

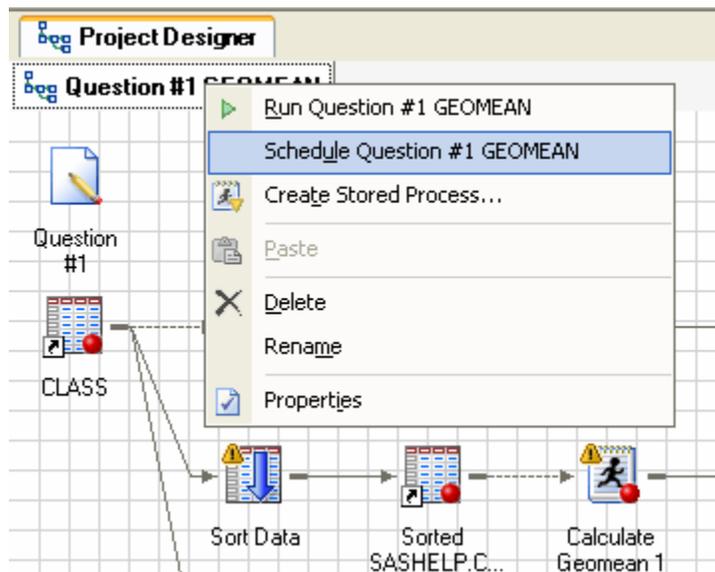
## Spring 2008 OASUS Q&A

The following answers are provided to the benefit of the OASUS Users Group and are not meant to replace SAS Technical Support. Also, an Enterprise Guide project is provided as a companion to this document. The project is available on the OASUS web site ([www.oasus.ca](http://www.oasus.ca)) under the Fall 2008 meeting since where these answers have been presented formally.

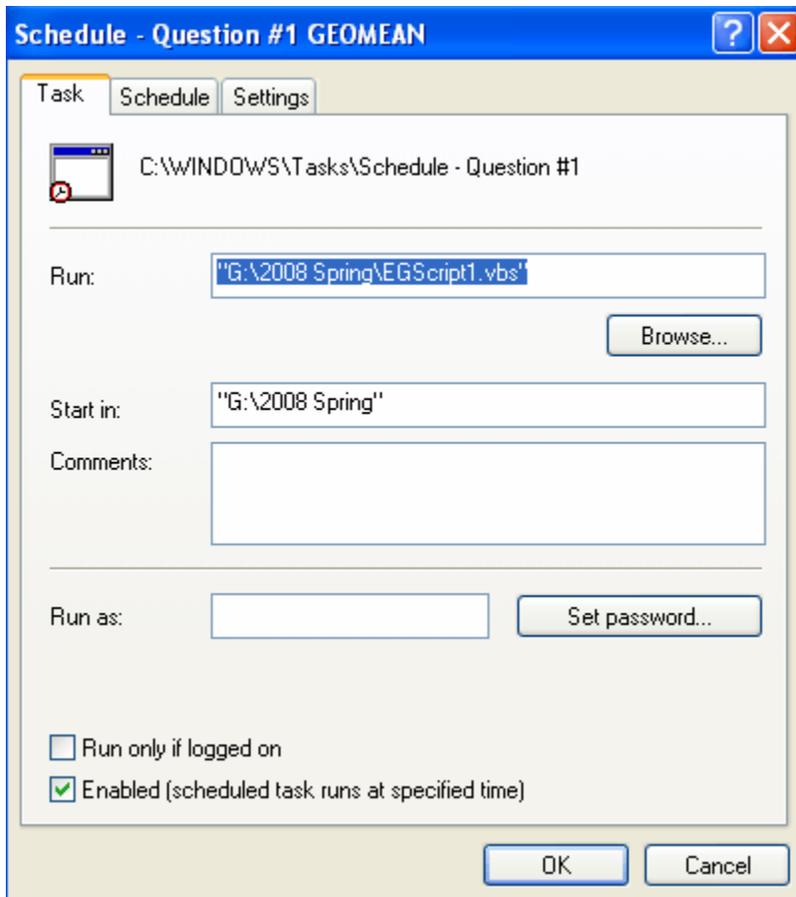
### **Question 1 : How to set up a recurring SAS job for multiple users?**

Windows Scheduling works well for recurring/batch SAS jobs. The easiest way to invoke the Windows scheduler is by using Enterprise Guide. In Enterprise Guide, an interface to the Windows scheduler is available. All scheduled tasks are recorded within an Enterprise Guide project. When a scheduled task is run, the results are recorded in the same Enterprise Guide project where it has been defined. Here is an example:

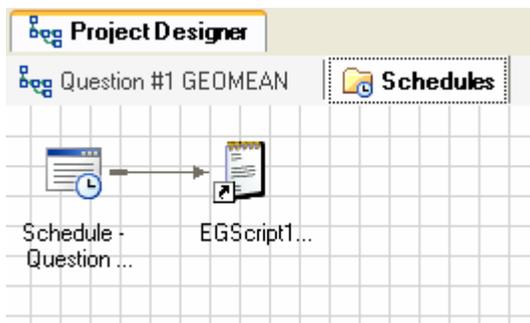
1) Invoke the scheduler with a right-click on a Process Flow.



