

DCSUG Newsletter

Washington, DC SAS Users Group

Fourth Quarter Meeting

When: Tuesday, December 3, 2002
Time: 8:30a.m. to noon
Place: Bureau of Labor Statistics
Postal Square Building
Room G440
First Street North East
Washington DC



Agenda:	8:30 - 9:00	Registration and refreshments
	9:00 - 9:15	Announcements
	9:15 - 10:15	David Shamlin, SAS V9 Nuts and Bolts
	10:15 - 10:30	Break
	10:30 - 11:00	Christopher Zogby, Zogby Enterprises, Inc. Proc SQL: Conditional Logic, Macro Variables and Subqueries
	11:00 - 11:15	User-to User Questions and Answers Moderators - DCSUG officers
	11:15 - 11:45	Andrew Karp, Sierra Information Services, Inc. Multilabel Formats Revisited Using ODS for Selecting "Optimal" Subsets of Independent Variables
	11:45 - Noon	Giveaway Drawings

***Presentation times subject to change**

Visitors must use the visitor's entrance on First Street NE. Directions for those taking Metro: Take the Red Line to Union Station. Exit the station via the Amtrak exit, but do not go up to the train level. Instead, after going up to and out through the farecard gates, turn to the left where there is an exit to the street (First Street NE). Directly across from the exit are the employee and visitor's entrances to the building. After checking in with security, take the main elevators down to the G level. **BRING PICTURE ID!**



If you have special needs and are planning to attend this meeting, please contact any DCSUG officer no later than 11/19/02

SAS® is a registered trademark of the SAS Institute, Cary, NC

NOTE: If you are not currently a paid member of the DCSUG and are planning to attend the meeting, please email kdennis@enterprise.foundation.org or call (410) 772-2780 by Nov 26th and leave your name. Security at the Bureau of Labor Statistics is very tight and having your name on a list of possible attendees will make your entrance to the building easier.

President's Message

Peter Knapp

I'd like to welcome you to the final meeting of 2002. As I reflect back on the past year, I realize that the aftermath of September 11 has been extensive and has impacted many people in so many ways. At the same time much healing has happened this past year, both individually and as a nation. That's a good sign that we can all overcome obstacles put in front of us and persevere.

A new challenge many of us will be facing in the new year is the transition to SAS V9. Our first speaker, David Shamlin from the SAS Institute, will be giving us a presentation on the SAS V9. In addition Christopher Zogby will be talking about Proc SQL and Andrew Karp will discuss Multilabel Formats and ODS. All of these presentations promise to be informative. I hope you get a chance to attend the meeting.

Speaker Abstracts

prior DCSUG meetings. (See page 7 for other presentations by Andrew Karp in the DC area)

V9 Nuts and Bolts

David Shamlin, SAS
Research and Development Director, Base Table Services

What's new in the V9 Release of Base SAS? Plenty of good stuff! We'll quickly review the direction SAS is taking its core product and briefly (I swear!) look at how the marketing message relates to the technological innovations R&D is putting into the product. But mostly we'll demo! demo! demo! A bunch of new tools have been added to the data step, ODS and other parts of Base that make life easier. Not only will we mention as many as we can, we'll show them in action.

David Shamlin joined SAS in 1987 as a member of the VMS Host group, where he helped develop low-level file systems. He also did a tour of duty with the SAS IO development team, becoming a pioneer in industry standard data access interfaces to SAS data stores. David brings these experiences to his current role as research and development director for Base Table Services where he leads the development of fundamental SAS technology related to the data step, BASE PROCs, the LIBNAME supervisor, data set and catalog IO, SAS/SHARE and other client/server components related to SAS data access. David holds a bachelor's degree in computer science from North Carolina State University.

Multilabel Formats Revisited

Andrew Karp, Sierra Information Services, Inc.

This 15 minute talk is a follow-on to the presentation Andrew gave on this topic at a previous DCSUG meeting. It shows how the new MULTILABEL Format facility can be used to create Formats with overlapping ranges, and also provides clarifications and additional insights in to MLFs that Andrew did not present previously.

