

# Statistics 210 - Introduction to SAS

(updated Fall 2009)

## References

It is my goal to provide all the information that you need to work with SAS in this course. This is one of two handouts. This one emphasizes how to run SAS. The other handout is a SAS language guide. I know of two additional good sources of information (though I'm sure there are many other as well). First, you can purchase *The Little SAS Book: A Primer* by L. D. Delwiche and S. J. Slaughter. Now in its 3rd edition, this book has been recommended by several students. Second, UCLA's Academic Technology Services maintains a very useful SAS website, [www.ats.ucla.edu/stat/sas](http://www.ats.ucla.edu/stat/sas).

## Accessing SAS

I am aware of three approaches to accessing SAS.

- **Purchase** - You can purchase a copy of SAS for your computer (or for a machine in a lab that you have access to). The cost is \$45 for a one year license. You can pay via the web if you are going to charge to a UCI account (go to [licenses.nacs.uci.edu](http://licenses.nacs.uci.edu) and then search for SAS to get instructions) or you can by pay cash/check in person at the NACS office (E2130 Engineering Gateway). Installation is quite easy.
- **Public Lab** - Unfortunately SAS is only available in one lab on campus. It has been a constant struggle to keep it available in campus labs. Currently ICS 364, a large computer lab, has SAS installed on 5 machines in the center of the room. (These are in one row that is in front of the lab monitor's desk.)
- **Public Computing Cluster** - SAS is available on the Broadcomm Distributed Unified Cluster (BDUC), a group of computing nodes run by a single "resource manager". This option is "free" (in terms of dollars) but is the most complicated in terms of setup. You should be fairly at home with your computer (downloading and installing free software, connecting via ssh, transferring files via ftp) before attempting this approach. All the info you need is at [www.nacs.uci.edu/computing/bduc/newuser.html](http://www.nacs.uci.edu/computing/bduc/newuser.html). A short description of how to proceed is provided here but you may want to refer to that site for more details. There are four steps:
  1. Request an account by emailing to [bduc-request@uci.edu](mailto:bduc-request@uci.edu) (you should probably tell them you are in Stat 210 and want to access SAS);
  2. You will need a secure way of connecting to the cluster (Mac and Linux have "ssh" builtin typically. Windows will require you to download an application to do this, PuTTY ([www.chiark.greenend.org.uk/~sgtatham/putty/](http://www.chiark.greenend.org.uk/~sgtatham/putty/)) is highly recommended.) Once you have the ability to connect, then you want to connect to the specific BDUC node that has SAS by typing 'ssh bduc-claw1.nacs.uci.edu'.
  3. You will need a way to transfer files to BDUC (data) and from BDUC (SAS results). A secure copy/ftp utility is helpful (Mac and Linux have "scp" and/or "sftp" builtin typically and with these you use put/get to move files from one machine to another. Windows will require you to download an application to do this and WinSCP ([www.winscp.net](http://www.winscp.net)) is recommended.)
  4. To run SAS interactively you will need an X11 graphics windowing capability on your home machine (Linux typically comes with this automatically; Mac installation DVDs have it but not sure it is installed automatically; Windows requires yet another application and here Xming ([www.straightrunning.com/XmingNotes/](http://www.straightrunning.com/XmingNotes/)) is recommended.)

If you get all this done, then congratulations because you are ready to analyze data on BDUC!

## Preliminaries - The SAS Display Manager System

Assume you are in front of a PC with SAS. Begin SAS by double-clicking. This will start *SAS* and open up the display manager system. (On BUC you must first start your X11 application and then connect to BUC. Just type 'SAS' to the prompt and if all is set up right the display manager system should open up.) The display manager system is made up of a set of windows that correspond to the various activities that you will perform during a *SAS* session. The windows are: **EDITOR** for writing and submitting SAS programs, **LOG** for displaying SAS commands as they are submitted and corresponding error/warning messages, **OUTPUT** containing SAS output, **RESULTS** which provides an outline of your output and an easy way to page through it, and **EXPLORER** which allows you to navigate around and provides a place to store data files. SAS will also generate separate graphics windows when you create plots. At any time, you can move among the windows (you will often want to flip back and forth among the log, program editor, output and graphics windows) by using the **View** or **Window** menu. Once a *SAS* session is initiated, the various tasks are accomplished by selecting appropriate options from the menus displayed on the menu bar at the top of the *SAS* windows. Note that the options change depending on which window you are working in.

## SAS Programming

**Introduction.** There are several ways to go about analyzing data with SAS. The primary way that we will work use SAS requires you, the data analyst, to create a program of SAS commands. There is a bit of a learning curve before you can do this but this is how sophisticated SAS analyses are done in major companies or research groups. Other options for running SAS are described later in the handout.

**Typing a program directly into SAS.** The **EDITOR** window is where you enter and edit the *SAS* statements that you want executed. To start entering *SAS* statements, activate the **EDITOR** window (by clicking the mouse pointer inside the window if that window is not already active), position the cursor at the start, and begin entering *SAS* statements. *SAS* language syntax will be covered as we go; see also the short language guide that is available on the course website. Remember to include a "RUN;" statement at the end of the code you are submitting for otherwise SAS will not run your final PROC.

**Calling up a previously created program.** It is also possible to prepare a file containing *SAS* statements outside of *SAS* using any editor that you wish and then retrieve it into *SAS*. (This is in fact highly recommended.) This file should have the filetype .sas. To open a text file full of *SAS* commands, use the **File** menu options at the top of the *SAS* window. First, select the **Open Program** option under the **File** menu. A new window will appear, in which you may select a file from the list of files displayed by clicking on the required file name. If the file you need to access is not in your current **SAS** directory, you need to locate it. After specifying the correct filename, click the mouse on the **Open Program** box at the bottom of the window. Once a file has been opened, you can edit it in the program window.

