



# THE SESUG INFORMANT

## INSIDE THIS ISSUE:

<i>What's Wrong With "IF"?</i>	2
Ian Whitlock	
<i>Graphical Excellence or Chartjunk?</i>	6
Archer Gravely	
<i>SESUG 2K</i>	10
<i>Copying SAS Data Sets and Indexes</i>	12
Kirk Lafler	
<i>A Review of "Cody's Data Cleaning Techniques Using SAS Software"</i>	16
Andrew T. Kuligowski	
<i>From the Institute</i>	17
<i>SAS Boot Camp War Games</i>	20
Randy C. Finch	
<i>Newsletter Sponsors</i>	20

The SESUG Informant; Spring 2000; Volume 2, Issue 1

©2000 SouthEast SAS Users Group, Inc.

The SESUG Informant is a publication of the SouthEast SAS Users Group, Inc. It is published twice a year; in the spring before the SAS Users Group International (SUGI) conference and again in the fall before the SESUG conference.

If you would like to help out by advertising or writing articles for this newsletter, please contact the editor, Randy Finch, by calling 256-386-2197 or E-mailing rcfinch@tva.gov.

## PRESIDENTIAL NOTE

Well, I now follow my esteemed predecessor, David Riba, in penning this column. In the year past we had our seventh annual conference, SESUG '99, in Mobile, AL. Conference co-chairs were Greg Barnes Nelson and George Matthews. We saw some excellent presentations in Mobile, and Greg and George had some creative solutions to a number of the seemingly routine problems that one encounters in organizing and running a conference. In addition to the published conference *Proceedings*, a CD is also available. This is the first time we have done this and we thank Chris Roper and Qualex Consulting Services for making this available. As always, we also thank SAS Institute for all the work they do to make the printed Proceedings available.

In the new year, we look forward to SESUG 2000 to be held in Charlotte, NC and co-chaired by Maribeth Johnson and Philip d'Almada. Maribeth and Phil are hard at work now on SESUG 2000; their conference planning team is in place and already at work as I write this. Knowing the conference co-chairs, as well as many of those who will be working with them, I am looking

forward to this year's conference. I know it will be great; and I expect to see a lot of the new SAS v8 features presented. If you will be moving on to version 8 and want to get the most out of it, I think you will want to see what others have already found. More information about Charlotte and SESUG 2000 is available elsewhere in this issue. As a final note, I always suggest attendees *participate* in the conference they attend. I believe they have a far richer experience, and leave with a greater knowledge for having done so.

You will be interested to know that in 2001, the SESUG conference will be replaced (this one time) by the SSU Conference. The SouthEast SAS Users Group (SESUG) and the South Central SAS Users Group (SCSUG) have taken the precedent-breaking step of joining together to put on the "Southern SAS Users Conference". The

(Continued on page 15)



F. Joseph Kelley

## EDITORIAL COMMENT

Here we GROW again! Don't you just hate that line? Back when it was new, it was kind of cute, but now it's just annoying. However, it does describe this newsletter. Our first issue, Spring 1999, was a mere 12 pages. Our second issue, Fall 1999, grew to 16 pages. This issue marks our third, and it has grown to 20 pages. If this keeps up, we'll have to convert to a magazine format by the year 2005.

Well, anyway, welcome to our newslet-

ter. As usual, we have some interesting articles for you. Check out our regular contributor, Ian Whitlock, talking about how horrific IF statements are (at least for some applications). Are your charts and graphs ready for the junkyard? Archer



Randy C. Finch

(Continued on page 4)

# Tech Tip



Ian Whitlock

*Note how it is hard from the code to see what looks so simple in the diagram. This is an indication that IF statements are the wrong tool for the task.*

*It takes less code, and it is easier for a human to understand and to insure the correctness of the code.*

## WHAT'S WRONG WITH "IF"?

What is wrong with the IF statement? It is a very powerful tool in the hands of a good programmer, but it is often asked to do lowly tasks like handling conditional assignment statements and consequently obscuring important coding issues. Consider assigning a color to each unit square in a 3 by 3 square. The diagram below specifies how to assign COLOR based on two variables X, designating the row, and Y designating the column. X and Y may have values 1, 2, or 3.

RED	RED	YELLOW
RED	BLUE	YELLOW
GREEN	BLUE	BLUE

For example the square indicated by X=3 and Y=1 is to be green, and the square indicated by X=2 and Y=3 is to be yellow.

To solve the assignment problem, an IF oriented programmer would write code like:

```
if x = 1 and y <= 2 then
  color = "RED";
else
  if x=1 and y=3 then
    color="YELLOW" ;
  else
    ....
```

or

```
if x = 1 then do ;
  if y <= 2 then color = "RED" ;
  else color = "YELLOW" ;
end ;
else
  ....
```

Note how it is hard from the code to see what looks so simple in the diagram. This is an indication that IF statements are the wrong tool for the task. Consider trying to change the code when the color specification changes for some square, or verifying that the code accurately reflects the specifications. Again, these thoughts, should lead you to conclude that there ought to be a better way.

Arrays can act like complex switches, i. e. as sequence of IF - THEN - ELSE statements. An array oriented programmer

might approach the above problem using code like

```
array c ( 3 , 3 ) $_temporary_
( "RED" "RED" "YELLOW"
  "RED" "BLUE" "YELLOW"
  "GREEN" "BLUE" "YELLOW" ) ;
if ( 1 <= x <= 3 ) and ( 1 <= y <=
  3 ) then color = c ( x , y ) ;
```

Why is this better? It takes less code, and it is easier for a human to understand and to insure the correctness of the code. It is also easier to remove the color specification from the code. For example, (Code should be responsible for the work, not the messy detailed specifications.)

```
array c ( 3 , 3 ) $_temporary_;
if _n_ = 1 then do ;
  infile specs missover ;
  input i j color $ ;
  c ( i , j ) = color ;
end ;
if ( 1 <= x <= 3 ) and ( 1 <= y <=
  3 ) then color = c ( x , y ) ;
```

Now the specifications for COLOR reside in a data file where they can be easily modified without touching the code. Even a non-programmer can look at a print of the specification file and verify that it accurately meets the requirements or fails to do so.

So where do formats come into this picture? The variables X and Y often don't have convenient values like 1, 2, 3. Suppose instead they took on character values, say "FIRST", "SECOND" and "THIRD" (any 3 words or character values would do). Do we have to give up the convenience of arrays that we have just learned? No, INVALUE informats can be used to convert character values to numbers.

```
proc format ;
  invalue trans
    "FIRST" = 1
    "SECOND" = 2
    "THIRD" = 3 ;
run ;
```

The assignment of color then is

*(Continued on page 4)*

DAILY COMPOSITE EXCHANGE TRANSACTIONS

Q

Table with multiple columns containing financial data, likely representing exchange transactions. The table is partially obscured by a large advertisement.

