



FROM THE PRESIDENT

First of all, congratulations to our newly elected SANDS officers for year 2016! Wei Cheng, Archie Medrano, and Song Lin were elected as President, Vice President, and Treasurer, respectively, to serve our local SAS user community for another year! SANDS members acknowledge the hard work of the SANDS officers and Executive Committee members over the years. We appreciate the continuous support from our members!

Secondly, I have very exciting news for the upcoming August meeting! We will have three SAS speakers coming to San Diego this time around! The meeting will start at 1:00pm and will last until 7:00pm, so it will be different from our regular meetings, which run from 3:00pm to 8:00pm.

Dr. Robert Rodriguez will give the first presentation "Statistical Model Building for Large, Complex Data: Five Directions in SAS/ STAT® Software", followed by Sanjay Matange's presentation "Clinical Graphs Using SAS®". Chevell Parker will then give the third presentation "A Ringside Seat: The ODS Excel Destination Versus the ODS ExcelXP". The last presentation in the afternoon is Sanjay's second presentation "Annotate Your SGPLOT Graphs ". All three SAS speakers are internationally well-known SAS presenters. We are so lucky to have all of them present in our August meeting! Dr. Rodriguez is an old friend of SANDS. He has presented multiple times at SANDS meetings, including keynote presentations for our first three one-day meetings in 2005, 2008, and 2012. He was also the keynote speaker for our joint meeting with SDASA in 2010. It is the first time for Sanjay and Chevell to present at SANDS meeting. Their talks on ODS Statistical Graphs and Excel files are always in the top list of topics in SAS conferences, so you don't want to miss them!

After the networking dinner, we will have a panel discussion with the SAS speakers. You will get a chance to ask questions to all three SAS experts! Or, if you can send your questions to me before the meeting, I can pass the questions to them.

Last but not least, I would like to give thanks to our sponsors and advertisers. This time, all four presentations are sponsored by the SAS Institute and, in addition, SAS is providing the networking dinner and giveaways for the meeting. Once again, Andrea Littleton from SAS Americas Marketing and Support Department will attend our meeting. Andy is the local user group liaison from the SAS Institute. She has been providing timely support for all of our SANDS meetings in the past years, so please say thanks to her when you meet her in person. As always, Pfizer will provide our meeting locations for the whole year. SANDS meetings are held with no cost to our members because of their generous sponsorship. Thank you, SAS and Pfizer!

SANDS meetings provide opportunities to present your work and network with other SAS users, so we encourage you to submit abstracts of your presentations to us, or submit either your or other people's articles to us to be published in our newsletter.

We look forward to seeing you all on August 10th! - Wei Cheng





Time	Session	Speaker
1:00 - 1:50	Presentation: Statistical Model Building for Large, Complex Data: Five Directions in SAS/STAT [®] Software	Robert N. Rodriguez
2:00 - 2:50	Presentation: Clinical Graphs Using SAS®	Sanjay Matange
3:00 - 3:50	Presentation: A Ringside Seat: The ODS Excel Destination Versus the ODS ExcelXP	Chevell Parker
4:00 - 4:50	Presentation: Annotate Your SGPLOT Graphs	Sanjay Matange
5:00 - 6:00	Dinner / Networking	
6:00 - 6:15	Welcome / SANDS Business	
6:15 - 6:45	Panel Discussion with the SAS Speakers	
6:45 - 7:00	Stump the Programmer; Door Prizes and Giveaways	



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



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

Need Help with SAS or Math?

Algebra	SAT Math
Geometry	
Trigonometry	ELM Exam
Pre-Calculus	
Calculus	GRE Math
Discrete Math	
Linear Algebra	GMAT Math
Differential	
Equations	TEAS Math
Number Theory	

Call or text Archie at 858-633-8486 or email info@askawaytutoring.com

ASK A WAY TUTORING





Robert N. Rodriguez

SAS Institute

Bob Rodriguez joined SAS in 1983 and is a senior director in SAS Research & Development with responsibility for the development of statistical software. He received his PhD in statistics from the University of North Carolina at Chapel Hill and worked as a research statistician at General Motors Research Laboratories before joining SAS. Bob is a Fellow of the American Statistical Association, and he was the 2012 President of the ASA.

