

Spring 2019
C.J. Anderson

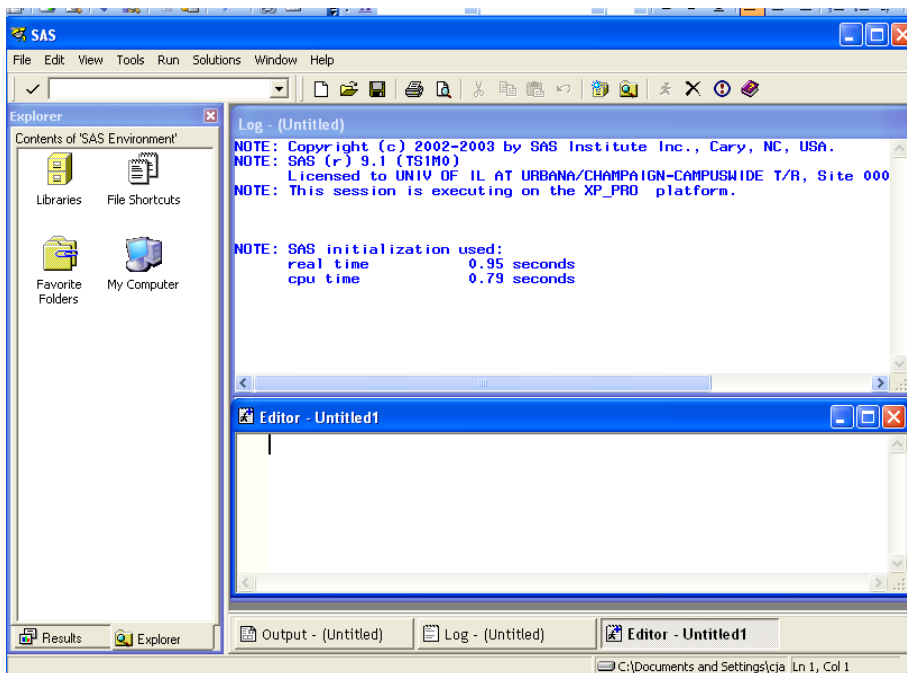
Introduction to SAS: General

Go to course web-site and click on **hsb-data.sas**

There are 5 main working environments (windows) in SAS:

- **Explorer window:** Lets you view data in SAS data sets or go to output for specific commands.
- **Editor window:** This is where you enter SAS commands (programs).
- **Log window:** Commands entered and run are repeated here along with warning and error messages can be found in the log window. Also, when you create a data set, information about the data set is printed (e.g., number of observations, variables, etc).
- **Output window:** A plain text window with the results of your analyses.
- **Results window:** If you have the option set to produce HTML output, a “nice” looking output can be found here.

When you open SAS you will see:



Some basic SAS syntax

- Just like sentences end with periods, “.”, all SAS commands must end with a semicolon “;”
- It is good practice to put comments in your SAS programs. Comments begin with an asterisk “*” and end with a semicolon “;”. For example.

* Homework 1, problem 2;

- SAS variable names:
 - 1 to 32 characters in length
 - They must begin with a letter (A-Z). Note that SAS does is not case sensitive; that is, it read “A” and “a” as the same.
 - The second and remaining characters in a variable name can be other letters, numbers, or underscores (i.e., “_”).
 - By default SAS assumes that variables are numeric. If a variable is character, then the name must be followed by a space and then a \$.

Two basic steps in a SAS program

- (1) The **DATA** step: create or read in data, modify variables, create new variables, etc.
- (2) **PROCEDURE** step (PROC for short): data analysis