**66 OUT OF 68 OF THE WORLD'S TOP EXCHANGES WILL TRADE ANYTHING. EXCEPT THEIR COMPUTERS.**

Up, down, up, down. Okay for markets. Not okay for their computer systems. So in an overwhelming 66 of 68 cases, their computer systems are Compaq NonStop™ eBusiness Solutions. The same kind trusted with 75% of all ATM transactions, 66% of all credit card sales, 60% of all funds transfers. All from Compaq, the people who know 24x7 computing best. Including the fact that it's not just for high finance anymore. Got an e-business? If it stops, for even a blink, you lose. Not just money, but customers. Maybe forever. Want to avoid that? Maybe you should trade up. For more: [www.compaq.com/NonStop](http://www.compaq.com/NonStop).

**24x7x COMPAQ**

COMPAQ NonStop

Table with multiple columns containing financial data, likely representing exchange transactions. This section is visible below the advertisement.

# Tech Tip (continued)

*So what is wrong with IF statements? They can easily block one from developing good easily maintained data driven programs.*

(Continued from page 2)

```
i = invalue ( x , trans. ) ;
j = invalue ( y , trans. ) ;
if (1 <= i <= 3) and (1 <= j <=
  3) then color = c ( i , j ) ;
```

Or perhaps it is time to do away with the array. We could make the specification file look like:

```
FIRST FIRST RED
FIRST SECOND RED
FIRST THIRD YELLOW
....
```

Now one could code

```
data specs ;
  retain fmtname "$trans" ;
  infile specs ;
  input x $ y $ label $ ;
  start = trim(x) || "," || y ;
run ;
proc format cntlin = specs ;
run ;
```

Now there is no array and the assignment code is simply:

```
color = put ( trim ( x ) || ", "
```

```
|| y , $trans. ) ;
```

So what is wrong with IF statements? They can easily block one from developing good easily maintained data driven programs. They are the lazy programmer's answer how to postpone thinking. But, in fact, they seduce him into working harder to debug and maintain his program.

# Editor (continued)

(Continued from page 1)

Gravely lets you know. Having trouble copying your indexes along with your data sets? Check out Kirk Lafler's quick tip. Got data that will not come clean even with the strongest detergent and heavy scrubbing? Maybe Ron Cody's new book is for you. Andy Kuligowski's reviews it for you. Also, SAS Institute provides some useful information including a complete list of names and contacts for all local user groups in the Southeast region. A preview of SESUG 2K is in the middle. Also, don't forget to check out our sponsors' ads and let them know you appreciate their support of SESUG.

finding it all in  
**SAS® SOFTWARE**  
training & consulting

**destiny**

**development**

customized applications utilizing the latest technology

**experience**

providers of SAS software training and consulting for over 10 years

highly skilled trainers & consultants

**service**

instant response

competitive pricing

free course customization

**training**

over 40 courses offered

portable training room

available

**interactive**

individual attention

one machine per student

hands-on training workshops

**need based**

we listen to your needs

we provide solutions

**you are invited to call**

860-721-1684 or

1-800-7TRAINING

fax: 860-721-9784

e-mail: [info@destinycorp.com](mailto:info@destinycorp.com)

visit: <http://www.destinycorp.com>

**destiny—it's all there.**

Destiny Corporation is a SAS Institute Quality Partner



**Qualex Consulting Services, Inc.** has the experience and the resources to harvest the knowledge contained within the vast amount of data generated by your company's web traffic.

Consulting exclusively in the SAS® System, our consultants have many years of SAS application experience applying the most demanding of information delivery disciplines for clients such as:

**SAS Institute**  
**General Motors**  
**Ernst & Young**  
**Philip Morris**  
**Washington Post**  
**LTV Steel**  
**Washington Post**

**Blue Cross/Blue Shield of Florida**  
**National Committee for Quality Assurance (NCQA)**  
**Office of the Comptroller of the Currency (OCC)**  
**First Union Bank of North Carolina**  
**Oppenheimer Financial Services**  
**Federal Reserve Board**

Visit our web site for a wide assortment of free SAS Tools, explanation of our services, and much much more! <http://www.qlx.com>

Qualex Consulting Services, Inc.  
 382 Fox Chase Drive, Collinsville, VA 24078  
 Phone: 919-380-8284 / e-mail [qualex@qlx.com](mailto:qualex@qlx.com)

# How To



**Archer Gravely**

*The analyst has three main tools to convey the message of his or her work: words, tables, and graphs.*

*It is an unfortunate fact that ugly, vibrating statistical graphics are all too common in the popular press, professional literature, and unpublished research reports.*

## GRAPHICAL EXCELLENCE OR CHARTJUNK?

It's interesting to note that we have very clear standards for proper English usage, but in the presentation of quantitative information in tables and graphs such standards are seldom understood and are often violated. Edward Tufte in his landmark work *The Visual Display of Quantitative Information* (1983) articulates a number of principles which enable users to create presentations that communicate complex ideas with clarity, precision, and efficiency while avoiding the pitfalls of *chartjunk*, graphical decoration, unintentional "optical art", and other forms of bad graphical design.

### Tables vs. Graphs

The analyst has three main tools to convey the message of his or her work: words, tables, and graphs. Like a woodworker's chisels, saws, and planes, these tools are best used to perform distinct functions, but all are usually required in combination to produce a quality piece of work. Some general guidelines for the effective use of tables, graphs, and words are:

- Tables are the media of choice for conveying numerical *values*
- Graphs are best suited for depicting *relationships* among variables
- When small numerical differences have important implications, graphs tend to do a poor job of communicating the findings
- Tables do a much better job than words in comparing more than two numbers
- For reporting small data sets, tables are more effective than graphs
- Sentences are used to *summarize* information

### Principles of Graphical Excellence

Tufte's (1983, p.51) work on the presentation of quantitative information provides a compelling set of principles to de-

fine the concept of graphical excellence:

- Well-designed presentation of interesting data - a matter of substance, of statistics, and of design
- Consists of complex ideas communicated with clarity, precision, and efficiency
- Gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
- Nearly always multivariate
- Requires telling the truth about the data

Tufte in his most recent work (1997) notes that "...clarity and excellence in thinking is very much like clarity and excellence in the display of data. When principles of design replicate principles of thought, the act of arranging information becomes an act of insight."

### Avoid Chartjunk and Graphical Decoration

It is an unfortunate fact that ugly, vibrating statistical graphics are all too common in the popular press, professional literature, and unpublished research reports. Tufte (1983, p. 107) coined the word *chartjunk* to describe poorly designed graphics that have the following characteristics:

- Use an overwhelming amount of ink to describe a few numbers
- Print label information vertically
- Use cross-hatching patterns that vibrate
- Attempt to describe differences with indecipherable graphical elements (e. g., pie charts, block charts, star charts)

Tufte makes the humorous point that chartjunk is often a form of "unintentional optical art". Many chartjunk applications arise from the naive "look what I did with a

*(Continued on page 8)*

## COME SEE US AT SUGI ON APRIL 9-12 IN INDIANAPOLIS

Save **15%** on SAS Institute books while at the conference.\*

# SAS® Publications

## We Get the Word Out Worldwide!



### Visit the Publications Booth to learn more about:

- **New Books**
- **Books by Users™** - users wrote them, users love them
- **SAS Online Tutor™** - web-based training contains 35 lessons that teach how to use version 8 SAS environment
- **SAS® Certified Professional Training Guide, V6** - contains instruction designed to prepare you for the SAS Certified Professional V6 Exam (A00-001)

SAS Institute's Publications Division offers you a comprehensive library of documentation for each component of the SAS Software System. Stop by the Publications Booth, place your order and receive a **15% discount** on your SAS documentation purchase. SAS Institute staff will be on hand to answer any questions and to assist you with your order.