2) The Windows Scheduler opens to schedule a task to run the Process Flow. Note that you can provide your Windows credentials to let the task run even if you're not logged on.



The schedule (along with any other schedules that may exist for the project) is then saved right in the project. EG creates a vb script to run you project unattended. It is possible to change those scripts by hand if need be.



References:

SUGI paper 019-30, Boost Your Programming Productivity with SAS® Enterprise Guide®, Chris Hemedinger, SAS Institute Inc., Cary, NC

**Question 2 : How can we create code to generate a geometric mean of "one" variable with "by" group? Ex. gmean of revenue by province**

The geometric mean (geomean) can be computed for a variable x with n values with the following formula:

$$\exp \left( \frac{(\log(x_1) + \log(x_2) + \dots + \log(x_n))}{n} \right)$$

SAS provides a geomean function but it is available in the DATA step, not directly in summary procedures. To compute the geometric mean for BY-groups, we propose three different approaches. Each approach is illustrated with an example that output the same result; that is the geomean of variable height by age using the standard data set sashelp.class (available in all SAS installations).

*1) Using BY-group processing in a DATA step*

This is the traditional approach that most people would use. If you are familiar with a DATA step, it works really well, is fairly concise and has no limitations with regards to the size of BY-groups.

```
data b(keep=age height_geomean);  
  
    retain sum_of_logs 0 i 0;  
  
    set sortedclass;  
        by age;  
  
    if first.age then do;  
        sum_of_logs = 0;  
        i = 0;  
    end;  
  
    sum_of_logs = sum(sum_of_logs,log(height));  
    i=i+1;  
  
    if last.age then do;  
        height_geomean = exp(sum_of_logs/i);  
        output;  
    end;  
  
run;
```

## 2) Using a compute block in PROC REPORT

If you are familiar with PROC REPORT, it is possible to compute the geomean using a COMPUTE block. It is less compact than the DATA step.

```
ODS RESULTS OFF;
proc report data=sashelp.class nowd out=a1;
  column
    age
    height
    height_log
    height_geomean;
  define age / group noprint;
  define height / display noprint;
  define height_log / computed noprint;
  define height_geomean / computed noprint;

  break after age / ;

  /* Accumulate the logs by group */
  compute height_log;
    height_log = log(height);
    sum_of_logs = sum(sum_of_logs, height_log);
    /* Count the number of lines in the group */
    i = sum(i, 1);
  endcomp;

  compute after age;
    height_geomean = exp(sum_of_logs/(i-1));
    /* Reset the counts for a new group */
    sum_of_logs = 0;
    i = 0;
  endcomp;
run;
ODS RESULTS ON;
```

## 3) Using PROC TRANSPOSE

PROC TRANSPOSE is a data manipulation facility that convert the columns of a data set into rows. To compute the geomean, you can place all the values pertaining to a BY-group in a single row. Once that is done, then a simple DATA step can do the computation one row at a time. The trick is to use the “of” operator to list all variables that start with a given prefix.

The big advantage with this approach is that you can use the SAS function GEOMEAN - you don't even need to know the formula! An important limitation however could be the size of a BY-group since all the values of interest must reside in memory for the computation to take place.

The example below should be run and the output of PROC TRANSPOSE be carefully analysed.

```
PROC TRANSPOSE DATA=sashelp.class
  OUT= TransposedClass
  PREFIX=height
  NAME=Source
  LABEL=Label;
  BY Age;
  VAR Height;
RUN; QUIT;

data c(keep=age height_geomean);
  set TransposedClass;
  height_geomean = GEOMEAN(of height:);
run;
```

References:

SAS Global Forum paper 222-2007, The Power of the BY Statement, Paul Choate, California Department of Developmental Services, Sacramento, CA, Toby Dunn, Stat Works Inc., San Antonio, TX

SAS Global Forum paper 234-2007, An Animated Guide: Proc Transpose, Russ Lavery: Contractor for Numeric Resources

SUGI 24 paper 60, Changing the Shape of Your Data: PROC TRANSPOSE vs. Arrays Bob Virgile, Robert Virgile Associates, Inc.

