



## HARTFORD AREA SAS® USERS GROUP

Volume 21 Issue 2

Second Quarter, May 2009

Table of Contents	
Notes from the Chairperson	2
Highlights of First Quarter 2009 Meeting	2-3
Calendar of Events	3
Announcement of Second Quarter 2009 Meeting	3-5
Continuing with Education	5-6
Tips & Techniques	6-8
Book Review	8-9
HASUG Email Distribution List	9
HASUG Sponsorship Policy & Recruiting Policy	10
HASUG Steering Committee	11
Map and Directions to Anthem BCBS	12

The HASUG Flash is available for reading on the World Wide Web at  
<http://www.hasug.org>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies

**Editor:** Barbara Moss  
**Layout:** Mahinda Yapa  
**Member Lists:** Charles Patridge  
**Web master:** Rob Krajcik

**Contributors:** Peter Prause  
Janet Stuelpner  
Charles Patridge  
Cindy Johnson

## **Chairperson's Notes** (Janet Stuelpner, SAS Institute)

This year is turning into a wonderful learning experience for all of the members of HASUG. Our next meeting will be held at Wellpoint/Anthem Blue Cross Blue Shield in North Haven. We welcome them as a host site and look forward to a wonderful schedule of speakers. Our first speaker is Tim D'Auria. He will be speaking about how Text Miner can save lives. Joining us after a long absence is Robert Virgile. Bob is best known as the person who has challenged us all at NESUG and SAS Global Forum with the SAS Programming Contest. He also is an author in the SAS Books by User Series of publications. Bob will be talking about efficiency techniques. It is sure to be a great meeting.

As a reminder, be sure to register for the meeting by 4pm on May 20. We need to provide our host with a list of attendees for security purposes. After this time, the registration will be closed. Don't miss out on this wonderful program.

We are trying something at our next meeting that has never been done before in the history of HASUG. We are one of the few user groups that does not collect any dues. We are a totally volunteer organization. In the past, we have had some very generous donors who have supplied the group with our web site, both hosting and name. We are at the crossroads right now where our funding is diminishing. We would like to stay an organization where dues are not needed. The Steering Committee has voted to hold a raffle at the next meeting to raise funds to pay for the web site and web hosting fees for the next year. Please help us out by donating generously at the

meeting. You have the chance to help a great organization and win some great prizes. We appreciate your attendance, cooperation and continuing support.

Our Google Groups list continues to grow. This is a reminder for you. If you want to get emails with messages and announcements from HASUG, you must join the HASUG Google Groups. The directions are located on the HASUG web site. Click the link and fill in the information. It is easy to become a part of the group. Don't miss anything, sign up today. We do not give out our membership list to anyone, so your information is safe with us.

As I come to the end of my term, I would like to thank everyone on the Steering Committee who made my reign as chairperson an easy one. We worked together to find host sites, put together interesting programs and keep HASUG the stellar organization that I know it can continue to be. During this time, my learning curve was steep but productive. The support of my colleagues was never ending. I couldn't have chosen to work with a better group and am glad that I have had the opportunity to do this in Connecticut. I say goodbye and thank you to all of the steering committee and the members who attend our meetings.

## **February Meeting Highlights** (Cindy Johnson, Bristol-Myers Squibb)

On Thursday, February 19 the first quarterly meeting in 2009 was held at Hartford Life in Simsbury, CT. The attendance was good and the speakers were informative yet entertaining.

*(... Continues on page 3)*

(...Continued from page 2)

Rick Langston from SAS Institute spoke to us on “*What’s New in Sas 9.2*”. He gave some history of the releases that currently exist and the major differences between versions. SAS V9.2 is unusual in that it was released in phases. The Phase I (called ‘Classic’) release has been available since 2008Q1 but the Phase 2 (called ‘Platform’) version of SAS 9.2 is being released during 2009Q2. Rick then gave us a preview of some interesting enhancements found in SAS 9.2 (Platform) such as GEODIST, a function to determine the distance between 2 places using longitude and latitude; ZIPCITY a function which returns the City name based on a zip code number or ZIPCITYDISTANCE which gives the distance between zip codes based on the USPS zip codes; and the SAS Code Analyzer which is a run-time utility to examine dependencies in a program. Rick has worked for the Institute for 20+ years. He brought his knowledge and experience to our group. Thank you, Rick.