**Enjoy the conference!**

SAS and all other SAS Institute Inc. product and service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates a USA registration.  
\* Certain restrictions apply. Please note that this offer does not apply to SAS Institute specialty items, SelectText, or the video-based training courses. No other discounts apply.  
32029.0100

# How To (continued)

*Without exception, the designers of graphical and presentation software, have made it very easy for unwary users to produce chartjunk based on program defaults or easy to use procedures that exemplify bad graphical design.*

*Based on the principles of graphical excellence, pie charts, block charts, and star charts are not recommended for presenting quantitative data.*

(Continued from page 6)

computer” syndrome rather than a well thought-out concept of the data presentation. Without exception, the designers of graphical and presentation software (SAS included), have made it very easy for unwary users to produce chartjunk based on program defaults (e.g., cross-hatched patterns) or easy to use procedures (e.g., pie charts, three-dimensional charts, block charts, bubble charts, etc.) that exemplify bad graphical design. Tufte (1983) provides us with a checklist of characteristics to describe *friendly* vs. *unfriendly* graphics.

**Figure 1: Characteristics of User-Friendly vs. Unfriendly Graphics**

Friendly	Unfriendly
Spell out words - avoid cryptic labels	Many abbreviations; require unnecessary reader effort to understand the intended message
Labels are printed left to right	Vertical labels; words in several directions
Include messages to help reader understand data	Cryptic graphic information
Labels are placed on graphic so reader can instantly understand graphic elements	Make reader go back and forth between legend and graphic. Use cross-hatching and garish colors
Graphic entices reader - arouses curiosity	Graphic is chartjunk
Colors (if used) are sensitive to color blind readers. Blue can be distinguished from almost all other colors by most color deficient people	Use red and green for side by side contrasts
Type is easily readable and not overbearing	Type is clotted and overbearing
Type is upper/lower case, with serifs	Type is all upper case, sans serif

Based on the principles of graphical excellence, pie charts, block charts, and star charts are *not* recommended for presenting quantitative data. Bertin (1981) and Tufte (1983, p. 178) argue convincingly that tables are far superior to *dumb* pie charts and the only thing worse than a single pie chart is a group of pie charts where:

... the viewer is asked to compare quantities located in spatial disar-

ray both within and between pies.. given their low data-density and failure to order numbers along a visual dimension, pie-charts should *never* be used.

A more complete treatment of this topic is provided in Gravely's 1998 *Your Guide to Survey Research Using the SAS System* published by SAS Institute and the 1999 SUGI paper "Achieving Graphical Excellence With SAS/GRAPH".

## References

Bertin, Jacques. *Graphics and Graphic Information Processing*. Berlin; New York: de Gruyter, 1981.

Gravely, Archer R. *Your Guide to Survey Research Using the SAS System*. Cary, NC: SAS Institute Inc., 1998.

Mahon, B.H., "Statistics and Decisions: The Importance of Communication and the Power of Graphical Presentation". *Journal of the Royal Statistical Society (A)*. 1977, **140** pp. 298-323.

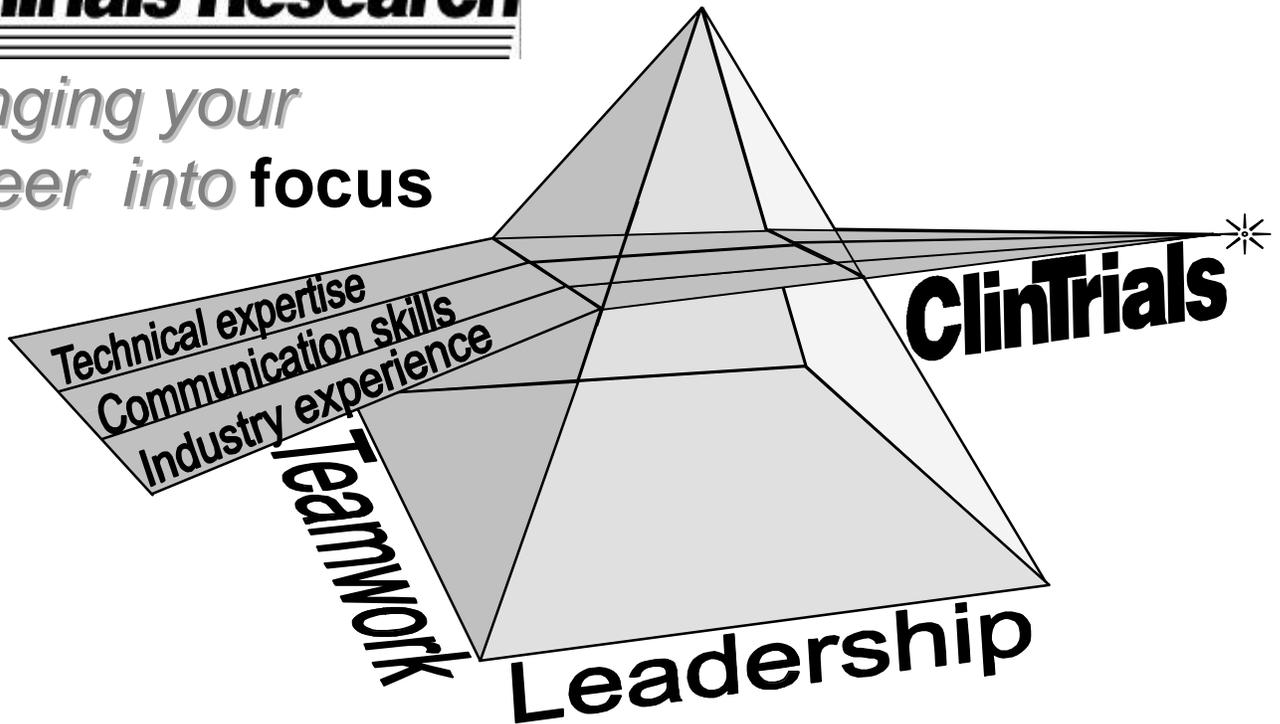
Tufte, Edward R. *The Visual Display of Quantitative Information*, Cheshire, Conn: Graphics Press, 1983.

Tufte, Edward R. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Conn: Graphics Press, 1997.

Archer R. Gravely  
Office of Institutional Research -CPO 1460  
UNC Asheville  
One University Heights  
Asheville, NC 28804-8503  
email: gravely@unca.edu  
phone: (828) 232-5118

# ClinTrials Research

Bringing your career into focus



ClinTrials Research is actively seeking experienced SAS programmers to lead our clinical trial programming teams. These teams are responsible for providing the statistical results of clinical trials, and are active throughout the clinical trial process.

ClinTrials Research is a leading research organization working with top pharmaceutical and biotechnology companies. We offer a very competitive salary and benefits package, and look forward to hearing from you. Please reference job code #HireMe4SAS.