SAS Global Forum paper 079-2008, A Step-by-Step Introduction to PROC REPORT, David Lewandowski, Thomson Healthcare, Evanston, IL

***Question 3 : How do you refresh the sysdate without quitting your SAS application? Basically you want the sysdate = the date and time you run.***

This is very easy to accomplish by using the SAS System option DTRESET. This option specifies that SAS will update the date and time in the titles of the SAS log and the procedure output file. Otherwise, SAS will use the date and time when the session started.

***Question 4 : Is SAS Enterprise Miner useful for narrative analysis?***

Yes, Enterprise Miner with the add-on product called Text Miner can perform very sophisticated text data analysis.

Example: Analysis of information such as customer letters and call center notes may provide valuable information about customer dissatisfaction or insights into service and product needs. SAS Text Miner provides a full range of predictive modeling tools that can unearth opportunities for timely exploitation.

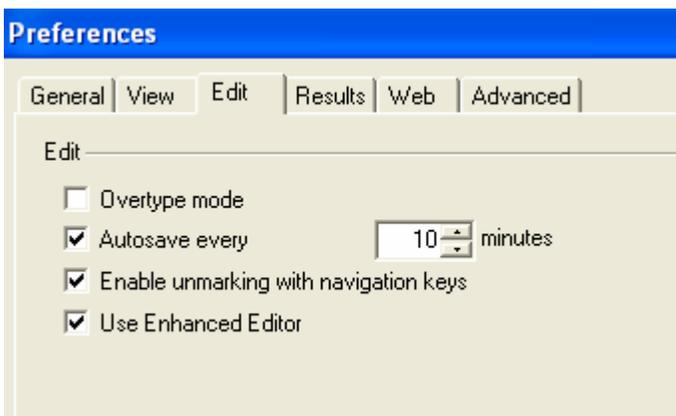
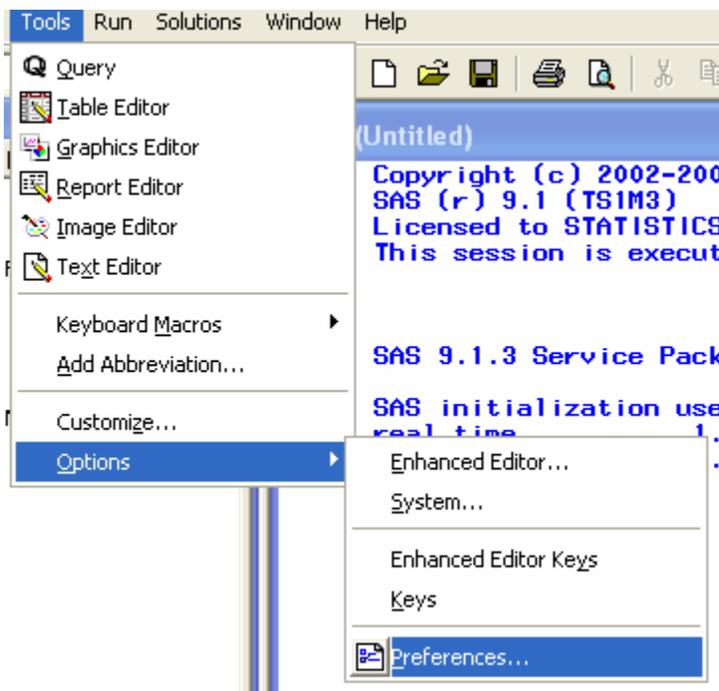
### **Question 5 : Is there a way to save a program automatically at fixed intervals when using the Enhanced Editor?**

The Enhanced Editor is available in both the SAS Display Manager and Enterprise Guide.

In the case of Enterprise Guide, there are no options at the moment to save a project automatically and at fixed intervals. Therefore manual saves are required.

In the case of the SAS Display Manager, it is possible to set an autosave option that will trigger an automatic backup to be taken at fixed interval time.

Both the autosave option and the time interval can be set under tools-options-preferences:



In SAS 9.1 under Windows XP, when the autosave option is turned on, SAS saves a copy of the program file being edited in the directory “C:\Documents and Settings\\Application Data\SAS\EnhancedEditor” where <username> corresponds to the user logged in. The name of the backup file is identical to the file being edited except that the extension is .\$AS instead of .SAS. If the SAS Display Manager terminates abnormally, then the latest backup is kept for the user to recover.

***Question 6 : In a keep statement, you can use Q: to keep all variables starting with Q. Is there any way to keep all variables ending in 08?***

This question was answered in the Spring 2007 Q&A.