Rob Krajcik from Bristol Myers Squibb led an interactive discussion about SAS quirks and tidbits cleverly named, “*Why does SAS run clockwise?*” There was a great deal of audience participation and Rob fully engaged the group with questions like, ‘Why does SAS use Best12. format as a default conversion from numeric to character?’. Do you know the answer? If not, check out his slides at [www.hasug.org](http://www.hasug.org) in the HASUG newsletters section. If you get a chance to hear Rob lead this interactive session, I highly recommend your participation. It was a stimulating, enjoyable talk.

## CALENDAR OF EVENTS

### Future General Meetings and Conferences:

May 21, 2009  
Wellpoint, North Haven, CT  
May 31-June3, 2009  
PharmaSUG, Portland, OR  
Aug 20, 2009  
SAS Institute, Glastonbury, CT

### HASUG Steering Committee Meetings:

May 21, 2009  
Immediately after HASUG meeting  
June 18, 2009  
Conference call 11:30-12:30  
Aug 20, 2009  
Immediately after HASUG meeting

## NEXT HASUG MEETING ANNOUNCEMENT

Our next meeting is on:  
**Thursday, May 21, 2009**  
at  
**WellPoint (Anthem BCBS)**  
**370 Bassett Road**  
**North Haven, Connecticut 06473**

**from 9:00 am to noon**

***Refreshments to be served & time for conversation beginning at 8:30.***  
***Please bring a photo ID to allow for access to the facility.***  
***Registration will be open until 4:00pm on Wednesday, May 20th.***

(...Continues on page 4)

(...Continued from page 3)

Our topics and Speakers are:

**“Using SAS Text Mining to Save Lives”**

**Timothy C. D’Auria, Creative Computing, Inc.**

What if there was a way for a computer to read clinician notes and predict patients most at risk for serious, life-threatening complications before they happen? Unstructured textual data often comprises the vast majority, upwards of 80% in cases, of information available for organizational analysis, particularly in the healthcare industry. During this presentation, we will explore how unstructured data can be incorporated into analyses alongside structured sources to predict future patient emergency room usage. We will then discuss why text mining may be a preferable alternative to structured healthcare data analysis using logistic regression and other statistical procedures. The presentation will conclude with a summary of how SAS text mining and broader data mining capabilities can ultimately be leveraged to save patient lives.

**Timothy D’Auria** leads the SAS Analytics Practice at Creative Computing, Inc., a SAS Partner. Timothy’s greatest interest is the application of data mining to improve patient care, but his work often takes him to the deepest corners of many other industries. Timothy’s contributions in the field of business data mining have been featured in leading industry publications and have been incorporated into educational programs at NYU, the University of Denver, and Cornell

University. He has led the successful development of three analysis-driven businesses from the ground-up, one of which pioneered the use of artificial intelligence in the design of hospitality forecasting technologies. Timothy received his Bachelors of Science degree in Statistics and Biology with Distinction in Research from Cornell University and is a certified Emergency Medical Technician. His recent article, “What are Analytics?” was published in the October 2008 edition of Managed Healthcare Executive, and elaborates on his vision of analytical methods within healthcare.

**“The Most Important Efficiency Techniques”**

**Bob Virgile, Robert Virgile Associates, Inc.**

What makes a program efficient? What makes one technique more important than another?