Statistical Model Building for Large, Complex Data: Five New Directions in SAS/STAT® Software

The increasing size and complexity of data in research and business applications require a more versatile set of tools for building explanatory and predictive statistical models. In response to this need, SAS/STAT software continues to add new methods. This presentation takes you on a high-level tour of five recent enhancements: new effect selection methods for regression models with the GLMSELECT procedure, model selection for generalized linear models with the HPGENSELECT procedure, model selection of generalized additive models with the GAMPL procedure, and building classification and regression trees with the HPSPLIT procedure. For each of these approaches, the presentation reviews its key concepts, uses a basic example to illustrate its benefits, and guides you to information that will help you get started.



Sanjay Matange

SAS Institute

Sanjay Matange is R & D Director in the Data Visualization Division at SAS, responsible for the development and support of ODS Graphics. This includes the Graph Template Language (GTL), Statistical Graphics (SG) procedures, ODS Graphics Designer and other related graphics applications. Sanjay has been with SAS for over 20 years and is author of two SAS Press books.

Clinical Graphs Using SAS®

Graphs are essential for many clinical and health care domains, including analysis of clinical trials safety data and analysis of the efficacy of the treatment, such as change in tumor size. Creating such graphs is a breeze with procedures from SAS[®] 9.4 ODS Graphics. This paper shows how to create many industry-standard graphs – such as Lipid Profile, Swimmer Plot, Survival Plot, Forest Plot with Subgroups, Waterfall Plot, and Patient Profile using Study Data Tabulation Model (SDTM) data – with just a few lines of code.

Annotate your SGPLOT Graphs

The SG procedures provide you multiple plot statements to create many different kind of graphs. These plot statements can be used together in creative ways to build your graph. However, even with this ability to customize, there are times when you need more than what you can get using just the plot statements. You need a way to add custom information anywhere on the graph.

With SAS 9.3, the SG procedures support the ability to annotate the graph using data set based information. This annotation functionality is designed in a way similar to the annotate facility available with the SAS/GRAPH procedures. There are a few differences and enhancements. If you already know annotation from SAS/GRAPH, or if you are new to it, this paper will show you how to add custom annotations to your graphs.







Chevell Parker

SAS Institute

Chevell is a member of the BASE/Graph group within SAS Technical Support. His support areas include the Output Delivery System and the XML Engine. Chevell has been a SAS user for over 20 years.

A Ringside Seat: The ODS Excel Destination Versus the ODS ExcelXP Tagset

The new and highly anticipated ODS destination for Microsoft Excel is finally here! Available as a production feature in the third maintenance release of SAS[®] 9.4 (TS1M3), this new destination generates native Excel (XLSX) files that are compatible with Microsoft Office 2010 or later. This paper is written for anyone, from entry-level programmers to business analysts, who use SAS and Microsoft Excel to create reports.

The discussion covers features and benefits of the new Excel destination, differences between the Excel destination and the older ExcelXP tagset, and functionality that exists in the ExcelXP tagset that is not available in the Excel destination. These topics are all illustrated with meaningful examples.

The paper also explains how you can bridge the gap that exists as a result of differences in the functionality between the destination and the tagset. In addition, the discussion outlines when it is beneficial for you to use the Excel destination versus the ExcelXP tagset, and vice versa. After reading this paper, you should be able to make an informed decision about which tool best meets your needs.

SANDS Sponsorship

We would like to thank <u>Pfizer</u> for having once again generously provided our meeting room. This month, the <u>SAS Institute</u> is sponsoring all four presentations and providing the networking dinner and giveaways. 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, <u>Tyler Smith</u> and <u>Vivian Huang</u>.

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 <u>Cathy Liu</u>: 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

In our last episode we explored a handy way to create a profile of our SAS data set so we can both document our data and also use the profile to verify the accuracy of our SAS data set. This time we will explore a method for rearranging our data so we can easily analyze it. (I know last time I said that this time we would explore an easy way to use SAS Enterprise Guide to create a data dictionary of our SAS data set. But, I think this time's topic will be more interesting.)

