

# DCSUG Newsletter

Washington, DC SAS Users Group

## Third Quarter Meeting

When: Tuesday, September 10, 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 - 9:45	Paul Gorrell, Social & Scientific Systems, Inc. <b>Guidelines for Storing SAS Numeric Data</b>
	9:45 - 10:15	Peter Knapp, U.S. Department of Commerce <b>Debugging 101</b>
	10:15 - 10:30	Break
	10:30 - 10:45	User-to User Questions and Answers Moderators - DCSUG officers
	10:45 - 11:45	Ian Whitlock, Westat <b>SAS® Macro Design Issues</b>
	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 8/29/02

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](mailto:kdennis@enterprise.foundation.org) or call (410) 772-2780 by Sept 4<sup>th</sup> 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 especially look forward to the September meeting, as I will be giving a presentation of a paper I will be presenting at NESUG. It will give me an opportunity to practice my presentation skills on a smaller, more intimate audience. I look forward to your feedback and comments so I can refine my presentation for NESUG. The other two presenters at September's DCSUG meeting, Paul Gorrell and Ian Whitlock, will also be presenting their NESUG papers. After hearing our presentations, I hope we will inspire those of you who have not already planned to attend NESUG this year to do so. The conference is from September 29 - October 2 in Buffalo, NY. Many members of DCSUG are working hard to make this conference a success. For a bit of Southern hospitality, SESUG is presenting their annual conference in Savannah, Georgia from September 22 - 24. Hopefully you'll be able to attend one or both conferences. Come out and show your support!

## Speaker Abstracts

### Guidelines for Storing SAS Numeric Data

Paul Gorrell, Social & Scientific Systems, Inc.

Understanding how SAS stores numeric data is important for making informed decisions regarding LENGTH specifications, numeric precision in computations, and efficient use of disk space. I will review aspects of floating-point representation that are relevant to programming decisions, and discuss situations where numeric length should not be reduced, even if the range of integer values would appear to permit it. For saving disk space, I'll demonstrate the advantages of the SAS V8 option COMPRESS=BINARY.

*Paul Gorrell is a Senior Programmer Analyst at Social & Scientific Systems, Inc. September 2002 is the 20th anniversary of his first SAS program. These days he works primarily with health-care survey data, especially the Medical Expenditure Panel Survey [MEPS].*

### Debugging 101

Peter Knapp, U.S. Department of Commerce

The aim of this paper is to show a beginning user of SAS how to debug SAS programs. New users often review their logs only for syntax errors that appear in red. They neglect to look for other types of coding errors. SAS identifies non-syntax errors in notes and warnings. Examples include notes that SAS has found uninitialized variables, generated missing values, or encountered more than one data set with repeats of BY values in a MERGE statement. In addition, once all coding errors are cleaned up, new users can find that their programs do not produce the desired results. To produce the desired results they need to uncover logic errors that are often more difficult to find than coding errors. By demonstrating various debugging techniques, I plan to show that with a little practice, one can master the art of debugging SAS programs.

*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 twelve 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. He is the current president of the DC SAS Users Group.*

### SAS<sup>®</sup> Macro Design Issues

Ian Whitlock, Westat.

Two questions motivated this paper. The first came more than twenty years ago at the end of my first programming course when the professor asked, "You now know how to

write programs, but do you know when it is appropriate to write a program?" For SAS macro the question might be rephrased. What kind of SAS problem is appropriate to macro? Or when should the solution involve macro code?

The second came from SAS-L in September 2001, "What sources discuss how to develop clear and good macro code?" There have been many Beginning Tutorials on macro, and many of the examples in them may illustrate some of the principles that these questions hint at, but I do not know of any source that concentrates on the principles.

To sum up, I will consider: 1) What makes a macro good? and 2) How do you make a maintainable, useful, clear system of macros? Do you need to know how to write %IF and %DO statements to understand this talk? No, all terms will be introduced, however, knowing something about these statements will increase your appreciation of the subject.