Most approaches to efficiency focus on speeding up the program. However, the programmer should consider much more than that. Is the program easy to understand and maintain? Does it require vast amounts of other resources (disk space, memory, tape drives to name a few). Does the analyst have to search through pages and pages of output to locate a few key numbers? How much time can you afford to learn and apply new techniques?

While much of this presentation will help speed up programs, it is still up to you to decide when and where it makes sense to incorporate the various techniques into your programs.

(...Continues on page 5)

(...Continued from page 4)

One worthwhile strategy is to start with a few good techniques and make them a habit. As you add to your list of good habits over time, it becomes easier and easier to automatically use better techniques when you program.

**Bob Virgile** is an independent SAS trainer and consultant, with over 25 years of experience developing and teaching SAS courses. He has presented numerous papers at SAS conferences, formerly developed problem-solving contests for NESUG and SGF, and has written two books for SAS Institute. He is the only person to be officially barred from participating in the NESUG SAS Bowl, due to superior SAS knowledge.

### **Continuing With Education...**

(Peter Prause, The Hartford)

**The SAS Institute Regional Training Center** in Glastonbury, CT is offering the following courses. Course information for these courses and many other non-local courses is available on the Internet at the support.sas.com web site. To register for a course or to get more information, phone SAS Institute at 1-800-333-7660.

SAS Macro Language 1: Essentials  
Apr 30 – May 1  
SAS Programming 1: Essentials  
May 13 – 15  
SAS Macro Language 2: Developing  
Macro Applications  
May 18 – 19  
Querying and Reporting Using SAS  
Enterprise Guide  
Jun 4 – 5

SAS Programming 2: Data  
Manipulation Techniques

Jun 10 – 12

SAS SQL 1: Essentials

Jun 25 – 26

When time and distance are an issue, SAS Institute offers Live Web classes and self-paced e-learning to help you get the training you need while accommodating your busy schedule. Call for more details.

**Destiny Corporation** is offering Instructor Assisted Virtual SAS® training courses in their Rocky Hill, CT office. A live instructor, in person, will support and answer all questions. The attendees are provided the supporting Online Virtual Course and materials. The Online Course is valid for 1 year from course dates.

Please visit the Destiny Corporation website for more training dates. To register for a course or to get more information, please contact Destiny on 1-800-7-TRAINING or [www.destinycorp.com](http://www.destinycorp.com).

### **Course Schedule**

**All courses are offered on the following dates in 2009:**

**June 10 – 12**

**August 17 – 19**

#### 2-day courses:

Introduction to Programming  
Programming III: Advanced Techniques  
Output Delivery System (ODS) Basics  
Version 9 Changes and Enhancements  
Enterprise Guide Querying and  
Reporting  
SQL Processing with SAS® Software

(...Continues on page 6)

(..Continued from page 5)

### 3-day courses:

Programming I SAS® Essentials  
Programming II: Data Manipulation  
Using the Data Step  
Report Writing: A Programming  
Approach  
Macros in SAS® Software

## Tips & Techniques

(Charles Patridge, ISO Innovative Analytics)

This macro takes 2 SAS datasets, runs a ONE-WAY Frequency on a list of vars common to both datasets, compares the frequencies of these variables, and produces a proc print of the distributions with their differences.

### Sample Call:

```
%procfreq_compare(  
    base= data1 ,  
    compare=data2 ,  
    _vars_=aa bb cc dd,  
    format=7.1,  
    cleanup=1);
```

### Parameters:

BASE [REQUIRED] - name of  
BASE sas dataset with the lib  
reference.  
COMPARE [REQUIRED] - name of  
COMPARE sas dataset with the lib  
reference.  
\_VARS\_ [REQUIRED] - a list of  
SAS variables common to both  
BASE and COMPARE sasdatasets,  
separated by blanks.  
XWHERE [OPTIONAL] - a valid  
WHERE clause to subset both BASE  
and COMPARE datasets.  
WEIGHT [OPTIONAL] - a sas  
variable common to both BASE and  
COMPARE datasets to weight the  
ONE-WAY Frequencies.  
FORMAT [OPTIONAL] - a valid  
format applied to all variables  
listed in \_VARS\_. Default is  
best.2 for numeric variables.