I often have need to analyze the same variable from different, but related types of data, and I want those related types of data side-byside in the same observation. So, for example, I want to relate indirect expenses by department to direct costs by department by month. I have one SAS data set that contains the indirect expenses in a variable named Dollars with each observation containing a variable for the Year-Month and another for the Department. I have another SAS data set with the same structure that contains the direct costs. To prep my data for the rearranging I need to do, I first created a temporary SAS data set from each of these two SAS data sets in which I added a new variable that I called D_I (for Direct_Indirect) which I populated with an "I" in each observation in the indirect expense SAS data set and with a "D" in each observation in the direct cost SAS data set. After summarizing the two data sets (PROC SUMMARY, one of my personal favorites) and combining the two data sets into a single SAS data set, the data looks something like this:

Department	Year-Month	D-I	Dollars
AB316	2013-01	D	\$1,973,620.45
AB316	2013-01	I	\$59,193.90
AB316	2013-02	D	\$2,691,211.10
AB316	2013-02	I	\$52,080.14
AB316	2013-03	D	\$3,491,730.76
AB316	2013-03	I	\$78,890.62
AB316	2013-04	D	\$2,671,943.48
AB316	2013-04	I	\$59,864.41
AB316	2013-05	D	\$2,778,048.90
AB316	2013-05	I	\$63,285.73
AB316	2013-06	D	\$3,319,321.89
AB316	2013-06	I	\$105,844.23
AB316	2013-07	D	\$2,410,532.01
AB316	2013-07	I	\$68,760.39
3B316	2013-08	n	C2 522 440 11

The first thing that I want to do to analyze this data set is to compute a ratio of the indirect dollars to the direct dollars and store that ratio in a new variable in the same observation along with the indirect and direct dollars. But, as you can see, the indirect and direct dollars are not stored in the same observation. So, here's where the rearranging is needed.

I heard somewhere that the TRANSPOSE procedure might do the trick. So, let's get some background on the TRANSPOSE procedure from the SAS Institute (Base sas(r) 9.4 procedures, 2016):

The TRANSPOSE procedure creates an output data set by restructuring the values in a SAS data set, transposing selected variables into observations. The TRANSPOSE procedure can often eliminate the need to write a lengthy DATA step to achieve the same result. Further, the output data set can be used in subsequent DATA or PROC steps for analysis, reporting, or further data manipulation.

Sounds like we found what we need. With a little help from SAS Enterprise Guide for some rapid application development, I have the following SAS code:

```
PROC TRANSPOSE DATA=WORK.D_I_Dollars
OUT=WORK.D_I_Dollars_Transposed
(LABEL="Direct/Indirect Dollars by Department")
PREFIX=Dollars NAME=Source LABEL=Label;
BY Department Year_Month;
ID D_I;
VAR Dollars;
RUN; QUIT;
```



The BY and VAR statements are probably obvious, but the ID statement is quite important. Getting some more help from the SAS procedures guide (Base sas(r) 9.4 procedures, 2016(2)):

ID Statement. Specifies one or more variables in the input data set whose nonmissing formatted values name the transposed variables in the output data set. When a variable name is being formed in the transposed (output) data set, the formatted values of all listed ID variables are concatenated in the same order that the variables are listed in the ID statement. The PREFIX=, DELIMITER=, and SUFFIX= options can be used to modify the formed variable name.

The PREFIX value is used in constructing names for transposed variables in the output data set. The NAME value specifies the name for the variable in the output data set that contains the name of the variable that is being transposed. The LABEL value specifies a name for the variable in the output data set that contains the label of the variable that is being transposed.