Using ODS for Selecting "Optimal" Subsets of Independent Variables

Andrew Karp, Sierra Information Services, Inc

While most attention has been paid to ODS capabilities to render/display output, this new feature of the SAS System offers ways to speed the work of data analysts and statisticians, too. This 15-minute presentation shows how a combination of PROC REG, ODS, Data Step and Macro Language tools are combined to have SAS automatically assess over 1,000 linear regression models, each with a different combination of independent variables, and identify one statistically optimal combination. That combination is then passed to a subsequent PROC REG task, via a Macro variable, to obtain parameter estimates and other measures of model performance.

Andrew Karp is President of Sierra Information Services, Inc., a SAS Software consulting and training firm located in California wine country. A 20-year SAS Software user, Andrew is a SAS Certified Professional for Version 8 of the SAS System and has given dozens of papers at SAS users group meetings around the world. A graduate of The George Washington University and former Dupont Circle area resident, Andrew regularly returns to DC and has spoken at several

Proc SQL Tips: Conditional Logic, Macro Variables and Subqueries.

Christopher Zogby, Zogby Enterprises, Inc.

Proc SQL offers a number of features that make it a significant coding asset. For a variety of coding tasks, proc SQL can offer the most compact and novel solution. In this paper I will discuss three features that have helped increase my SQL mileage.

Chris Zogby is a Washington, DC area data mining consultant and proprietor of Zogby Enterprises, Inc.

DCSUG 2002 Steering Committee Officers

CHAIR -- Peter Knapp

(202) 482-1359 (202) 482-1388(fax)
peter_knapp@ita.doc.gov

SECRETARY -- Larry Altmayer

(301) 763-2569 (301) 457-4437 (fax)
Lawrence.W.Altmayer@census.gov

TREASURER -- Arlene Siller

(301) 458-4498 (301) 458-4032 (fax)
asiller@cdc.gov

SAS LIAISON -- Mike Rhoads

(301) 251-4308 (301) 294-2040 (fax)
rholdsm1@westat.com

OTHER COMMITTEE MEMBERS

Karen Dennis

(410) 772-2780 (410) 294-715-2310 (fax)
kdennis@enterprisefoundation.org

Frank Fry

(202) 452-2666 (202) 452-6433 (fax)
frank.fry@frb.gov

Christopher Zogby

(703) 671-6330 (703) 671-6360 (fax)
czogby@erols.com

Linda Atkinson

(202) 694-5046 (202) 694-5718 (fax)
Atkinson@ers.usda.gov

Howard Schreier

(202) 482-4180 (202) 482-4614 (fax)
Howard_Schreier@ita.doc.gov

DESKTOP SPECIAL INTEREST GROUP

CO-CHAIR -- Joe Mirabal

(703) 265-4620
jmmirabal@aol.com

CO-CHAIR -- Yesvy Gustasp

(301) 625-9576 (202) 307-5888 (fax)
Yesvy_Gustasp@yahoo.com

SAS Talk

By Ian Whitlock

It is time to take a look at %EVAL. As you probably know, all macro data is character. Why? Well macro is a programming language for writing programs to write SAS code. It would be awkward in the extreme if every bit of SAS code, i.e. data to macro, had to be quoted. So why does SAS use quotes for character data? Well in SAS, expressions that begin with a letter or underscore have a SAS meaning as key words or the names of things. For that matter so do expressions that begin with a number as literal numbers. Consequently literal character data must be distinguished in some way. SAS uses quotes for this purpose. So do many other programming languages because the amount of character data in a good program should be relatively small, because code is a very poor way to store data. In macro, it is the instructions that are distinguished, not the data. Macro instructions begin with a percent sign and end with a semicolon, while macro variable references contain ampersands. Consequently every thing else, that is not part of an instruction or a macro variable reference, can be and is macro data.