## Harbourside Travel Services

“A Full Service Travel Agency”

- ⇒ Airline Tickets
- ⇒ Rental Cars
- ⇒ Hotels
- ⇒ Management Reports
- ⇒ Leisure Packages

Providing 7 years service for the South-east SAS Users Group

1-800-330-8646  
 727-724-9094  
 727-724-9247 Fax  
[www.harboursidetravel.com](http://www.harboursidetravel.com)

Human Resources  
 PO Box 13991  
 Research Triangle Park, NC 27709  
 Voice: (919) 462-2224  
 Fax: (919) 469-8534  
 E-Mail: [hr@clintrialsresearch.com](mailto:hr@clintrialsresearch.com)  
[www.clintrialsresearch.com](http://www.clintrialsresearch.com)



*SESUG 2K -- Charlotte, NC  
 October 15-17, 2000  
 It's going to be Grand...2 Grand!*

Had you heard of Charlotte 10 to 15 years ago? What a difference a few years can make!

Plan to come and check things out at the 8<sup>th</sup> annual SESUG conference at the Adam's Mark Hotel.



Charlotte is now known across the country as:

- a sports Mecca, with the NFL Carolina Panthers, NBA Charlotte Hornets, and NASCAR racing leading the way.
- a center of the performing arts with music, visual arts, and dance all in top form.
- a dynamic center of finance and commerce, second only to New York City.

Named for Queen Charlotte, the city takes pride in treating every visitor as a royal guest.

Find out more about Charlotte at the Convention and Visitors Bureau web site  
[www.charlottecvb.org](http://www.charlottecvb.org)

*SESUG 2K  
 Sneak Preview*

	<b>Saturday</b>	<b>Sunday</b>	<b>Monday</b>	<b>Tuesday</b>
<b>Daytime</b>	<ul style="list-style-type: none"> <li>• Weekend Workshops</li> </ul>	<ul style="list-style-type: none"> <li>• CONSUG Meeting (SAS Consultants Group)</li> <li>• Weekend Workshops</li> </ul>	<ul style="list-style-type: none"> <li>• Hands on Workshops</li> <li>• Concurrent Sessions (all day)</li> <li>• Meet the Presenter Luncheon</li> <li>• SAS Institute Formal On-line Demos</li> <li>• SAS Demo Room</li> </ul>	<ul style="list-style-type: none"> <li>• Hands on Workshops</li> <li>• Concurrent Sessions (until 3pm)</li> <li>• Meet the Presenter Luncheon</li> <li>• SAS Institute Formal On-line Demos</li> <li>• Closing Session / Giveaways</li> </ul>
<b>Evening</b>		<ul style="list-style-type: none"> <li>• Hosted Dinners (various Charlotte restaurants)</li> <li>• Opening Session / Keynote Speaker</li> <li>• Social Hour</li> </ul>	<ul style="list-style-type: none"> <li>• Mixer</li> <li>• Hosted Dinners (various Charlotte restaurants)</li> </ul>	

The *SESUG 2K* Conference will provide presentations geared toward all SAS programming levels from novice to advanced. Planned concurrent sessions for Monday and Tuesday will include:

- Advanced and Beginning Tutorials
- Applications Development
- Emerging Technologies
- Coders' Corner
- Information Visualization
- Posters
- Statistics
- Systems Architecture
- Professional Development and User Support
- and more!

*Want to present at SESUG? How do I submit an idea for a paper or a poster?*

Simply send us an abstract along with contact information by April 28, 2000. You will be mailed notification in May, 2000. If accepted, a final camera-ready copy will be due by July 24, 2000, to be published in the Proceedings.

Please submit your abstract by April 28 to: <http://www.sesug.org> (or to the Conference Co-Chairs). Be sure to include the following information:

- ✓ Author contact information (address, phone, e-mail, affiliation)
- ✓ Proposed Paper or Poster title
- ✓ Abstract (max. 150 words)
- ✓ Author biography (a summary of your experience and background, max. 75 words)
- ✓ Proposed section (See the *SESUG 2K* Call for Papers and Participation)

*Want to attend?*

You should receive your registration packet in July, 2000. If you don't receive one, e-mail us and we'll see that you get one. Also, check the SESUG web site for updated information throughout the year!

*What about Corporate Sponsorships?*

If you'd like to help support the *SESUG 2K* conference, we are offering 4 levels of sponsorship again this year. Please contact Randy Finch at [rcfinch@tva.gov](mailto:rcfinch@tva.gov) for more information.

*Want to Volunteer?*

There are many opportunities for you to volunteer to help with the conference. If you're interested, please contact Clara Waterman, our Volunteer Coordinator, at [clwaterman@aol.com](mailto:clwaterman@aol.com) or the *SESUG 2K* Conference Co-Chairs.

*SESUG 2K Conference Co-Chairs & Weekend Workshop Chair*

**Operations Chair**  
 Philip J. d'Almada  
 EDS (c/o CDC)  
 1600 Clifton Road, E-45  
 Atlanta, Georgia 30333  
 (404) 639-6120  
[pxd2@cdc.gov](mailto:pxd2@cdc.gov)



**Program Chair**  
 Maribeth Johnson  
 Office of Biostatistics AE-3035  
 Medical College of Georgia  
 Augusta, Georgia 30912-4900  
 (706) 721-3785  
[maribeth@stat.mcg.edu](mailto:maribeth@stat.mcg.edu)



**Weekend Workshop Chair**  
 Dan Bruns  
 Tennessee Valley Authority  
 1101 Market St (MP2B)  
 Chattanooga, TN 37402  
 (423) 751-6430  
[debruns@tva.gov](mailto:debruns@tva.gov)



*See you in Charlotte!*

## Quick Tip



**Kirk Lafler**

*A better and more efficient approach that will achieve all of your copying objectives is to use the COPY procedure.*

## COPYING SAS DATA SETS AND INDEXES

SAS users often need to copy or backup one or more SAS data sets. There are several ways to copy a data set. Unfortunately not all methods are the same and most definitely don't produce the same results. Here's a tip that may provide some help and avoid unsuspected surprises when copying data sets that have one or more indexes associated with it.

The SAS® System provides several ways to copy a data set, but not all methods are designed to copy a data set with one or more indexes associated with it. Let's first look at a commonly used approach and one

that may not produce the desired results. The DATA and SET statements provide an easy way to copy a data set by performing a simple process of reading and writing one observation at a time to either an existing or new data set. Although this approach is convenient at times, it is not the most efficient technique for making data set copies and more importantly can be dangerous, particularly when the data set has one or more indexes associated with it. The DATA and SET statements can be used to copy the observations in one data set to another (or the same data set), but it has no way of copying indexes that are associated with the data set. All it does is copy the data minus any indexes. The only real advantage for ever using this approach is that it copies data sets within the same library or changes the name of a data set as it is being copied.

A better and more efficient approach that will achieve all of your copying objectives is to use the COPY procedure. It's purpose is to copy data sets quickly and completely with any associated indexes. Although it cannot make copies within the same library or change the name of the data set as it is copied, these limitations should not be reasons for not using this procedure for all your copying and backup needs.