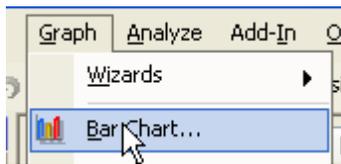
**Question 7 : I am trying to create complex graphs (histogram with proportional columns) with SAS EG. This type of graph can be built using the G100 option in SAS code but I was wondering if this could be done within the SAS EG Interface?**

Enterprise Guide has built-in interfaces (called tasks) for many SAS procedures. Unfortunately, not all procedures and options are supported by these tasks. If you find a procedure for which there is no EG task, you can always invoke it through code or you could ask your IT group to build a custom task or a stored process.

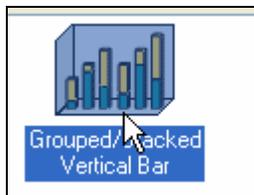
However, if a procedure has a corresponding EG task, it is likely that not all options and statements have been made available to the users. It is desirable to keep the number of options implemented to a manageable amount; otherwise the interface would become very clumsy.

This is not the case for the G100 option referred to in the question: it is implemented in the Bar Chart task!

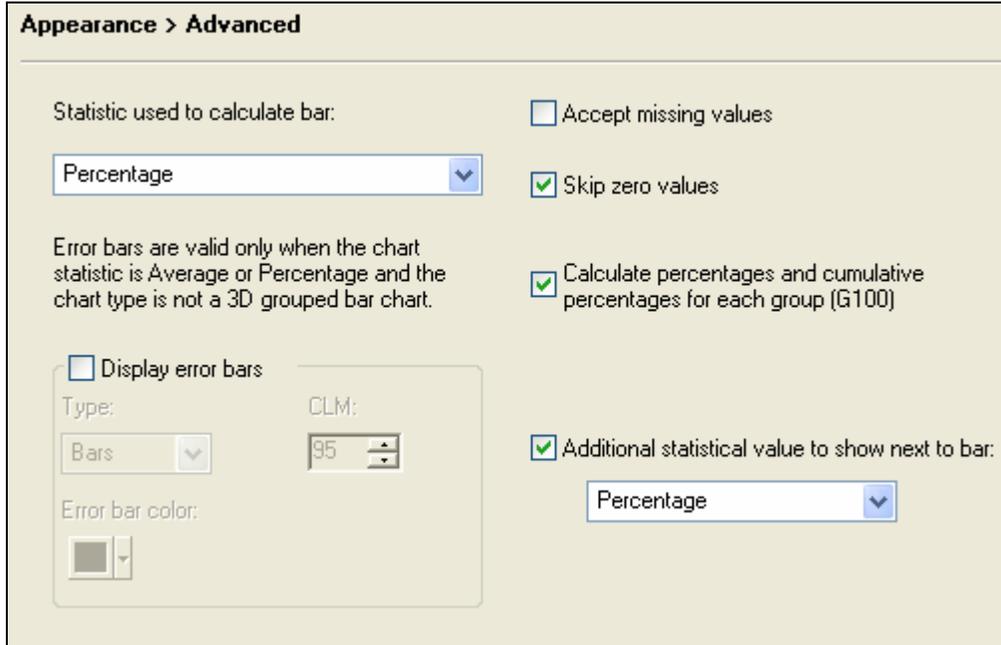
1) Select the Bar Chart task.



2) Pick a Grouped/Stacked Vertical Bar.

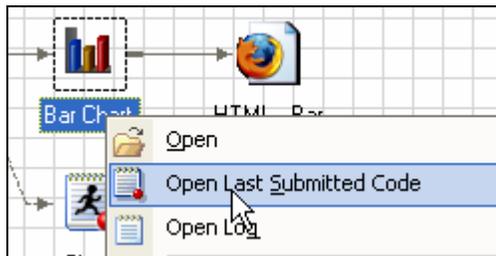


4) Under the Advanced tab, select the G100 option.



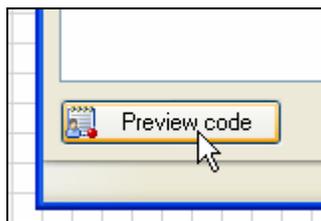
If one or many options are missing in an EG task, you have two alternatives:

1) Generate the code using the EG task. Then, copy the generated code into a code node and add the missing option(s).

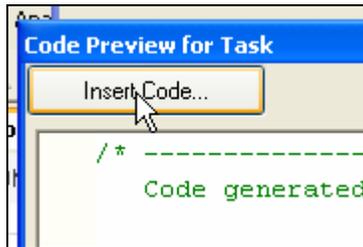


2) Use a little known feature of Enterprise Guide which allows you to insert custom code into the generated code of a task. The advantage of this approach is that the custom code is retained with the task.