CLEANUP [OPTIONAL] - 1 will  
cleanup all work files not  
needed after comparison is  
completed. Any other value will  
keep work files.

### Notes:

1. The SAS Lib References need to be issued before calling this macro.
2. Do not mix character and numeric variables in the same list of \_VARS\_.
3. Need to choose a format appropriate for all variables listed in \_VARS\_.

```
%macro procfreq_compare ( base=,  
compare=, _vars_=, xwhere=,  
weight=, format=best.2,  
cleanup=1);
```

```
%if %length(&_vars_) gt 0 %then  
%do;  
%let _vars_ = %cmpres(&_vars_);  
%global howmany_vars;
```

```
data _null_;  
howmany = 1 +  
(length(compbl("&_vars_")) -  
length(compress("&_vars_")));  
call symputx('howmany_vars',  
put(howmany,3.));  
run;  
%end;
```

```
/**/ FREQ on BASE or COMPARE  
Table ***/
```

### %macro

```
run_freq(Base_Compare=BASE);
```

```
proc freq data=%%&&Base_Compare;  
%if %length(&xwhere) %then %do;  
    where &xwhere;  
title2 "Where Clause = &xwhere";  
%end;
```

```
%do _i_ = 1 %to &howmany_vars;  
table %Scan(&_vars_, &_i_,%STR(  
) / missing outpct noprint  
out=&Base_Compare.&_i_;
```

(...Continues on page 7)

(..Continued from page 6)

```
format %sysfunc( translate(
%Scan(&vars_, &i_,%STR( )), '
', '*' ) ) &format ;
%end;
%if %length(&weight) gt 0 %then
%do;
weight &weight;
%end;

%mend run_freq;

%run_freq(BASE_COMPARE=BASE);

%run_freq(BASE_COMPARE=COMPARE);

%macro Join_Base_Compare;

%do _i_ = 1 %to &howmany_vars;
proc sort data=base&i_
out=base&i_ ; by
%Scan(&vars_, &i_,%STR( ));
run;

proc sort data=compare&i_
out=compare&i_; by
%Scan(&vars_, &i_,%STR( ));
run;

data base&i_;
set base&i_;
%Scan(&vars_, &i_,%STR( )) =
input( put( %Scan(&vars_,
&i_,%STR( )), &format),
&format);
run;

data compare&i_;
set compare&i_;
%Scan(&vars_, &i_,%STR( )) =
input( put( %Scan(&vars_,
&i_,%STR( )), &format),
&format);
run;

data Diff&i_;
length var_name $32.;
label var_values =
"%Scan(&vars_, &i_,%STR( ))";

merge base&i_ (rename= (
count=base_cnt
percent=base_pct ) )
compare&i_ (rename= (
```

```
count=compare_cnt
percent=compare_pct) );

by %Scan(&vars_, &i_,%STR( ));

var_name = "%Scan(&vars_,
&i_,%STR( ))";
var_values = %Scan(&vars_,
&i_,%STR( ));

label
base_cnt = "Base Count"
base_pct = "Base Percent"
compare_cnt = "Compare Count"
compare_pct = "Compare Percent"
diff_cnt = "Difference Count"
diff_pct = "Difference
Percent" ;

if base_cnt = . then base_cnt =
0;
if compare_cnt = . then
compare_cnt = 0;
diff_cnt = compare_cnt -
base_cnt;

if base_pct = . then base_pct
= 0;
if compare_pct = . then
compare_pct = 0;
diff_pct = compare_pct -
base_pct;

drop %Scan(&vars_, &i_,%STR(
));
run;
%end;

/** join the difference files
into a single file
delete _diff_ if it exists */

proc datasets library=work
nolist; delete _diff_; quit;
%do _i_ = 1 %to &howmany_vars;
proc append base=_diff_
data=diff&i_;
run;

%if &cleanup = 1 %then %do;

/** delete the individual base
and compare files to clean up
***/
(...Continues on page 8)
```