The following syntax illustrates how the COPY procedure is used:

```
PROC COPY IN=libref OUT=libref MTYPE=(DATA
PROGRAM) ;
SELECT dataset1 dataset2 ... datasetn ;
```

The IN= and OUT= parameters specify the input and output locations using the assigned librefs. The MTYPE= parameter, if specified, informs the SAS System to copy only data sets (and indexes) and not any other type of file that can be in a SAS data library. Other values that can be specified for MTYPE= include CATALOG for catalogs, PROGRAM for compiled DATA steps, and VIEW for views. The SELECT statement is optional and if present tells the SAS System what data sets, or files, to copy from the input location to the output location. If SELECT is omitted, the COPY procedure copies all files, including indexes, from the input to output location.

Another way to copy data sets and indexes is by using the DATASETS procedure. The following syntax illustrates how the DATASETS procedure is used:

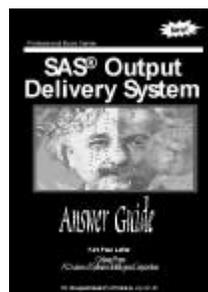
*(Continued on page 15)*

**Have an ODS question? Reach for the book with the answers! (Updated 2nd Edition)**

### SAS® Output Delivery System Answer Guide

This book is the ideal tool for getting the help you need to take full advantage of the SAS® Output Delivery System (ODS). It's clearly written and organized into specific subject areas so you can be more productive and find what you need fast. You'll get straightforward answers on how you can use ODS to produce great looking SAS output. Specific sections include:

- ✓ Top Ten FAQs
- ✓ Introduction to ODS
- ✓ Creating ODS Output
- ✓ HTML Fundamentals
- ✓ ODS and HTML
- ✓ Creating Output Data Sets
- ✓ ODS and the DATA Step
- ✓ Customizing Output
- ✓ ODS and the TEMPLATE Procedure
- ✓ Miscellaneous Questions
- ✓ And much more!



**Order your copy today! Available March 2000**

\$34.95 plus \$4.95 S/H

Software Intelligence Corporation  
P.O. Box 1390

Spring Valley, California 91799-1390  
Phone: (619) 660-2400

E-mail: [KirKLafler@CompuServe.com](mailto:KirKLafler@CompuServe.com)  
Web Site: [www.software-intelligence.com](http://www.software-intelligence.com)

SAS is the registered trademark of SAS Institute Inc., Cary, NC, USA

Statisticians/  
SAS® Programmers:



for something  
better

Smith Hanley offers a diversity of placement opportunities for statisticians and SAS® programmers.

Choose from permanent employment opportunities with leading companies worldwide, or join Smith Hanley Consulting and work at assignments with many of the same companies.



**SMITHANLEY**

New York • Chicago • Philadelphia • Southport • Orlando

Remember: Smith Hanley for something better.  
Call or email/fax your resume to:

Permanent: Jean Kim - 800.989.5627  
 Fax: 212.818.9067  
 E-mail: permanent@smithhanley.com  
 Contract: Keith Shelly - 800.684.9921  
 Fax: 407.805.3020  
 E-mail: contract@smithhanley.com

Look to  
Smith  
Hanley

*A Diversity of  
Opportunities*

**Permanent & Contract  
Placement Services**  
 Pharmaceutical  
 Marketing Science  
 Data Management  
 Research  
 Manufacturing  
 SAS® Programming

**Data Warehouse/  
Data Mart Services**

- Design
- Implementation
- Exploitation
- Training and Support
- Data Validation

*At ASG,  
our job is  
to help you with your numbers!*

For data management, biostatistics, data warehousing, data mining, web development and data systems development, we create effective solutions for complex problems. These solutions can be as varied as your needs. Our job is to help your company realize the full benefit of your technological investments. We can help you over short-term hurdles with our contract consulting, system development and/or applications development teams. With our highly skilled personnel a partnership with ASG adds up to success. For more details just call (919) 467-0505 today!



2000 Regency Parkway • Suite 355 • Cary, NC 27511  
 (919) 467-0505 • Fax (919) 467-1195 • e-mail: info@asg-inc.com  
[www.asg-inc.com](http://www.asg-inc.com)



IBM is a registered trademark of IBM Corporation in the USA and in other countries.

**SSU 2001**

**THE CONFERENCE FOR SOUTHERN SAS USERS**

Mark your calendars now for

the First Ever Joint Conference  
 of  
 the SouthEast SAS Users Group  
 and  
 the South Central SAS Users Group

New Orleans, LA

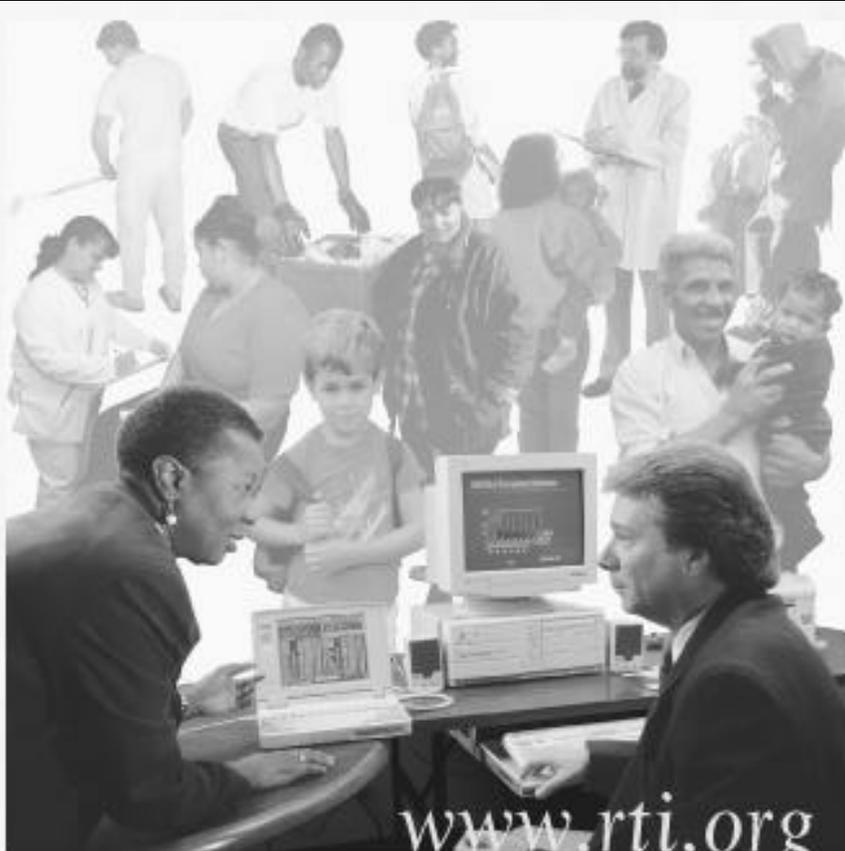
August 19-22, 2001

Co-Chairs:

Debbie Buck  
 &  
 Dave Riba



[WWW.SESUG.ORG/SSU2001/WELCOME.HTM](http://WWW.SESUG.ORG/SSU2001/WELCOME.HTM)



**RESEARCH  
TRIANGLE  
INSTITUTE**

RTI goes beyond the expected to deliver an uncommon mix of R&D and technology services. RTI provides worldwide services in research computing, statistics, health research, and surveys. We also develop innovative technologies for electronic data capture, including computer-assisted interviewing tools such as ACASI and T-ACASI.