Department	Year-Month	UoM	DollarsD	Dollarsi	Ratio
AB316	2013-01	Dollars	\$1,973,620.45	\$59,193.90	2.999%
AB316	2013-02	Dollars	\$2,691,211.10	\$52,080.14	1.935%
AB316	2013-03	Dollars	\$3,491,730.76	\$78,890.62	2.259%
AB316	2013-04	Dollars	\$2,671,943.48	\$59,864.41	2.240%
AB316	2013-05	Dollars	\$2,778,048.90	\$63,285.73	2.278%
AB316	2013-06	Dollars	\$3,319,321.89	\$105,844.23	3.189%
AB316	2013-07	Dollars	\$2,410,532.01	\$68,760.39	2.852%
AB316	2013-08	Dollars	\$2,533,449.11	\$50,422.95	1.990%
AB316	2013-09	Dollars	\$3,031,591.88	\$38,019.68	1.254%
AB316	2013-10	Dollars	\$2,147,996.37	\$31,150.15	1.450%
AB316	2013-11	Dollars	\$2,926,558.18	\$36,775.70	1.257%
AB316	2013-12	Dollars	\$3,528,323.77	\$42,117.82	1.194%
AB316	2014-01	Dollars	\$2,182,557.92	\$35,073.61	1.607%
AB316	2014-02	Dollars	\$2,752,186.34	\$79,968.42	2.906%
30216	2014-03	Dellere	02 015 052 06	060 050 16	1 6479

After transposing my data, I use a simple SAS Data Step to create a ratio variable, and voila!, my results look like this:

(Behind the scenes I used a variable label to change the column specified in the LABEL= to display the text "UoM", Unit of Measure.) Notice how the TRANSPOSE procedure created the names for the two variables containing the dollars that were identified by the D_I code as "D" or "I". The PREFIX= value is "Dollars" so the variable names both start with "Dollars" followed by the value in the D_I variable, which is the variable that we identified in the ID statement.

Next time we will explore an easy way to use SAS Enterprise Guide to create a data dictionary of our SAS data set. (This time, I mean it.) Thanks for reading.

Curtis A. Smith DoD Project Manager <u>ca.smith86@att.net</u>

(2016). BASE SAS(R) 9.4 Procedures Guide, Fifth Edition, Overview: TRANSPOSE Procedure. Cary, NC: SAS Institute, Inc. Retrieved from https://support.sas.com/documentation/cdl/en/proc/68954/HTML/default/viewer.htm#p1n9l1b2dg00w9n1h3tovj0rwlos.htm

(2016). BASE SAS(R) 9.4 Procedures Guide, Fifth Edition, TRANSPOSE Procedure, ID Statement. Cary, NC: SAS Institute, Inc. Retrieved from https://support.sas.com/documentation/cdl/en/proc/68954/HTML/default/viewer.htm#p16njbxciji0qin1x6y7mzktujp4.htm



KIRK'S KORNER: THE 5 CATS IN THE HAT – SLEEK CONCATENATION STRING FUNCTIONS

SAS functions are an essential component of the SAS Base software. Representing a variety of built-in and callable routines, functions serve as the "work horses" in the SAS software providing users with "ready-to-use" tools designed to ease the burden of writing and testing often lengthy and complex code for a variety of programming tasks. The advantage of using SAS functions is evident by their relative ease of use, and their ability to provide a more efficient, robust and scalable approach to simplifying a process or programming tasks.

SAS functions span a number of functional categories, including character, numeric, character string matching, data concatenation, truncation, data transformation, search, date and time, arithmetic and trigonometric, hyperbolic, state and zip code, macro, random number, statistical and probability, financial, SAS file I/O, external files, external routines, sort, to name a few. This SAS tip highlights an old, an alternate, and new way of concatenating strings and/or variables together. The code, results and analysis appear below.

SAS Code: data null ; length NUM 3. A B C D E \$ 8 BLANK \$ 1 ; A = 'The';NUM = 5; B = ' Cats';C = 'in';D = ' the';E = 'Hat' ; BLANK = ' ' * Old way of concatenating with TRIM and LEFT functions and concatenation operator ; OLD = trim(left(A)) || BLANK || trim(left(NUM)) || BLANK || trim(left(B)) || BLANK || trim(left(C)) || BLANK || trim(left(D)) || BLANK || trim(left(E)) ; * Using the STRIP function and concatenation operator ; STRIP = strip(A) || BLANK || strip(NUM) || BLANK || strip(B) || BLANK || strip(C) || BLANK || strip(D) || BLANK || strip(E) ; * Using the CAT functions to concatenate character and numeric values together ; CAT = cat (A, NUM, B, C, D, E);CATQ = catq(BLANK, A, NUM, B, C, D, E) ; CATS = cats(A, NUM, B, C, D, E); CATT = catt(A, NUM, B, C, D, E); CATX = catx(BLANK, A, NUM, B, C, D, E);put OLD= / STRIP= / CAT= / CATQ= / CATS= / CATT= / CATX= / ; run ;