So what does this have to do with %EVAL? Well, what about numbers? Every programming language has to have some concept of numbers. How can you loop 3 times if there is no concept of 3? There are times when macro instructions expect a number. For example in:

```
%if &cond %then %do ;  
%do i = &start %to &stop %by &increment ;  
%let x = %substr ( &string , &start , &length ) ;
```

The macro variables, COND, START, STOP, INCREMENT, and LENGTH in the context of the above macro instructions all have to represent integers (usually positive or zero). What happens when their values are expressions, say STOP has the value 4 + 3? Well the expression has 5 characters 4, +, 3, and two spaces. So how do we get the number 7? We need a macro function to do arithmetic on arithmetic expressions! Aha, you see there is %EVAL. Where? Each place a number is expected %EVAL is implicitly called to do the arithmetic producing a number. Of course you can also use the function explicitly, as in

```
%let i = %eval ( &i + 1 ) ;
```

But most of the time it is not explicit. What happens if the expression isn't a valid arithmetic expression? Well, the same thing that happens in any program where you do not play by the rules. The program doesn't work. With a little luck this means you get an error message instead of just wrong results. %EVAL is obliging and responsible for reporting many of my errors in macro programs, usually in the form:

```
ERROR: A character operand was found in the %EVAL function or %IF condition where a numeric  
operand is required. The condition was: _____
```

Sometimes one gets a surprise. For example,

```
%put %eval ( 1.2 + 3 ) ;
```

produces the error message shown above with the blanks filled in with 1.2 + 3. The problem is that 1.2 is not a number (integer). A language for manipulating text, i.e. SAS code, has little need for any numbers other than integers and %EVAL refuses to cooperate with those who expect more. Another surprise is:

```
%put %eval (lt) ;
```

Here, the value 0 is written to the log. So how do the letters LT become 0? Well, sometimes a symbol or group of letters will have a meaning in the context of some macro instruction. Note that it is still true that LT is just a list of letters and does not have meaning to macro. It has a meaning to %EVAL, within the context of an argument to %EVAL. In this context, it means LESS THAN as in the logical expression "1 lt 1" or "a lt a". Both logical expressions have the value 0 because they are false; something cannot be less than itself. Even nothing cannot be less than itself, and that is why %EVAL(LT) is 0. %EVAL sees the LESS THAN and asks - what is to the left, nothing; what is to the right, nothing. So the value is 0. However, there is a subtle point here. What kind of comparison was done? In "1 lt 1" we compare numbers and in "a lt a" we compare characters. So %EVAL did have to make a decision to do one kind of comparison or the other. Which one? Well, nothing is not number. If you disagree, take a look at %EVAL (3 +). you will get the dreaded error message. What was the character encountered? It must be the nothing that follows the plus sign. So in %EVAL(LT) there could not be a number comparison, because there are no

numbers. It must have been a character comparison. So what is %EVAL(LT + 4)? Well, nothing is on the left of the LT as we have already seen. What about to the right? Well "+ 4" is there and that is something. So the value must be 1, i.e. true, because nothing is less than something.

We still have not answered the question in general for one comparison operator – how does %EVAL decide which kind of comparison to do? If one of the operands is character then a character comparison should be done. Most of the time the rule is good, but it is not always so clear when something is character as opposed to part of a numeric expression. Consider, %EVAL(+ LT -). Clearly the plus and minus signs are not numeric expressions, so you might think it clearly requires a character comparison. But again, the result of a %PUT is the dreaded error message:

ERROR: A character operand was found in the %EVAL function or %IF condition where a numeric operand is required. The condition was: + lt -

Here things get a little sticky. I do not think the error is coming from the primary %EVAL. Instead the plus and minus signs indicate two arithmetic expressions to be evaluated with implied %EVAL's. In other words, the expression is equivalent to %EVAL(%eval(+) LT %eval(-)), where the lowercase %eval's are implied. It is the first implied %eval that encounters a character operand.