To learn about the scope of RTI's work, visit our World Wide Web site at [www.rti.org](http://www.rti.org) or contact **Dr. Michael Samuhel** at (919) 541-5803.

[www.rti.org](http://www.rti.org)




**Abt Associates Inc.**

*Bringing the energy of ideas to research and consulting services for businesses and governments worldwide since 1965.*

***Our U.S. office locations:***

55 Wheeler Street, Cambridge, MA 02138  
Tel: (617) 492-7100; Fax: (617) 492-5219

4800 Montgomery Lane, Suite 600, Bethesda, MD 20814  
Tel: (301) 913-0500; Fax: (301) 652-3618

640 N. LaSalle Street, Suite 400, Chicago, IL 60610  
Tel: (312) 867-4000; Fax: (312) 867-4200

[www.abtassociates.com](http://www.abtassociates.com)

## Quick Tip (continued)

(Continued from page 12)

```
PROC DATASETS LIBRARY=libref MTYPE=DATA ;
  COPY OUT=libref ;
  SELECT dataset1 dataset2 ... datasetn ;
```

The LIBRARY= parameter specifies the input location using the assigned libref. The MTYPE= parameter works the same as in the COPY procedure. The COPY OUT= statement specifies the output location using the assigned libref. The SELECT statement works the same as in the COPY procedure.

If you would like more information or have any questions about this tip please contact Kirk Paul Lafler at Software Intelligence Corporation, e-mail: KirkLafler@CompuServe.com.

## Prez (continued)

(Continued from page 1)

conference will be held in New Orleans, LA, and will be co-chaired by Deborah Buck (SCSUG) and David Riba (SESUG). There has never been a joint conference of any of the regional user groups, but we are all looking forward to this, and believe it will be an outstanding conference.

Work has already begun on SESUG 2002. You will hear more about this and SSU 2001 in Charlotte this fall. I look forward to seeing you there!

If you are new to SESUG, I want to welcome you (and if you're not new, then it is good to see you again). And perhaps I should tell you a bit about what this organization is and does. We are what is called a "regional users group". We have an annual conference that is organized by and for the user community and which deals (generally) with the use of SAS Software. We provide assistance to local user groups, usually in the form of speaker sharing (one of our speakers will do a presentation at a local group's meeting, we usually limit this to one/year/group). If you want to form a users group or if we can help your current group, let me know ([jkelly@arches.uga.edu](mailto:jkelly@arches.uga.edu)). Well, it looks as though the time/space keepers are signaling me to stop, so I'll call a halt here.

**We're on the Web**  
**www.sesug.org**




Would you like to achieve  
**maximum performance**  
from your **SAS** business  
intelligence solutions?

With **SGI™** hardware  
and **SAS®** software you can!

- Improved information access makes corporate data available throughout the organization more quickly, cheaply, and efficiently:
  - The current version of SAS software, V8, is optimized for the SGI platform, offering enhanced performance and functionality
  - The most powerful client/server solution is available with the strength and depth of the SGI product family, including workstations and servers for UNIX®, Linux®, and Windows NT®
- Expanded use of systems increases the availability and use of corporate data:
  - The current release of SAS software allows more effective utilization of emerging industry technologies that optimize investments made in data warehouses and support a variety of user needs and skill sets
  - SGI servers provide unbeatable scalability, flexibility, and bandwidth, resulting in faster time to insight

For more information  
about SGI, please call  
**(800) 800-7441**  
or visit [www.sgi.com](http://www.sgi.com)



©2000 Silicon Graphics, Inc. All rights reserved. SGI and the SGI logo are trademarks of Silicon Graphics, Inc. Windows and Windows NT are registered trademarks of Microsoft Corporation. UNIX is a registered trademark in the U.S. and other countries, licensed exclusively through X/Open Company Limited. SAS and the SAS logo are registered trademarks of SAS Institute, Inc. All other trademarks mentioned herein are the property of their respective owners.

## Book Review



**Andrew T.  
Kuligowski**

*The author, Ron Cody,  
has a well-deserved  
reputation in the SAS  
world as a writer and as  
a speaker*

*Even the more  
experienced SAS  
programmer can find  
techniques that they  
might not have thought  
of independently.*

## “CODY’S DATA CLEANING TECHNIQUES USING SAS SOFTWARE” BY RON CODY

Recently, one of those prime-time television news hours devoted a segment to a man who devoted his life to finding alternate uses for everyday items. He demonstrated how peanut butter can be used in place of shaving cream, and how Alka-seltzer tablets can be used to clean toilets, just to cite a few examples.

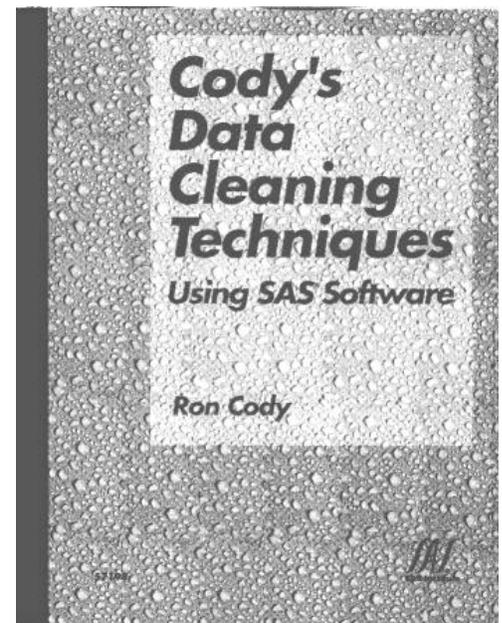
I found myself thinking about this gentleman while reading *Cody's Data Cleaning Techniques Using SAS Software*, the most recent publication from SAS Institute's Books By Users program. The author, Ron Cody, has a well-deserved reputation in the SAS world as a writer and as a speaker, with several other books and professional papers to his credit. In *Data Cleaning Techniques*, Ron uses an informal, almost chatty style to present various methods to verify the contents of input data sets. (To call him "Mr. Cody" would be a violation of the friendly atmosphere that he sets up in his book.) Ron is not content to merely show the syntax of a given command; he ensures that the reader understands WHY a particular solution was suggested and HOW it actually works.

However, if the reader only takes the material at face value, he is doing himself a great disservice. Just as the aforementioned "alternate use guru" could use a can of cola to remove dead bugs from a car windshield, *Cody's Data Cleaning Techniques Using SAS Software* can be used as a fantastic tool to learn some general SAS usage techniques! The beginning SAS programmer can find numerous examples of how to apply various DATA step tools such as formats and functions, and PROCs such as FORMAT, TABULATE, and UNIVARIATE, to real-world examples. The principles discussed in Ron's book can easily be understood and applied by those who have no interest in the stated topic of "data cleaning"; it would be a shame for someone who can benefit from the educational material to pass on it simply because of the book's

title. Even the more experienced SAS programmer can find techniques that they might not have thought of independently. And, for those who want further information about the coding techniques, Ron frequently includes references to outside materials that delve further into the current topic.