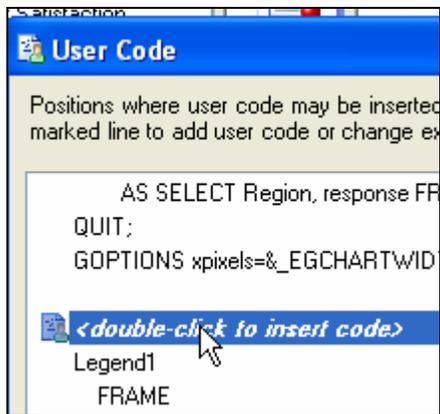
Select Preview code from the task window.



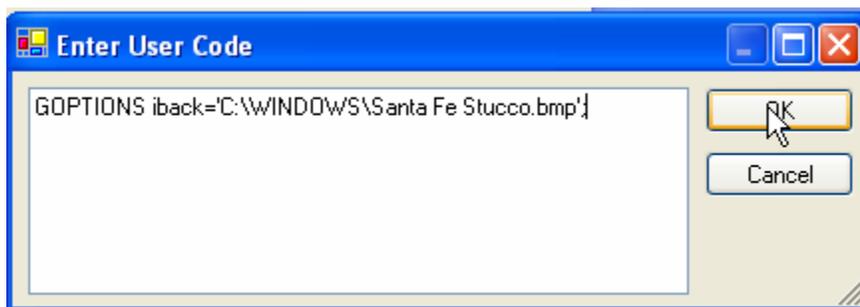
Select Insert Code...



Double-click an insertion point. There may be several of insertion points in the code; you have to select the one that is compatible with the option you want to add.



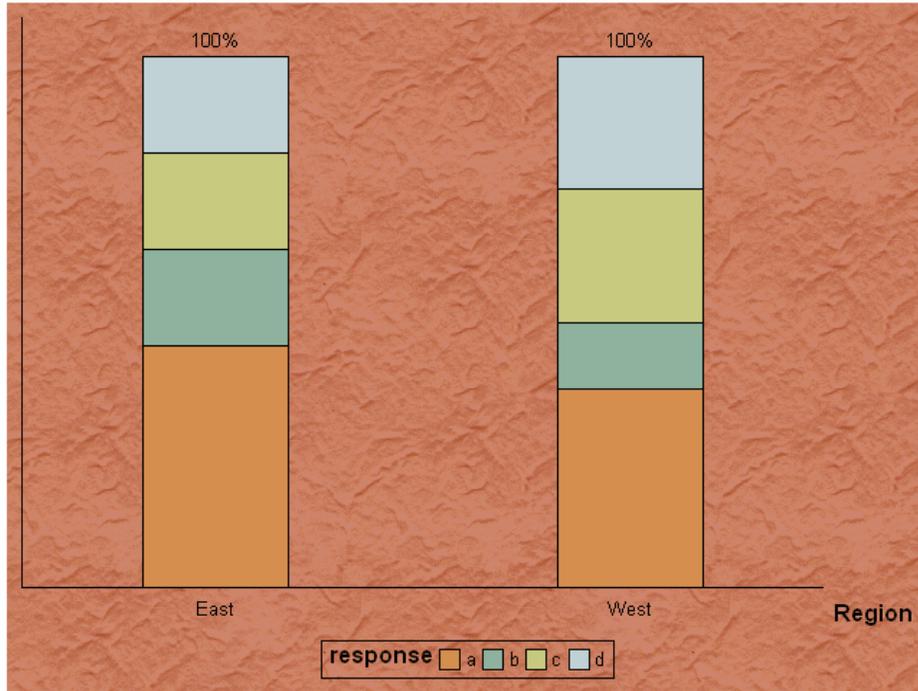
Type the desired option. Make sure you follow proper syntax; this code will be inserted as is. In the example below, we insert the option "iback" and specify an image file to display in the graph's background area.



Here is the result (with the G100 option selected):



**Bar Chart with custom code**



**Generated by the SAS System on 10NOV2008 at 5:05 PM**

**Question 8 : ODS Text - How do we suppress the space between the text and the Proc report table output? I want the text on the left corner of this table.**

Although the question is somewhat vague, it is possible to answer by saying that ODS should do the trick. With ODS, it is possible to:

- 1) Left justify a title.  
Title1 justify=left 'First Line of the Title';
- 2) Add a text before the report generated by PROC REPORT.  
text = 'Example for OASUS';
- 3) Left justify the report itself.  
style(report)=[just=left]

The following is the complete example for the PDF destination:

```
ods pdf
  file='c:\temp\temp.pdf'
  startpage=no
  text = 'Example for OASUS';
  title1 justify=left 'First Line of the Title';
  title2 'Second Line of the Title';
proc report data=sashelp.class nowd
  style(report)=[just=left];
  column age weight height;
run;
ods pdf close;
```