So it is time to do a little quoting and hide the meaning of the plus sign from %EVAL. In other words, we want to hide the meaning of the plus sign as an arithmetic symbol in an arithmetic expression. The function %STR will hide some constant symbols. So look at %EVAL (%STR(+) LT %STR(-)). The result is 1 on an ASCII machine. Good, the plus sign has the value 43 and the minus sign has the value 45 in the ASCII collating sequence. So a plus sign should be less than a minus sign. Now, we all know that since A is less than B, AX is less than BC in spite of the fact that X comes after C. Why? Because that is the way dictionary order works. So now what is %EVAL (%STR(+3) LT %STR(-7))? Well, it is 0 because negative numbers come before positive numbers. What happened?

The %STR function did hide the signs when only signs were involved. But when the numbers were added to the signs then the arithmetic expressions were compared. So something removed the quoting of the signs. Test it out with %EVAL (%STR(+3) + %STR(-7)). Sure enough the result is -4. We get the same result even when we separate the symbols with %EVAL (%STR(+3) + %STR(-7)). So %STR will hide these symbols, but only some of the time. Upon reviewing this, Quentin McMullen noted %EVAL (%STR(2+15) LT 3) is true so the comparison is character. By what rule does %EVAL decide comparison type? I do not know.

%STR also hides the meaning of blanks; hence %LENGTH(%STR()) is 0 and %LENGTH(%STR()) is 5. However, when we compare the two quantities with %EVAL(%STR() eq %STR()) the result is 1. Again, the quoting is removed in making the decision about what form of equality to use in the comparison. Thus both are reduced to blanks, with the leading and trailing blanks deleted. So, again, we are comparing nothing with nothing.

In conclusion, we end where we began. It is time to look at %EVAL.

Election of 2003 DCSUG Officers



Peter Knapp — Candidate for Chair

Peter Knapp interned at the Bureau of Labor Statistics (BLS) where he learned the joys of SAS on an MVS mainframe environment. After contracting with BLS for three years, he joined the International Trade Administration (ITA) at the Commerce Department. For ten years Mr. Knapp has helped administer international trade law at ITA. During this time he has guided SAS processing from a central mainframe shop to an end user PC based system.

Arlene Siller -- Candidate for Treasurer

Arlene Siller is a Computer Specialist with the National Center for Health Statistics and has used SAS for 24 years. Her responsibilities include programming and analysis using SAS/Base and SAS/Stat for National health surveys. She is the SAS Site Representative for her agency and co-chair of the in-house NCHS Users Group (NCHSSUG). She is very involved in providing support for NCHS SAS users, and

enjoys this personal interaction part of her responsibilities the most. She has programmed using the TSO, Wylbur, VMS, DOS, Windows and UNIX operating systems.

Candidate for Secretary - OPEN

Secretary responsibilities: Update member database, confirm meeting room reservation, forward membership checks to treasurer – please contact any steering committee member to volunteer.

Mike Rhoads -- Candidate for SAS Liaison

Mike Rhoads has served on DCSUG's steering committee for the past several years. He has also contributed to various SUGI and NESUG conferences as a paper presenter and session coordinator. 2003 marks the 24rd anniversary of his initiation into the wonderful world of SAS software. He is currently a Vice President in the Computer Systems and Applications group at Westat, where he still manages to sneak in an occasional DATA step or two when not overburdened by his other responsibilities.

DCSUG 2003 BALLOT

Place an "X" In the Box beside the candidate of your choice. You may write in the name of someone who has not been nominated if you wish.

DCSUG Officers

Chairperson

Description: Lead DCSUG, delegate assignments, chair group meetings, preside over nomination and election of officers.

Peter Knapp

Treasurer

Description: Manage the group finances: collect dues and fees, make disbursements, maintain record of income and disbursements. The Treasurer will also prepare a report of income and disbursements for distribution at the fourth quarter general meeting.

Arlene Siller

Secretary

Description: Record minutes of group meetings, maintain membership list, prepare meeting notices.