*Dr. Ian Whitlock is a senior systems analyst at Westat working on survey research problems. He has been programming largely in SAS since 1981 using an interactive environment to develop programs for execution in a batch environment. He is active on SAS-L, and has presented many different papers at national, regional, and local SAS User Group meetings.*

## Questions and Answers

Moderated by DCSUG Officers



Bring your SAS programming or procedure questions to DCSUG and its members. During this open session, you can ask questions of and get answers from experienced SAS users. All are welcomed to participate

Due to the high cost of mailing out newsletters, DCSUG will be emailing notices of the availability of accessing the quarterly newsletter (and other important announcements) at the DCSUG Internet site. If you are a member, please make sure your email address is current in our database. Please notify Larry Altmayer at [Lawrence.W.Altmayer@census.gov](mailto:Lawrence.W.Altmayer@census.gov) with the email address you want to use to receive notification of the publication of the newsletter. Members will still have the option of continuing to receive a hard copy of the newsletter upon request.

## DCSUG 2002 Steering Committee

### Officers

#### CHAIR -- Peter Knapp

(202) 482-1359 (202) 482-1388(fax)  
peter\_knapp@ita.doc.gov

#### SECRETARY -- Larry Altmayer

(301) 457-2581 (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)  
rhoadsm1@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

### Desktop Special Interest Group

The desktop SIG of DCSUG will have a presentation on EDA [exploratory data analysis] using JMP at the FDA, some afternoon in late summer or early fall (Please keep checking the web site).

The primary speaker is David L DesJardins. He will cover material that users have found beneficial from his NESUG & SUGI papers. The theme will be 'Using easily generated "live" graphs to effectively edit and analyze statistical data.' Questions, regarding corresponding use of JMP's point and click "sister" package, SAS/ Insight, will also be answered.

Check the DCSUG home page for further information: <http://www.ita.doc.gov/industry/otea/dcsug> Please contact Yesvy Gustasp or Joe Mirabal (see Steering Committee column) with questions, requests, suggestions, and offers.

### Nominations for 2003 DCSUG Officers

It's that time of year again. We need to elect officers for DCSUG for next year and we need your help. Please nominate people to fill the following positions. Give your nominations to any current DCSUG officer or send them to:

DCSUG  
P.O. Box 44670  
Washington, DC 20026

President

Secretary

Treasurer

SAS Liaison

←-----→

## 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



# Interleaving a Dataset with Itself

*By Howard Schreier*

When two or more SAS datasets are combined by means of a SET statement and an accompanying BY statement, they are said to be interleaved. For example:

```
data ab;
  set a b;
  by id;
run;
```

The observations in AB appear in the following order: any observations from A having the lowest value of ID, any observations from B having that same value of ID, any observations from A having the next-lowest value of ID, etc. The last observations in AB would be those from B having the highest value of ID, or (if it happened that B had no observations with that value of ID), those from A having the highest value of ID.

But the SET statement does not have to designate different datasets. It can designate the same dataset multiple times. For example:

```
set a a;
by id;
```

Why might you want to do that? A fairly common type of SAS programming problem involves the need to treat an entire BY group in a way which depends on some property of the group. The difficulty is that you need to process the entire group to determine the property, so when the DATA step is processing the first observation, the property is unknown.

In a very simple example, you want to create a new variable indicating the number of observations in the group. In other words, given

```
Obs  id
 1  P
 2  P
 3  Q
you want
Obs  id  count
 1  P   2
 2  P   2
 3  Q   1
```

The traditional SAS approach is to do it in two steps, one to prepare the counts and the other to combine the counts with the original data. There are also solutions using PROC SQL. But here is a solution using a single DATA step:

```
data counted;
  set given(in=preview)
    given;
  by id;
  * Initialization;
  if first.id then count = 0;
  * Evaluation;
  if preview then count + 1;
  * Processing;
  else output;
run;
```

Here dataset GIVEN is being interleaved with itself, since it is named twice on the SET statement and the SET statement is under control of a BY statement. So each BY group will be seen twice. The use of the IN= dataset option is very typical in this situation, as it allows the program logic to distinguish between the first and second pass through each group.

