.**



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](mailto:Tyler.Smith@nu.edu) and/or [Vivian Huang](mailto:Vivian.Huang@biotheranostics.com).

Catering	Vivian Huang		Vivian.Huang@biotheranostics.com
Facilities Coordinator	Yiyun Tang		Yiyun.Tang@pfizer.com
	Ren-Yu Tzeng		renyutzeng@gmail.com
Job Bank	Robert Hull	760-268-8003	RHull@synteract.com
Membership Coordinator	Cathy Liu	760-476-8800	Cathy.Liu@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
sasCommunity.org
[LinkedIn](#)
[Facebook](#)