OPEN

SAS Liaison

Description: Represent and increase the influence of DCSUG to the SAS Institute

Mike Rhoads

DCSUG Dues

Dues are set annually by the steering committee and voted on by the member ship at the fourth quarter meeting. Dues will be payable by the first quarter general meetings. Dues will be used to finance mailings, telephone charges, newsletter and other expenses as approved by the group of officers.

Keep proposed (current) rate (\$10.00 per year individual, \$50.00 per year corporate)
 Increase Decrease

Please mail ballot to:

**Washington DC SAS Users Group
P.O. Box 44670**

Washington, DC 20026-4670

Calendar of Meetings and Events

This feature is meant to give users an idea of some of the activities of area users groups and special interest groups, as well as regional and international groups. If your group would like to be regularly included in this feature, please call Mike Rhoads or one of the DCSUG officers. There is a DCSUG directory with phone, fax, and E-mail contact information elsewhere in this newsletter.

December

3 **DCSUG** DC SAS Users Group, General Meeting
 Time: 8:30 am - 12:00 noon
 Site: Bureau of Labor Statistics
 2 Massachusetts Ave. NE
 Program: See front page announcement
 Contacts: Mike Rhoads at (301) 251-4308
 RhoadsM1@Westat.com
 Peter Knapp at (202) 482-1359
 peter_knapp@ita.doc.gov

February 2003

6 **DCSUG** DC SAS Users Group, Special Evening Meeting
 Time: 6:00 pm - 8:00 pm
 Site: George Washington Univ.,
 K St. Center for Professional Education
 2020 K Street NW
 Program: See article
 Contacts: Mike Rhoads at (301) 251-4308
 RhoadsM1@Westat.com
 Peter Knapp at (202) 482-1359
 peter_knapp@ita.doc.gov

March 2003

4* **DCSUG** DC SAS Users Group, General Meeting
 Time: 8:30 am - 12:00 noon
 Site: TBA
 Program: TBA

* - tentative date

Northern Virginia SAS Users Group (NOVASUG) -

General Meeting

These meetings are open to all people free of charge. No registration is necessary.

Date: Thursday November 21, 2002

Location: Freddie Mac 8200 Jones Branch Dr. McLean, VA 22102

Time: 6:30 - 8:30 P.M.

March/April 2003

30-2 **SUGI 28** 28th Annual SAS Users Group International Conference
 Site: Seattle
 Chair: Art Carpenter
 Further Info:
www.sas.com/usergroups/sugi/sugi28/index.html

Other Groups

CENSUG Census SAS Users Group
 Contact: Andrea Potter at (301) 457-1687

MDSUG Maryland SAS Users Group
 Contact: Mr. R.H. Miller, (410) 740-4229
 millerrh@juno.com

NCHSSUG National Center for Health Statistics SAS Users Group
 Contacts: Arlene Siller at (301) 458-4498
 asiller@cdc.gov
 Linda Tompkins at (301) 458-4533

NIHSUG National Institutes of Health SAS Users Group
 Contact: Ray Danner at (301) 496-6037

NOVASUG Northern Virginia SAS Users Group
 Contact: Gary McQuown at (703) 628-5681
 gmcquown@dasconsultants.com

SIGSTAT Special Interest Group in Statistics, CPCUG and WINFORMS
 Contact: Charlie Hallahan, 202-694-5051
 hallahan@ers.usda.gov

VASUG Virginia SAS Users Group (Richmond area)
 Contact: Mark Jordan
 President@vasug.org

Programming for Job Security Revisited:

Even More Tips and Techniques to Maximize Your Indispensability - Art Carpenter, California Occidental Consultants and Tony Payne, SPSInfoquest Ltd.
 Using MULTILABEL Formats - Andrew H. Karp, Sierra Information Services, Inc.

For more information: <http://www.novasug.org/>

Calendar of Meetings and Events (continued)



Special Evening Meeting

Andrew Karp, a frequent presenter at DCSUG meetings, has offered to present at an evening meeting

Thurs Feb 6, 2003 at the GWU K Street Center, 2020 K Street NW

The meeting will run from 6 to 8 pm. Topics to be determined.