I admit that there are few things that I wish Ron had done differently. My biggest issue was in his explanation of how to use SAS to find data outliers using statistical measures such as standard deviation. The techniques that he discusses appear sound; however, Ron does not devote the same attention to explaining the meaning of the statistics he employs that he does to explaining the SAS code that drives them. Some of us non-stats types could have used a little more background in that area. (Oh, and Ron, that photo of you on the back cover has GOT to go!)

Overall, I found *Cody's Data Cleaning Techniques Using SAS Software* to be an outstanding publication, and it will be an excellent addition to my professional library.



## IMPORTANT INFORMATION FROM SAS INSTITUTE

### Institute Support Contacts

Would you believe that there is still snow on the ground at the Cary headquarters in isolated spots? Just over three weeks ago (25JAN00), North Carolina's Triangle experienced a record-breaking snowfall of 20+ inches. Consequently the Institute was closed for business 3 full days, public schools for over a week, and Institute employees were trapped in their homes. Cabin fever became an epidemic! All of us in User Group Support are now happy to be back at work, and each of us is looking forward to spring! Let us know what we can do to support you.

- SESUG Liaison - Michael Smith, 919-677-8000, ext 7462; email michael.smith@sas.com
- SESUG Local User Group Liaison - Patsy Harbour, 919-677-8000, ext 2855; email patsy.harbour@sas.com

### Plans for 2000 from SAS Institute

- User group newsletters on the Institute's WWW site
- Local user group meeting announcements on the Institute's WWW site
- The formation of a User Group Council, whose purpose is to help evaluate and improve Institute services to SAS user group conferences. SESUG's representatives to the council are Joe Kelley and Andy Kuligowski.

### SUGI 25 News

Make plans now for SUGI 25, scheduled for 09-12APR00 in Indianapolis! Conference Chair Nancy Patton has announced great new conference features for SUGI 25 such as:

- an Executive Conference for your organization's leaders.
- luncheon keynote speaker Dr. Robert Kriegel, an expert in innovative strategies for keeping ahead of the changes

and challenges in today's marketplace

- an expanded and exciting First-Timers Festival to welcome new attendees
- the SUGI Cyber Cafe with internet access for all attendees
- a SUGI museum for a look back over the last 25 years
- access to SUGI papers, via the web, prior to the conference.

You can register online and find out more about SUGI 25 by visiting [www.sas.com/usergroups/sugi/sugi25/](http://www.sas.com/usergroups/sugi/sugi25/). Remember, the last day for preregistration is 17MAR00.

### Subscribe to UGNEWS

If you'd like to receive periodic information about the User Group Support program, subscribe to the Institute's electronic newsletter, UGNEWS. Sent via email and posted on the web at <http://www.sas.com/ugnews>, the newsletter

- announces the formation of new user groups and new services from the User Group Support program
- contains features about Institute user group services
- includes reminders for upcoming user group conferences
- lists Institute and BBU speakers scheduled for upcoming user group meetings
- provides tips and news from leaders in the SAS user group community.

To start receiving UGNEWS: send e-mail to the User Group & User Publishing Support division at [sugweb@sas.com](mailto:sugweb@sas.com) saying "SUBSCRIBE UGNEWS-L" and include your full name or use the form on our web site at <http://www.sas.com/usergroups/ugnews-form.html>

## From the Institute



*Conference Chair Nancy Patton has announced great new conference features for SUGI 25*

*You can register online and find out more about SUGI 25 by visiting [www.sas.com/usergroups/sugi/sugi25/](http://www.sas.com/usergroups/sugi/sugi25/). Remember, the last day for preregistration is 17MAR00.*

# From the Institute (continued)



*The user groups are in  
alphabetical order first  
by state and then by user  
group name.*

## USER GROUP SUPPORT INFORMATION

Get involved in your local SAS user group! Listed below are the registered local user groups in the SouthEast Region. To request Institute support for your local group, contact Patsy Harbour at (919) 677-8000, ext 2855 or by email at [patsy.harbour@sas.com](mailto:patsy.harbour@sas.com). Patsy can also assist users in starting local or in-house users groups.

### BIRMINGHAM USERS GROUP FOR SAS

Maddox, Mr. David  
BellSouth Telecommunications  
South E3E1  
3535 Colonnade Parkway  
Birmingham, AL 35243  
phone: 205-977-0474  
email: [David.Maddox1@bridge.bst.bls.com](mailto:David.Maddox1@bridge.bst.bls.com)  
URL: <http://www-epi.soph.uab.edu/bugs/>

### FL GULF COAST SAS USERS GROUP

Riba, Mr. S. David  
JADE Tech, Inc.  
PO Box 4517  
Clearwater, FL 34618  
phone: 727-726-6099  
email: [dave@jadetek.com](mailto:dave@jadetek.com)  
URL: <http://www.jadetek.com/fgcsug.htm>

### NORTH FL USERS OF SAS SOFTWARE

Jordan, Mr. Brad  
Consolidated Systems, Inc.  
1977 Woodleigh Drive West  
Jacksonville, FL 32211  
phone: 904-954-7491  
email: [jordanb@mediaone.net](mailto:jordanb@mediaone.net)  
URL: <http://www.nofuss.com>

### SOUTH FLORIDA SAS USERS GROUP

Markovitz, Ms. Heidi  
Simply Systems  
161 Crandon Blvd  
Suite 325  
Key Biscayne, FL 33149  
phone: 305-365-0439  
email: [simplysyst@aol.com](mailto:simplysyst@aol.com)

### TALLAHASSEE SAS USERS GROUP

Butcher, Ms. Beth  
Information Systems of Florida  
1203 Governor's Square Blvd  
#500  
Tallahassee, FL 32301  
phone: 850-656-7600 X209  
email: [bbutcher@isf.com](mailto:bbutcher@isf.com)

### ATLANTA SAS USER GROUP

Whitehurst, Mr. Joe  
341 Eureka Dr  
Atlanta, GA 30305  
phone: 404-240-0729  
email: [joe@whitehurst-associates.com](mailto:joe@whitehurst-associates.com)

### NORTHEAST GEORGIA SAS USERS GROUP

Matthews, Mr. George  
The University of Georgia  
UCNS  
Computer Services Annex  
Athens, GA 30602-1911  
phone: 706-542-5359  
email: [gmatthew@uga.cc.uga.edu](mailto:gmatthew@uga.cc.uga.edu)

### MARYLAND SAS USER GROUP

Miller, Mr. R.H.  
6246 Sunny Spring  
Columbia, MD 21044  
phone: 410-740-4229  
email: [millerrh@juno.com](mailto:millerrh@juno.com)  
URL: <http://research.umbc.edu/chpdm/mdsug.htm>