Which results in :

<i>First Line of the Title</i>	<i>Second Line of the Title</i>	20:31 Wednesday, November 12, 2008 1
Example for OASUS		
<b>Age</b>	<b>Weight</b>	<b>Height</b>
253	1900.5	1184.4

References :

“ODS Tips & Tricks” , [http://www4.ncsu.edu/~vmdepuy/ODS\\_tips.doc](http://www4.ncsu.edu/~vmdepuy/ODS_tips.doc)

Using SAS® Output Delivery System (ODS)”  
<http://www2.sas.com/proceedings/sugi26/p058-26.pdf>

## **Question 9 : How can we create more than one tab in Excel in a stored process using SAS Add-In for Excel?**

To achieve this, the stored process needs to stream Excel data directly to Excel using the `_Webout` file reference and the HTTP protocol. In other words, the stored process output type should be “streaming”.

The streaming results generated by the stored process must start with a HTTP header that identifies the content. This can be done by using a special Data Step function called `STPSRV_HEADER`. For example, in the case of Excel you would call the function to generate the appropriate header as followed:

```
data _null_;
  rc = stpsrv_header('Content-type', 'application/vnd.ms-excel;');
  rc = stpsrv_header('Content-disposition',
                    'attachment filename=test.xls');
run;
```

To feed the stream with the appropriate Excel results, in our case 2 or more sheets, the use of ODS is recommended. One approach with ODS is to use the `MARKUP` destination with the `ExcelXP` tagsets. This facility allows generating XML output that can be opened by Excel into one or many worksheets as followed:

```
ods listing close;
ods tagsets.excelxp body="test.xls";
ods tagsets.ExcelXP options(sheet_interval='none' sheet_name='Sheet 1
- Print');
proc print data=sashelp.class ;
run;
ods tagsets.ExcelXP options(sheet_interval='none' sheet_name='Sheet 2
- Freq');
proc freq data=sashelp.class;
  tables age / nocum nopercent;
run;
ods tagsets.excelxp close;
```

References:

SAS Global Forum Paper 229-2007, Creating Multi-Sheet Excel Workbooks the Easy Way with SAS®, Vincent DelGobbo, SAS Institute Inc., Cary, NC

SAS Global Forum Paper 023-2007, Converting SAS/IntrNet® Programs to SAS® Stored Processes, Heather Weinstein, SAS Institute Inc., Cary, NC

***Question 10 : I have a client who runs SAS on UNIX through Excel. His work space always ends up under the root directory. How can he permanently point to a defined workspace path?***

The best solution to this problem is to change the SAS config file to set the –Work option to whatever location that you find convenient and this change would apply to all users.

The config file to be updated can be found in <<root>>\SAS\SAS 9.1\nls\en\sasV9.cfg in the case of Foundation or in some other location for the Intelligence Platform server. Check with your IT support team before changing the SAS config file in the later case.

## **Question 11 : How to get rid of the warning messages in the SAS log when missing numeric values are read.**

A specific example of the warnings would help to give a precise response. Sometimes missing values are legitimate, sometimes they are not. However, when a variable is missing for every record in the file, there may be a problem with the program. More often, when your dealing with missing values within your program, SAS will issue a note underlining the statement it thinks you should need to know has missing values. A best practice is to eliminate the cause of a problem rather than the warnings.

There are many ways to avoid warnings for missing values that you consider to be “valid”. One method is to use a DATA step function such as SUM that treats missing without issuing warnings. You have to make sure that the behaviour of the function meets your requirements. Another method is to control the treatment of missing values using an IF construct. In this case you have to decide how a missing value is to be used. Is it equivalent to a zero or does the calculation return a missing? Again, it all depends what your requirements are. The code below shows both methods.

```
* First create a dataset with some missing values;
data test;
  set sashelp.class;

  if age>13 then do;
    height=.;
    weight=.;
  end;
run;
* Do some calculations with missing values with no warning;
data test2;
  set test;

  * Use the sum function;
  weight=sum(weight,10);

  * Use an if statement;
  if (height = .) then height=.;
  else height=height+10;
run;
```

### *References*

SUGI paper 025-31, MISSING! - Understanding and Making the Most of Missing Data  
Suzanne M. Humphreys, PRA International, Victoria, BC (Canada)

SAS Global Forum paper 082-2008, Tools of Miss-Calculation: Managing Missing Values with SAS®, William C. Murphy Howard M. Proskin and Associates, Inc., Rochester, NY

***Question 12 : Can the hash object remain in memory and be reused in the next DATA step without having to be reloaded? (I think the hash object can only be used in a DATA step therefore when it terminates, the object is deleted from memory.) I would like to use the hash object through the life of the program not just during a data step.***

The hash object is an in-memory lookup table accessible from the DATA step. Unfortunately, the hash object can only persist for the duration of DATA Step. It is not possible to retain the hash object in memory once a DATA Step completes.