The event is hosted by Sierra Information Services, Inc. and light refreshments will be served.

Upcoming Seminars:

January/ February 2003 GW Univ. K Street Center for Professional Education
 Farragut West and North Metrorail Stations 2020 K Street NW
 Time: 9am – 5pm

<u>Date</u>	<u>Seminar Topic</u>	<u>Instructor</u>
1/29	<i>Manipulating Data with SAS Functions and Arrays</i>	Ben Cochran
1/30	<i>Treatment of Missing Data Via Multiple Imputation and Maximum Likelihood (Day 1)</i>	Paul Allison
1/30	<i>Introduction to SAS/SQL Processing</i>	Ben Cochran
1/31	<i>Treatment of Missing Data Via Multiple Imputation and Maximum Likelihood (Day 2)</i>	Paul Allison
1/31	<i>Advanced SAS/SQL Programming</i>	Ben Cochran
2/1	<i>ODS for Data Analysts and Statisticians</i>	Andrew Karp
2/3	<i>Intro to Data Summarization and Reporting Data Using the SAS System</i>	Andrew Karp
2/4	<i>Advanced Reporting with the SAS Output Delivery System</i>	Andrew Karp
2/5	<i>SAS System Programming Efficiencies: Tips & Techniques</i>	Andrew Karp
2/6	<i>Building and Applying Predictive Models Using the SAS System</i>	Andrew Karp
2/6	<i>Introduction to ODS Templates: No Experience Necessary for Success</i>	Bernadette Johnson
2/7	<i>ODS Templates for the Web: Using HTML Tagsets and Cascading Style Sheet (CCS) Tools with PROC TEMPLATE</i>	Bernadette Johnson
2/10	<i>Intro to SAS/GRAPH Software</i>	Art Carpenter
2/10	<i>SAS ODS for Data Analysts and Statisticians</i>	Andrew Karp
2/11	<i>From %MACRO to %MEND: An Intro to the SAS Macro Language</i>	Art Carpenter
2/12	<i>Advanced Topics in the SAS Macro Language</i>	Art Carpenter
2/13	<i>Introduction to SAS/STAT Software</i>	Art Carpenter

Registration and payment details for these training sessions are available at www.SierralInformation.com, via email at SierralInfo@aol.com or voice 707 996 7380

2003 Meetings

Plan ahead and mark your calendars now. The first two DCSUG meetings in 2003 are TENTATIVELY scheduled for March 4 and June 3. Meeting times and locations to be announced. If at any time you have a question about an upcoming meeting or other DCSUG information, call one of the members of the Steering Committee.



To visit the DCSUG home page

The URL for the DCSUG home page on the Web:
<http://www.ita.doc.gov/industry/otea/dcsug>

Remember to Join DCSUG!!!



Don't forget to renew your DCSUG membership for 2003. Dues are used to defray the costs of producing and mailing the newsletter and provide refreshments at our meetings. You'll make sure that you keep receiving the newsletter and keep DCSUG going. If you have not joined DCSUG, now is a great time to do so. Joining DCSUG is as easy as completing the membership form included in this newsletter. Individual memberships are only \$10; corporate memberships are \$50



Membership Application/Renewal

Date: _____

Name: _____

Company Name: _____

Mailing Address: _____

City, _____

State _____ Zip _____

Phone # _____

Fax # _____

E-Mail: _____

Preferred mode of contact (check one):

Fax

e-mail

SAS Version(s) _____

SAS Product(s): _____

Operating System(s) _____

Check one:

Individual membership

for 1 year is \$10.00

Corporate sponsorship

for 1 year is \$50.00

Please indicate a contact person for Corporate Sponsorships!

Check one:

New membership

Renewal membership

Mail to:

Washington DC SAS Users Group

P.O. Box 44670

Washington, DC 20026-4670