(..Continued from page 7)

```
proc datasets library=work
nolist; delete base&i_
compare&i_; quit;
    %end;
    %end;

proc sort data=_diff_
out=_diff_; by var_name; run;

title1 "Base File = &base";
title2 "Compare File =
&compare";

proc print data=_diff_;
by var_name;
id var_name;
pageby var_name;
run;
%mend Join_Base_Compare;

%Join_Base_Compare;

run;
%mend procfreq_Compare ;
```

## Book Review

(Charles Patridge, ISO Innovative Analytics)

**Title:** *Just Enough SAS®: A Quick-Start Guide to SAS® for Engineers*

**Author:** Robert Rutledge

**Publisher:** SAS Press

**Pages:** ~400 pages

**Availability Date:** Spring 2009

**Chapter 1 – Getting Started**

**Chapter 2 – Data Step Programming**

**Chapter 3 – Data Out, Data In -  
Spreadsheets**

**Chapter 4 - Data Out, Data In – Relational  
Databases**

**Chapter 5 – Summarizing Your Data**

**Chapter 6 – Plotting Your Data with  
SAS/GRAPH**

**Chapter 7 – The Output Delivery System**

**Chapter 8 – Plotting Your Data with ODS  
Graphics**

**Chapter 9 – Analyzing Quality Data**

**Chapter 10 – Analyzing Reliability Data**

**Chapter 11 – Macro Programming**

So, do you have to be an Engineer to want/read this book??? No!

This is an excellent book/guide (over 400 pages) to getting started with the SAS Software when you want to analyze your data – no matter where your data resides. This is the first book by users I have had the privilege to review where such a broad spectrum of SAS topics are discussed and illustrated for people just getting started with SAS.

The book discusses various ways to get your data into SAS, be it from Excel, text files, or databases. It also shows you how to push your data from SAS into a variety of outputs such as Excel, databases, HTML, and RTF using the ODS facility.

I also liked Rutledge's creative way of comparing various ways of doing certain routines – such as using SQL vs Datastep, SQL vs Proc Means, etc. This I hope will spark creativity within the reader to ponder over other ways of doing a particular task within SAS.

Finally, I like the idea of introducing the reader to SAS macros and how to use them to make SAS applications /programs more robust. The examples are simple and yet easy to follow, and I hope will encourage more SAS professionals into adopting and learning the SAS macro facility as part of their tools of the trade.

I would highly recommend this book for any person – beginner to expert (there are some tidbits which I discovered and did not know).

(...Continues on page 9)

(..Continued from page 8)

The fact this book covers a variety of SAS Products such as Base SAS, SAS/Stat and SAS Graph provides a nice way to see what these products have to offer if you do not currently license them.

There are plenty of examples and codes to copy and try out for yourself. I would not use this book necessarily as a reference but it definitely would encourage me to look more closely at some online documentation for those examples I found useful within the book.

## **HASUG EMAIL DISTRIBUTION LIST**

This service is provided free of charge to all who wish to subscribe to the email notification process. In order to sign up for this service, you need to point your browser to the HASUG web site, <http://www.hasug.org>, and click the hyperlink, "Join HASUG Email List".

Your name and email address are not displayed nor given out to any party, including members of the HASUG steering committee. Only the list administrators have access to subscription information. Only List administrators will be able to send out email messages to the HASUG members.

If you should change your email address under which you subscribed, you will need to unsubscribe with your old email address, and then re-subscribe using your new email address. Typically, a

notice is sent approximately 3 weeks before meetings, as well as a couple of days beforehand to remind people of the date, time, and location.