If you need to share a hash table between DATA steps, you can load the hash table from a data set that has been pre-loaded in memory. This is possible by using the SASFILE statement.

The SASFILE statement reduces the open/close operations to just one and also I/O processing by holding the data in memory. When the SASFILE statement executes with the LOAD option, SAS opens and reads the file in memory. Then, when subsequent DATA and PROC steps execute, SAS does not have to open the file each time and issue I/O requests, since all the data resides in memory. The file remains open until a second SASFILE statement closes it, or until the SAS session ends.

Try the following code that uses the SASFILE statement and a hash table.

```

/* The content of sashelp.zipcode will be loaded into a hash table for
every lookup. */
/* To improve performance, the entire dataset will be loaded in memory
first using the */
/* SASFILE statement. */

/* Macro to lookup 100 Zip codes generated randomly */
%macro LookupZip(outdata);

    data &outdata(keep=Zip City StateCode);

        length Zip 8 City $35 StateCode $2;

        /* Load hash table from dataset*/
        if _N_ = 1 then do;
            declare hash h(dataset='sashelp.zipcode');
            h.defineKey('Zip');
            h.defineData('City','StateCode');
            h.defineDone();
        end;

        /* Lookup 100 Zip codes generated randomly */
        do i=1 to 100;
            zip=round((100000*rand('UNIFORM')),1);

            /* If no match, then unknown Zip code */
            if h.find() NE 0 then do;
                City='Unknown';
                StateCode='??';
            end;

            output;
        end;

    run;

%mend;

/* First load the zipcode dataset in memory */
SASFILE sashelp.zipcode LOAD;

/* Lookup Zip codes */
%LookupZip(zipmatch1);
%LookupZip(zipmatch2);
%LookupZip(zipmatch3);

/* Close dataset and free memory */
SASFILE sashelp.zipcode CLOSE;

```

Before you decide to use the SASFILE statement, make sure that you have taken into account the time required to load the data in memory. Also, don't forget that I/O processing can be reduced only if there is sufficient real memory allocated to your SAS session. If the SAS file does not fit in real memory, the operating system will use virtual memory. Because virtual memory resides on disk, you could be back to square one

since you will be replacing data access I/O requests with swapping I/O requests. You should always set up a test in your environment to measure performance with and without the SASFILE statement.

References:

SAS Global Forum paper 271-2007, Getting Started with the DATA Step Hash Object, Jason Secosky, SAS Institute Inc., Cary, NC, Janice Bloom, SAS Institute Inc., Cary, NC.

Insights Newsletter 1<sup>st</sup> quarter 2007, One thing every SAS programmer should know, Yves DeGuire, Statistics Canada.

Paper 095-2008, Using Table Lookup Techniques Efficiently, Jane Stroupe & Linda Jolley, SAS Institute Inc., Cary, NC

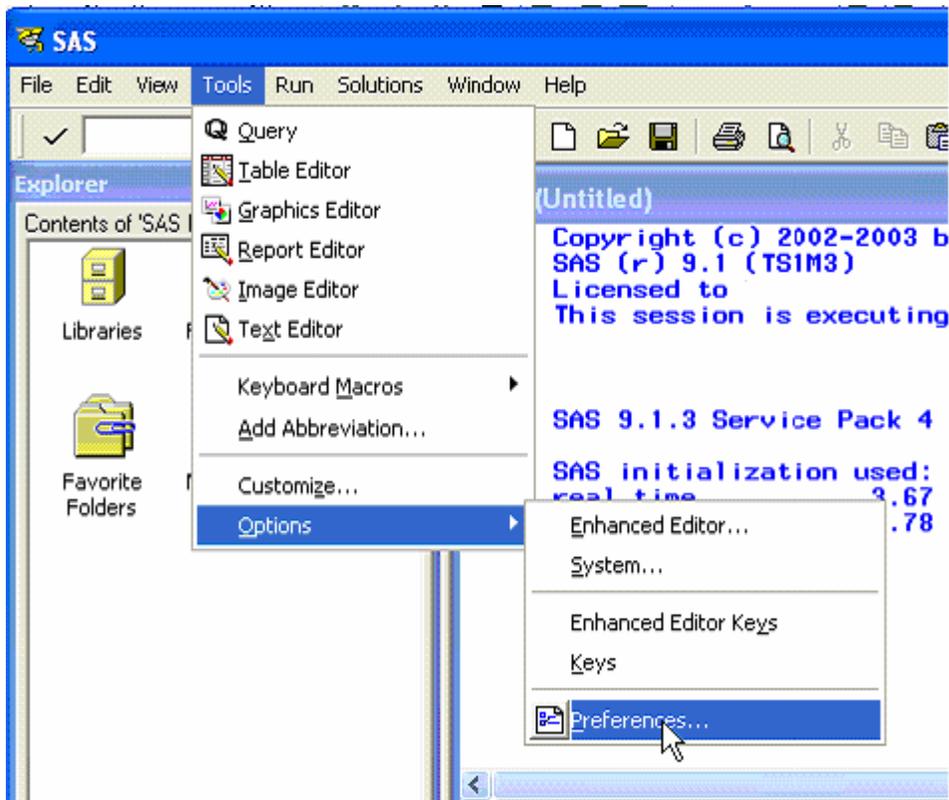
***Question 13 : In SAS 9 two things that were very handy in SAS 8 are missing. Using ALT & Click & drag mouse to select desired column or rectangle. Using CTL F or H - you can no longer copy a variable phrase into the Find or Replace box.***