Results:

```
OLD=The 5 Cats in the Hat

STRIP=The 5 Cats in the Hat

CAT=The 5 Cats in the Hat

CATQ="The "5" Cats ""in ""the ""Hat

CATS=The5CatsintheHat

CATT=The5 Cats in theHat

CATX=The 5 Cats in the Hat
```

Analysis:

In the SAS code, above, a single numeric variable, NUM, and six character variables: A, B, C, D, E, and BLANK are defined with their respective values as: NUM=5, A='The', B=' Cats', C='in', D=' the', E='Hat' and BLANK=' '. The oldest way of concatenating two or more strings or variables together is then specified, using the TRIM and LEFT functions with the concatenation operator " | | " in an assignment statement. An alternate approach using a TRIM function with the concatenation operator " | | " is then specified in an assignment

...



statement to join two or more strings or variables together. Finally, the newer and more robust concatenation approach is then specified using the CAT family of functions: CAT, CATQ, CATS, CATT, and CATX.

- 1. CAT, the simplest of concatenation functions, joins two or more strings and/or variables together, end-to-end producing the same results as with the concatenation (double bar) operator.
- 2. CATQ is similar to the default features of the CATX function, but the CATQ function adds quotation marks to any concatenated string or variable.
- 3. CATS removes all leading and trailing blanks and concatenates two or more strings and/or variables together.
- 4. CATT removes trailing blanks and concatenates two or more strings and/or variables together.
- 5. CATX, perhaps the most robust CAT function, removes leading and trailing blanks and concatenates two or more strings and/or variables together with a delimiter between each.

Kirk Paul Lafler Senior SAS Consultant, Application Developer, Entrepreneur, Educator and Author

Software Intelligence Corporation E-mail: <u>KirkLafler@cs.com</u> LinkedIn: <u>http://www.linkedin.com/in/KirkPaulLafler</u> Twitter: @sasNerd <u>Google Search Complete!</u>

Stump the Programmer #76 by Art Carpenter

Missing Function

```
* MISSING function detects missing values;
* Justina;
data miss;
 input v1 v2;
  if missing(v1) then v1miss='Is Miss';
 if missing(v2) then v2miss='Is Miss';
  * Note when neither is missing;
  if missing(v1 or v2) then
    note='One or both missing';
datalines;
0 0
. 0
1 2
3.
. 5
. .
run;
proc print data=miss;
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: <u>Art@DataXplorations.com</u> Richard: <u>RSmith@DataXplorations.com</u>

Web: <u>http://www.DataXplorations.com</u>







ELECTED OFFICERS

President • Wei Cheng

Vice-President • Archie Medrano

Secretary/Treasurer • <u>Song Lin</u>

Advertising Coordinators • Tyler Smith & Vivian Huang

Туре	Dimensions (w x h)	Cost	
Business Card	3.5″ × 2″	\$25	
Quarter Page	3.5″ x 4.5″	\$50	
Half Page	7″ × 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 <u>Tyler Smith</u> and/or <u>Vivian Huang</u>.

Catering Coordinator	• <u>Vivian Huang</u>
Facilities Coordinators	 <u>Ren-Yu Tzeng</u> & <u>Yiyun Tang</u>
Job Bank Coordinator	• <u>Jia Hu</u>
Membership Coordinators	• <u>Cathy Liu</u>
SAS Institute Liaison	 <u>Andrea Littleton</u>
Speaker/Program Coordinator	 Kirk Paul Lafler
Sponsorship Coordinators	 <u>Vivian Huang</u> & <u>Tyler Smith</u>
Members At-Large	 <u>Scott Leslie</u>, <u>Charlie Shipp</u>

Newsletter Editor/Webmaster • Archie Medrano

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:	<u>sandsuq.orq</u>	sascommunity.org	LinkedIn	<u>Facebook</u>	<u>Twitter</u>