### CHARLOTTE AREA SAS USERS GROUP

Spell, Mr. Randy  
Cadmus Direct  
1123 S. Church St., Suite 1000  
Charlotte, NC 28203  
phone: 704-344-7963  
email: [spellr@cadmus.com](mailto:spellr@cadmus.com)  
URL: <http://www.charlotteweb.com/chasug/>

**RESEARCH TRIANGLE SAS USER GROUP**

Brown, Mr. Keith  
 UNC - General Administration  
 Planning Division  
 PO Box 2688  
 Chapel Hill, NC 27515-2688  
 phone: 919-962-4547  
 email: keith\_brown@unc.edu  
 URL: <http://rtpnet.org/~rtsug/>

**TRIAD SAS USERS GROUP**

Slone, Mr. Keith  
 Wachovia  
 NC 37221  
 100 N Main Street  
 Winston-Salem, NC 27150  
 phone: 336-732-5853  
 email: keith.slone@wachovia.com

**WILMINGTON AREA SAS USERS GROUP**

McCoy, Ms. Suzanne D.  
 7816 Fisher Lane  
 Wilmington, NC 28412  
 phone: 910-452-6497  
 email: sdmccoy@bellsouth.net  
 URL: [http://members.tripod.com/~Suzanne\\_McCoy](http://members.tripod.com/~Suzanne_McCoy)

**SOUTH CAROLINA SAS USERS GROUP**

Heidari, Mr. Khosrow  
 SC Dept of Hlth & Environ Ctl  
 129 Park Shore East  
 Columbia, SC 29223-6025  
 phone: 803-898-0688  
 email: khosrow@mailexcite.com

**UPSTATE OF SC SAS USER GROUP**

Holeman, Mr. Eddie  
 Advocate Computer Services  
 Suite 301  
 1754 Woodruff Rd  
 Greenville, SC 29607  
 phone: 864-962-8234, x301  
 email: eholeman@AdvocateInc.com  
 URL: <http://www.advocateinc.com/upsug.htm>

**MEMPHIS AREA SAS USER GROUP**

Nestor, Ms. Jennifer  
 Mid-South Foundation for  
 Medical Care, Inc.  
 6401 Poplar Ave, Ste. 400  
 Memphis, TN 38119  
 phone: 901-682-0381 x334  
 email: tnpro.jnestor@sdps.org

**MIDDLE TENNESSEE SAS USERS GROUP**

Plumlee, Mr. George  
 Tennessee Dept of Health  
 Tennessee Tower 8th Floor  
 312 8th Ave N  
 Nashville, TN 37247-0340  
 phone: 615-741-1954

**NORTHERN VA SAS USER GROUP**

Mirabal, Mr. Joe  
 America On-Line Inc.  
 12100 Sunrise Valley Drive  
 Reston, VA 20191  
 phone: 703-265-4620  
 email: jmmirabal@aol.com

**VIRGINIA SAS USERS GROUP**

Rikard, Mr. Peter  
 VCU Computer Center  
 Suite 1141  
 900 E. Main Street  
 Richmond, VA 23219  
 phone: 804-828-4878  
 email: prikard@mail.ucc.vcu.edu  
 URL: <http://www.vasug.org>

**WASHINGTON DC SAS USER GROUP**

Rhoads, Mr. Mike  
 Westat  
 1650 Research Blvd.  
 Rockville, MD 20850  
 phone: 301-251-4308  
 email: rhoadsm1@westat.com  
 URL: <http://www.ita.doc.gov/industry/otea/dcsug/>

# From the Institute (continued)



*There is one user group  
 in AL, four in FL, two  
 in GA, one in MD,  
 four in NC, two in SC,  
 two in TN, two in VA,  
 and one in Washington  
 DC.*

# RUN;



Randy C. Finch

### Our Newsletter Sponsors:

Abt Associates	14
ASG, Inc.	13
ClinTrials Research	9
Compaq	3
Destiny Corp.	4
Harbourside Travel Services	9
Qualex Consulting Services	5
RTI	14
SAS Institute	7
SESUG 2K	10
SIG	15
Smith Hanley	13
Software Intelligence Corp.	12
SSU 2001	13



### SouthEast SAS Users Group

POB 8345  
Clearwater, FL 33758

## SAS BOOT CAMP WAR GAMES

“Report!” the Red Team sergeant exploded as he and an MP raced to the de-con tents. He was not a mean sergeant, but as ranking officer, he had to take action frequently. Now he needed his staff’s view of the suspected ongoing plot to disable their SAS programs. “I need a summary of what you have.”

The MP responded, “We have Jim Matthews, Mic Porter, and Philip d’Souza in custody. They insist they have no part in the plot, and we can’t build a strong case against them. Shall I release them, sir?”

“What other options do we have? Perhaps our sources were wrong. I’ve done a lot for Matthews. It’s hard to believe he’s the sort of man to betray me. And d’Souza has always been a man of high standards. And my old buddy Mic, important is his middle name. We can’t stomp men underfoot like this. Let them go!”

See if you can find the names of the 36 SAS procedures (some common, some obscure) I planted in the following story. The answers will appear in the Fall 2000 newsletter. Heck, there may be a few in there that I didn’t even intend to put there. So, if you find more, don’t be alarmed. Rules: The names appear as consecutive letters. They may be a full word, a partial word, or they may span words. Good luck!

“I’m in!” yelled a private. His display indicated he had hacked into the Green Team’s server. The staff’s calculations had been correct. The enemy’s server was protected by an encoded version of an old Fortran table of keywords. With his database (DB) loaded, he quickly issued a few SQL commands and the server’s primary app ended. This private was as stealthy as a bobcat. A log appeared onscreen charting every maneuver he had made.

The sergeant smiled slyly. “Our data set should be safe now. I need a printout of this. Send another print to Lieutenant Sanders. Do you copy?”

“YES SIR!” snapped the private.

The sergeant lifted the handset on the radio. “Data Pel Abel, this is Data Peco Python. Mission accomplished.”

The Red Team was victorious. They had passed the hardest test ever given them.

SESUG EXECUTIVE COUNCIL MEMBERS	TELEPHONE	E-MAIL
Gregory Barnes Nelson (1999 co-chair)	919-460-1174	greg.barnesnelson@statprobe.com
Dan Bruns, Vice President (1997 co-chair)	423-751-6430	debruns@tva.gov
Philip J. d’Almada (2000 co-chair)	404-639-6120	pxd2@cdc.gov
Frank C. Dilorio (1994 & 1996 co-chair)	919-942-2028	fcd1@mindspring.com
Randy C. Finch, Treasurer, Editor (1998 co-chair)	256-386-2197	rcfinch@tva.gov
Maribeth H. Johnson (2000 co-chair)	706-721-3785	maribeth@stat.mcg.edu
F. Joseph Kelley, President (1998 co-chair)	706-542-5359	jkelly@arches.uga.edu
Andrew T. Kuligowski (1997 co-chair)	727-738-3000 x6545	kuligoat@tvratings.com
Grace Lossman (1996 co-chair)	908-740-4570	grace.lossman@spcorp.com
George Matthews (1999 co-chair)	706-542-5359	gmatthew@arches.uga.edu
S. David Riba, Secretary (1993 & 2001 co-chair)	727-726-6099	dave@jadetek.com