In this simple example, the remainder of the DATA step comprises just three conditional statements. However, each of these is representative of what is often a larger block of code.

The first IF statement resets the counter at the start of each BY group. It is representative of whatever initialization is needed between groups.

The second IF statement distinguishes between the two passes and increments the counter during the first pass. It is representative of the code which evaluates the data and determines the required property or properties of a BY group.

The ELSE statement is related to the preceding IF and so is triggered by observations in the second pass. It is typical to have an OUTPUT statement here so that observations are written out only in the second pass (that is, after the required property of a BY group has been determined). In a more complicated "real world" program, this block would probably have assignment statements and other additional code for processing.

Thus, this simple example serves as a template for more complicated problems. One would set up a DO; END; block for each of the three conditional blocks (initialization, preview/evaluation, processing/output) and code each one according to the requirements of the problem.

## ***SAS PROGRAMMER***

**Description:** Statistical Programmer fluent in SAS (Statistical Analysis System) to support epidemiologic studies involving large file manipulation of databases of the Department of Veterans Affairs and statistical analyses of longitudinal data.

**Requirements:** Competence using the statistical application SAS including the writing of SAS MACROS; experience with mainframe and PC. College or graduate school courses in mathematical statistics (at least 12 credit hours) and 2-3 years relevant experience.

**Salary:** Civil Service position GS 9/11 : \$ 38,406 - \$ 60,405

**Contact:** Send CV to Han K. Kang, Dr PH, Director, War Related Illness and Injury Study Center (WRIISC), VA Medical Center, 50 Irving St. Washington DC 20422

## 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.

### September/October

10     **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

22-24 **SESUG 2002** Southeast SAS Users Group 9th Annual Conference  
 Site:        Hyatt Regency Savannah  
 Co-Chairs:  David Maddox, Heidi Markovitz  
 Further Info: [www.sesug.org/SESUG02](http://www.sesug.org/SESUG02)

29-2   **NESUG 2002** NorthEast SAS Users Group 15th Annual Conference  
 Site:        Adam's Mark Hotel, Buffalo NY  
 Co-Chairs:  Michael Mace, Chris Roper  
 Further Info: [www.nesug.org](http://www.nesug.org)

### 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:    TBA

### 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:        Peter Rikard at (804) 828-4878  
                   liaison@vasug.org

**SAS® Software Training Seminars**  
**Falls Church, VA**  
**November 2002**

Sponsored by NOVASUG (Northern VA SAS Users Group)

The Northern VA SAS Users Group is pleased to sponsor presentation of eleven one-day training seminars by three well-known SAS Software professionals. You can find out more about the seminars and register for them at [www.SierralInformation.com](http://www.SierralInformation.com).

These sessions will be held at the University of Virginia's Northern VA Center, conveniently located just a short walk from the West Falls Church Metrorail stations. Pay-your-own parking is also available nearby, as are several restaurants at which you can purchase your lunch.

Tuition, payable to Sierra Information Services, Inc. includes a full set of course notes, continental breakfast and afternoon refreshments.

The one-day sessions, and their instructors are:

<b><i>Date</i></b>	<b><i>Seminar Title</i></b>	<b><i>Instructor</i></b>
11/14	<b>Summarizing and Reporting Data Using the SAS System</b>	Andrew Karp Sierra Info. Services, Inc.
11/15	<b>Advanced Reporting with the SAS Output Delivery System</b>	Andrew Karp
11/18	<b>SAS System Programming Efficiencies: Tips &amp; Techniques</b>	Andrew Karp
11/18	<b>Introduction to SAS/SQL Programming</b>	Ben Cochran The Bedford Group
11/19	<b>Advanced SAS/SQL Programming</b>	Ben Cochran
11/19	<b>Introduction to SAS/STAT Software</b>	Art Carpenter Calif. Occidental Consultants
11/19	<b>Building and Applying Predictive Models with the SAS System</b>	Andrew Karp
11/20	<b>From %MACRO to %MEND: An Intro. to the SAS Macro Language</b>	Art Carpenter
11/20	<b>The SAS Output Delivery System for Data Analysts and Statisticians</b>	Andrew Karp
11/21	<b>Advanced Topics in the SAS Macro Language</b>	Art Carpenter
11/22	<b>Introduction to SAS/GRAPH Software</b>	Art Carpenter

Note SAS and all other SAS Institute product names are registered trademarks of SAS Institute, Inc. in the USA and other countries. ® indicates USA registration.