Both commands work in SAS 9! Also note that this is true for both the Display Manager and Enterprise Guide.

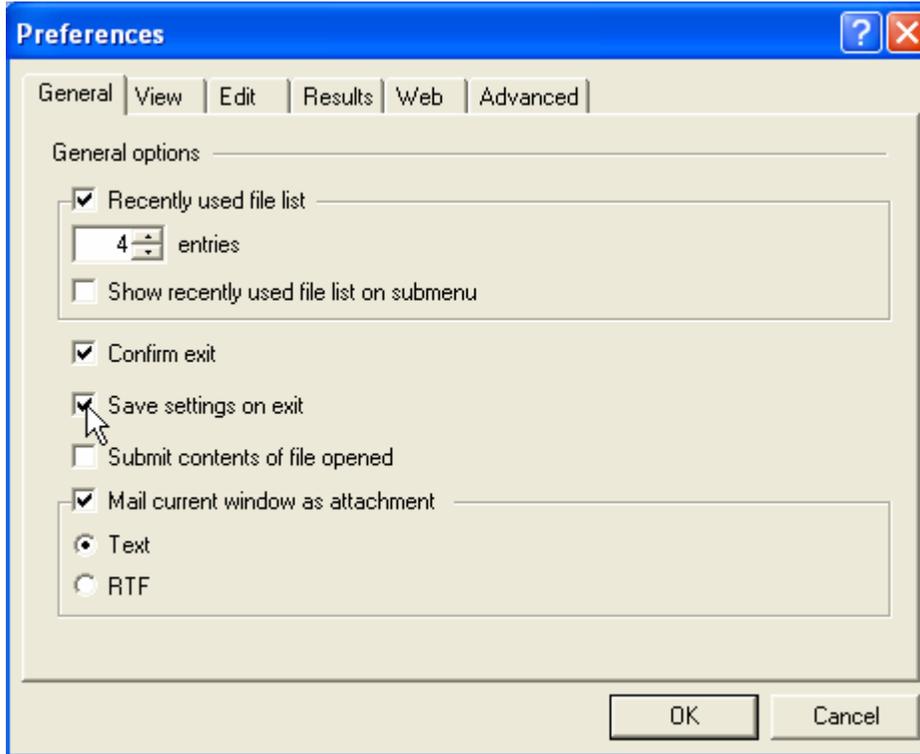
## Question 14 : How to get SAS to open as a full window when SAS opens?

This is very easy to do in the SAS 9.1 Display Manager.

- 1) First, get to the preferences screen



- 2) Turn on “Save Settings on Exit”

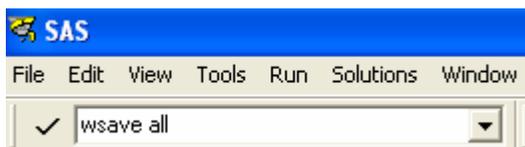


- 3) Arrange the Log, the Program and the Output windows the way you like (for example maximise the Program window).
- 4) Finally, exit from the Display Manager. Your settings should be retained for future sessions.

You should turn off the “Save setting on exit” option as soon as you have arranged the Display manager windows the way you like. Otherwise, you will overwrite these as soon as you re-arrange any windows.

Rather than using the preferences windows, you can instead issue the WSAVE ALL commands to save the settings. This will achieve similar results and need only 2 steps:

- 1) Arrange the Display Manager windows the way you like.
- 2) Issue the WSAVE ALL command:



Regardless of the approach you use, the information about window settings is saved in the user profile catalog, so it may be wise to backup the catalog in case you need to recover your personal settings.