Introduction to SAS

SAS Principles

SAS follows a 'write a program / run a program' model. To get data analyzed using SAS, you must:

1. Write commands that tell SAS how to read the data
2. Write commands that tell SAS what kinds of analyses you want
3. Tell SAS to execute those commands
4. Look at the .log file or the Log window to see whether SAS understood your commands
5. If you made a mistake,
   (a) edit your program
   (b) resubmit all or part of it
6. Look at the output and interpret it.

Access to SAS at ISU

Please check course syllabus or go to the web site

http://stat.iastate.edu/statistical-software-sas

to read more information on access to SAS at ISU.

SAS Programs

SAS programs are organized into:

- DATA steps. These convert your data file(s) into a form that is usable by SAS.
- PROC steps. These ask SAS to run a particular analysis.

You may have more than one DATA step and more than one PROC step. The typical Stat 500 program has one DATA step and multiple PROC steps.

SAS is a very powerful database manager. We will use only the simplest parts of its data management capability.

Data structure: In this course, we will assume that a data file has a 'flat-file by rows' format; that is:

- Each row contains one observation.
- Data values are separated by one or more spaces.
- There may be a header line giving the variable names. This can be omitted.

Here is an example:

```
group yield
a 29.9
a 11.4
b 26.6
b 23.7
a 25.3
b 28.5
b 14.2
b 17.9
a 16.5
a 21.1
b 24.3
```

There are 11 observations (one per row). Each observation has two variables. The first identifies a treatment group; the second is the yield.

Data sets: SAS does not work directly on the original data file. The analysis parts of SAS (the PROC steps) use data that are stored in a SAS data set. So the first part of a typical SAS program creates a SAS data set from the original data. The minimum DATA step has a DATA line, an INFILE line (or a CARDS line, or a DATALINES line), and an INPUT line.

```
data tomato;
  infile 'C:\Documents and Settings\pliu\My Documents\Courses\500\data\tomato.txt';
  input group $ yield;
  run;
```

This creates a SAS data set called tomato. The data come from the file 'tomato.txt' in the

```
C:\Documents and Settings\pliu\My Documents\Courses\500\data
```

folder. You need to specify the full path to a file, as shown in the above example. The easiest way to get the full path correct is to highlight a file name, right click on the mouse, and select "Properties." You want the information labeled "Location." Then you can highlight it, copy it to the clipboard, and paste it into your SAS program. Remember, you also need the file name (tomato.txt).

For this example, each row has two values. The first value will be stored with the variable name GROUP; the second value will be stored with the variable name YIELD.

By default, SAS treats all variables as numeric. That is fine for YIELD, but the values of GROUP are not numbers. You tell SAS to read a variable as a character variable using the dollar sign ($) **after** the variable name on the input line. If you omit the $, you get a lot of errors (values not numeric) in the Log window.

PROC steps: Each PROC step tells SAS to perform a particular analysis. Each PROC step includes commands that describe the specific analysis. I will introduce appropriate commands as needed throughout the semester.

```
proc ttest;
```

```
class group;
var yield;
title 'T-test of tomato yield';
run;

proc sort;
by group;  /* sort the data by group before doing the boxplot */
proc boxplot;
plot yield*group;
title 'High resolution boxplots for each group';
run;
```

Each command to SAS MUST end with the semicolon (;). One command can span multiple lines, or one line can have multiple commands. SAS only cares that each command ends with the semicolon (;).

HINT: If SAS gives you a bunch of errors, and the commands look correct, check to see if you left out the semicolon (;). This is a common problem, and omitting the semicolon (;) really confuses SAS.

While not necessary, it is good practice to end each PROC or DATA step with a "run;" command.

HINT: If you submit some SAS commands but don't get any output, or you don't get any output from the last PROC step, you forgot the "run;" command. You don't need to resubmit everything. Just type "run;" in your program Editor window (probably at the end), highlight it, and then click "Submit" from the "Run" menu. Then you should get the output from your last PROC step.

There are almost always multiple ways to do the same thing in SAS. If you see a different PROC or DATA step used, it may just be a different way to accomplish the same thing.

Working with Windows SAS on a PC

To start SAS: Click on the SAS icon.

The windows: Five windows will appear on your screen.

- Explorer: This window provides a point-and-click interface to SAS data sets. This is less important and usually can be ignored.
- Results: This window contains a running record of the output from your SAS session. When you click the output names in the Results window, you link directly to the output in the Output Results window for the HTML destination (default).
- Editor: You enter your SAS commands here. When you give the "Submit" command, SAS will execute the commands that are in this window. You can execute a subset of the commands here by highlighting the desired commands before using "Submit."
- Log: This window contains SAS's responses to your commands. Any errors and various informative messages will appear here.
- Output: This window contains the output from your SAS commands.

Editing files: The SAS Editor window works just like a word processor. Type in your commands and use the enter key to get a new line.

If you use Word or other standard word processors to edit your file(s), you need to save the files in text format (.txt) using "Save As ...". SAS can not read files saved as Word documents (.doc files).

Loading a class SAS program: The class web site will have examples of SAS programs. You can download the programs using a web browser, save them, and then load them into SAS using "Open Program ..." under the "File" menu.

The default extension for a SAS program is .sas. Your browser may not allow you to save the file as a .sas file. You may have to save the file as a text file (.txt). Then you rename the file (highlight the file name and right click "Rename") and replace the "filename.txt" with "filename.sas". This is a Windows/Internet Explorer problem.

Saving, Running, and Exiting To save the contents of any window to a file, click on the "Save As ..." command under the "File" menu in the desired window, type in (or click on) where you want to save it, name the file, and click "OK."

Reminder: Always check the path name to make sure it will save the file where you want to save it. The default directory is the SAS program directory, which may not be easy to find.