## SESUG in Savannah

If the idea of increasing your SAS knowledge in a gracious southern atmosphere appeals to you, mark your calendar now for the 9th Annual Conference of the Southeast SAS Users Group. SESUG 2002 will take place September 22-24 in Savannah, Georgia. Voted one of the "Worlds' Top Ten Trendy Travel Hot Spots" by the New York Times and named one of the "Top 10 U.S. Cities to Visit" by Condé Nast Traveler. Savannah captivates, intrigues, surprises and romances millions of visitors each year with her gracious hospitality and subtle charm (according to [www.savannah-visit.com](http://www.savannah-visit.com)). David Maddox and Heidi Markovitz are the conference co-chairs.

The conference offers 10 different paper presentations, as well as weekend workshops, and the conference leaders are now actively seeking people to present papers on almost any topic related to the usage of SAS software. The deadline for submitting paper abstracts is March 31. More details about the abstract presentation process can be found on the web at [www.sesug.org/SESUG02](http://www.sesug.org/SESUG02).



The fifteenth annual NorthEast SAS Users' Group conference will be held at the beautiful Adam's Mark in Buffalo, New York –home of the Buffalo Bills, Buffalo Bisons (baseball), and of course, the ORIGINAL buffalo wings and beef-on-weck. Served by more than ten airlines, with free shuttle service, and on-site car rental, the conference site is within two hours of all NESUG participants, as well as our friendly neighbors north of the border, eh !

Buffalo, the "twenty-minute city", is the gateway to the Niagara region, featuring world-class arts and culture, architecture, sports, wineries, dining, and great autumn weather. Come visit Lake Erie, Lake Ontario, the Niagara River and Falls, the historic Erie Canal, Buffalo and Erie County Naval and Military Park, Pedaling History Bicycle Museum, Old Fort Niagara, Original American Kazoo Company, Our Lady of Victory Basilica and National Shrine, Toy Town Museum, Chautauqua, and much, much more. In addition, Buffalo is just a short drive away from our Canadian neighbors, including the Shaw Festival, Niagara wine country, Casino Niagara, and the international metropolis of Toronto.

In addition to the opening night celebration, come partake in all the activities that have made NESUG one of the best valued conferences: SAS Demo Area, Presentations by SAS, Vendor Pavilion of service and solution providers, pre- and post- conference training, meeting and sharing with fellow SAS users. Planned concurrent sessions for Monday through mid-day Wednesday include: Advanced and Beginning Tutorials, Applications Development, Code Clinic, Coders' Corner, Emerging Technologies, Graphics & Information Visualization, Hands-On-Workshops, Posters, Statistics, Data Analysis, and Econometrics, Training and User Support.

These sessions will optionally follow the following tracks:

Beginner, Financial, Government, Management, Pharmaceutical, and Web

NESUG ([www.nesug.org](http://www.nesug.org)) has thrived for fifteen years because of your faithful attendance and the volunteer efforts of folks like you; serving as session coordinators, working the registration desk, assisting as Code Clinicians .....

Please contact the conference co-chairs for further information on how you can help continue the tradition.

<b>Michael Mace</b>	<b>Chris Roper</b>	<a href="http://www.nesug.org">www.nesug.org</a>
---------------------	--------------------	--

nesug15@att.net  
michael.a.mace@att.net  
330-796-3981

croper@ncqa.org  
202-955-3562

[www.buffalocvb.com](http://www.buffalocvb.com)  
[www.adamsmark.com/buffalo](http://www.adamsmark.com/buffalo)

## 2002 Meetings

Plan ahead and mark your calendars now. The DCSUG will meet on the following date in 2002: December 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 2002. 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

Washington DC SAS Users Group  
P.O. Box 44670  
Washington, DC 20026-4670