SAS Example

1. If you had this sas program in a file, you would need to find **File** on the main tool bar: File > Open program > ... Use Browse to find the name of the file. You would then need to click on this file and push the OPEN button. If all went well, a file should have opened in your “program” window and look something like this....

```
* SAS commands that create data set of the high school and beyond data;
*libname sasdata 'c:\data';
data hsb;
  input id sex race ses sctyp hsp locus concpt mot
        car rdg wrtg math sci civ;

  label id = 'ID number '
        sex = 'Gender: 1=male, 2=female'
        ses = 'Socio-Economic Status'
        sctyp = 'School Type: 1=public, 2=private'
        hsp = 'High School Program'
        locus = 'Locus of Control'
```

```

concept = 'Self Concept'
mot = 'Motivation'
car = 'Career Choice'
rdg = 'Reading T-Score'
wrtg = 'Writing T-Score'
math = 'Math T-Score'
sci = 'Science T-Score'
civ = 'Civics T-Score';
datalines;
1 2 1 1 1 3 0.29 0.88 0.67 10 33.6 43.7 40.2 39.0 40.6
2 1 1 1 1 1 -0.42 0.03 0.33 2 46.9 35.9 41.9 36.3 45.6
3 2 1 1 1 1 0.71 0.03 0.67 9 41.6 59.3 41.9 44.4 45.6
4 2 1 2 1 3 0.06 0.03 0.00 15 38.9 41.1 32.7 41.7 40.6
.
.
.

```

RUN;

2. Click on the run icon on the main tool bar to submit the commands in the program window (it looks like a little person running).
3. Check the log file to make sure everything ran OK. You should see

```

NOTE: The data set WORK.HSB has 600 observations and 15 variables.
NOTE: DATA statement used (Total process time):
      real time           0.29 seconds
      cpu time            0.03 seconds

```

4. Try some procedures:

- a) Create a cross-classification of two (or more or less) variables:

```

PROC FREQ DATA= hsb;
  TABLES race*sex / NOROW NOCOL ;
RUN;

```

- b) Compute the mean of SES and math achievement scores;

```

PROC MEANS DATA= hsb;
  VAR ses math;
RUN;

```

- c) Sort the data by gender.

```

PROC SORT DATA= hsb;
  BY sex;
RUN;

```

- d) Compute the means of SES and math achievement for each gender and save the results to a file, which we then print to see what's in the save (working file).

```

PROC MEANS DATA=hsb;
  CLASS sex;
  VAR ses mathach;
  OUTPUT OUT=mymeans MEAN=mses mmath STD=stdses stdmath;

PROC PRINT DATA=mymeans;
RUN;

```

5. Some more descriptive statistics, including box plots, stem-n-leaf (histogram), normal probability plot:

```

proc univariate data=hsb plots;
  var math rdg;
run;

```

6. To compute correlations and covariances, try

```

proc corr data=hsb cov;
  var math rdg wrtg sci civ;
  title 'All the achievement tests';

proc corr data=hsb cov;
  var math rdg wrtg sci civ;
  with locus concpt mot;
  title 'Achievement with the Psychological Measures';
run;

```

7. Save your program commands to a file:
- Make sure that your program window is the current/open window.
 - File > Save As >give it a name....
8. Save your output to a file:
- Make sure that your listing/output window is the current/open window.
 - File > Save As >give it a name with either type .lst or .txt
- 9 If you want to save your output in rich text format (ie, if you want to use it in a MS word document), use (for example):

```

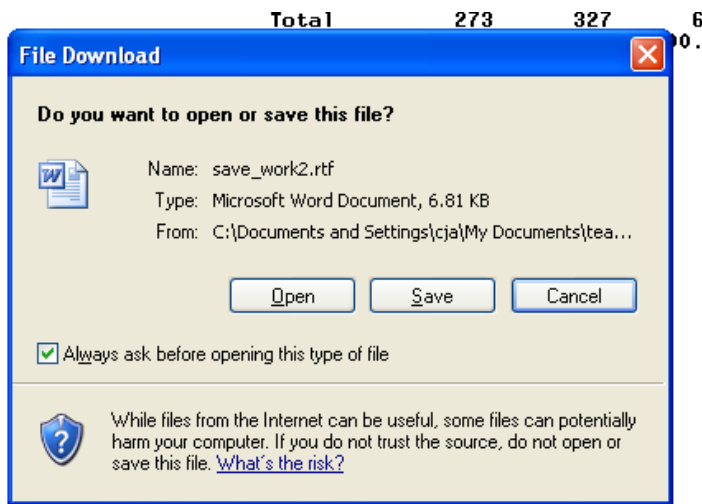
* Create an RTF file ;
ods rtf file='C:\Documents and Settings\cja\My
Documents\teaching\multivariate 584\save_work.rtf';

PROC FREQ DATA= hsb;
  TABLES race*sex / NOROW NOCOL ;
RUN;

* Close the RTF file;
ods rtf close;

```

After this runs, you will see the window:



Click on Save