To subscribe to the HASUG list you will need to follow the following steps:

1. From the HASUG website pick "Join HASUG Email List". Enter your email address and click on the SUBSCRIBE button. You will receive an email from [oreply@googlegroups.com](mailto:oreply@googlegroups.com).
2. Open the email and click on the confirmation URL in the email.
3. There are no questions to answer. Just click on "Apply to this Group" .
4. You will receive an email saying you have been added to the distribution list.

You must complete steps 1-3 to be added to the list.

## **THE HASUG RECRUITING POLICY**

HASUG's primary mission is to provide a forum for SAS professionals to meet and share experiences. HASUG also recognizes that searching for new employment opportunities is a normal activity when professionals meet. Since we depend on our respective employers to support HASUG by providing time for members to attend, prepare presentations, and provide meeting facilities, we do not wish to jeopardize those relationships we have all nurtured.

Therefore, when attending any of our quarterly meetings, we request that all parties engaged in recruiting activities be "professional and discreet", and suggest that such activities be carried outside and after our meetings. Those seeking or offering positions may display materials at a location designated by HASUG at each quarterly meeting. The materials may include resumes, job openings, contract opportunities, business cards, etc. None of the materials may include salary information. All such materials left after the meeting will be discarded. Violators of this policy may be asked to leave the meeting.

## **THE HASUG SPONSORSHIP POLICY**

All parties are encouraged and welcome to attend our quarterly HASUG meetings scheduled during the months of February, May, August, and November of each year. Usually these meetings are scheduled the 3rd Thursday from 9:00 am to noon at various locations throughout Connecticut. Past locations have included Hartford, Meriden, Middletown, Norwalk, Glastonbury, Stamford and New Haven. We are always looking for new hosts within the state of Connecticut. If you feel your organization can host a meeting of up to 70 or so people, please contact one of the Steering committee members.

## HASUG Steering Committee

<b><u>Name</u></b>	<b><u>Organization</u></b>	<b><u>Telephone</u></b>	<b><u>E-Mail</u></b>
Azhagarasan Annadorai	Kaizentric Technologies	(510) 343-5135	azhagarasan@kaizentric.com
Jim Blaha	The Hartford		James.Blaha@hotmail.com
Cindy Johnson	Bristol-Myers Squibb	(203)-677-6547	gjcj372@comcast.net
Karol Katz	Yale University	(203) 785-7572	Karol.Katz@yale.edu
Robert Krajcik	Bristol-Myers Squibb	(203) 677-6125	robert_krajcik@yahoo.com
Barbara Moss	The Hartford	(860) 547-7396	barbara.moss@ thehartford.com
Chuck Patridge	Independent Consultant		charles_s_patridge@prodigy.net
Peter Prause	The Hartford	(860) 547-7484	Peter.Prause@thehartford.com
Janet Stuelpner	SAS Institute	(203) 966-7520	jstuelpner@usa.net
Mahinda Yapa	Pitney Bowes	(203) 922-4662	myapa@yahoo.com
Helen Wang	Ingenix	(860) 231-8241	Helen.Wang@ingenix.com
Gerald Zuckier	Aetna	(860) 636-4053	GZuckier@alum.mit.edu

## DIRECTIONS to Wellpoint (Anthem BCBS)

370 Bassett Road

North Haven, Connecticut 06473

### From the North:

91S

Exit 13 Rt5 (Wallingford~North Haven) (ABCBS is about 1 mile from exit)

Turn Left onto South Colony Road/Rt5 South

Turn Left onto Bradley St

Stay straight and follow to Bassett Rd

Sign notes both Quinnipiac University and WellPoint (Anthem BCBS) *Event Parking*

### From the South:

91N

Exit 12 Washington Ave (ABCBS is about 1 mile from exit)

Turn Right onto Rt5 North

Turn Right onto Glenn Rd

Turn Left on to Bassett Rd

Sign notes both Quinnipiac University and WellPoint (Anthem BCBS) *Event Parking*