**Running a SAS program.** Once a program is complete then you may submit the program for execution. To submit a program you should be in the **EDITOR** window. Go to the **Run** menu and select the **Submit** option to run the program. You can submit the entire program or just highlight a section to submit. The program should produce output in the **LOG**, **RESULTS**, and **OUTPUT** window. The **LOG** window displays the SAS program statements as they are executed and any error/warning messages from SAS. If you don't get the output you expect, then you should check the **LOG** window to see what went wrong. If the program ran as expected, then the actual numerical results are in the **OUTPUT** window. Any graphical results (scatterplots, histograms) show up as separate windows. A list of the results produced can be found in the **RESULTS** window; this can be used to navigate through the output. If there has been an error you can edit the program and then resubmit. If you lose the program (i.e., delete it by accident and then want it back), then you can use the Recall Last Submit option from the **Run** menu. Use Recall multiple times to get back to earlier versions of a SAS program.

**Getting data for your SAS program.** A SAS program starts with a DATA step that sets up the data

for analysis (see the SAS Language Guide handout for more on the DATA step). Here we focus on getting the data into SAS. Data for your *SAS* program may be included within the same text as your program, preceded by a “CARDS;” statement (see details in the SAS Language Guide handout). This is rarely done however as it is only practical for very small data sets.

More often the data to be analyzed is available in the form of a text file (.txt), Excel spreadsheet (.xls), or comma-delimited file (.csv). There are two ways to retrieve data from a file. First, you can read it in as part of the program using an INFILE statement. Enter INFILE '*filename*'; before the INPUT statement in your program, where *filename* (enclosed in single quotes) refers to a file name if it is in the current directory or a path name if the file is in another directory. The FILENAME statement can be used to associate a SAS “abbreviation” with the longer file name. The “abbreviation” can then be used as a shorthand reference to the file in *SAS* programming statements that access external files, such as INFILE or FILE. For example, the following program reads data lines from a file named shopping.txt (assumed to be in the current working directory) identified with the *fileref* FOOD and creates a *SAS* dataset named MARKET:

```
FILENAME FOOD 'SHOPPING.TXT';
DATA MARKET;
INFILE FOOD;
INPUT DAY MMDDYY MEAT GROC DAIRY;
RUN;
```

The second approach to retrieving data from a file uses the Import Data option under the **File** menu. You can import many formats but we will primarily use .xls (Excel spreadsheet) or .csv (comma delimited file). Use Import Data to to navigate to the file. You should place the data in your WORK directory (when prompted) and give it a name. This is now saved as a SAS dataset. You can access it in SAS programs as WORK.name. For example you can use 'SET WORK.name;' in a DATA step or use the 'DATA=WORK.name' option in any procedure.

## Other Options for Running SAS

**Batch programming.** This section may date me (it's old-style computing) but you will also find it useful if you are using BDOC and having trouble with X11. Here you prepare a SAS program in a file, let's say 'program.sas' and then when you log into BDOC you just type 'sas program' to the prompt. This will run the program and produce files 'program.lst' (containing the OUTPUT file) and 'program.log' (containing the LOG file). One difficulty here is graphics - you can't get the pretty X11 graphics this way but you can get line graphics (see the SAS language guide handout for more details on this).

**ANALYST.** There are alternatives to writing SAS programs in the Display Window Manager or for batch programming. One is called ANALYST and the other is Enterprise Guide. After you open SAS there is a “command” box (it is the one with the check mark next to it directly under the File menu. You can type “Analyst” in this box and then hit the Enter/Return key. You can also open Analyst by going to Solutions and then selecting Analysis and then selecting Analyst. ANALYST will open a new “project” and show you a blank data table. You need to load data now. If you have already imported a data file then you can easily load it into ANALYST. Just select OpenBySASName under the file menu and navigate to the data set. If not, you will have to import one using the instructions given above. Data analysis is window-driven in ANALYST. You will find many common tasks in the Graph and Stat menus. Each will lead to a box that must be filled in, e.g., identifying the variable that is the response and the variable that defines the two samples. SAS does the programming for you! ANALYST will evidently not be available in future SAS releases.

**Enterprise Guide.** The latest innovation from SAS Institute is something called “Enterprise Guide”. I don't know very much about it but, like ANALYST, it makes data analysis relatively easy and window-based. I don't believe you are likely to run into it using the ICS lab or your own copy of SAS or BDOC.

## **Saving/Printing Files from SAS**

**Saving files.** To save a text file during a *SAS* session, make sure you are in the relevant window select the **Save as** option under the **File** menu. You can now specify the name of the file you want to save the information in. It will be saved in your current directory. If you want the file to be saved in a different directory, you need to navigate around until you find the right place. Note that you may save files from any of the three primary windows (**EDITOR, LOG, OUTPUT**) thus enabling you to save copies of the *SAS* program, the *SAS* log and the *SAS* output. If you want to save a graph, then I recommend that you save using the Export option (right click on the graph and choose Export) to save as a JPG or GIF which you should be able to use in your writeup. One difficulty is that the SAS files are all text files which can be difficult to incorporate in a document. You can write SAS directly to a .rtf file (which can be read by MS Word). This is done by adding a first line ‘ods rtf file = “c:\directory\hw.rtf” ;’ before your SAS commands and a last line ‘ods rtf close;’ after them.

**Printing files.** You will need to save files to the storage medium of your choice (likely a memory stick) and print at a location that works for you. Be careful before printing an output file because *SAS* listings are extremely long with lots of blank space. You will likely want to edit before printing (but see hint above about writing output directly to a .rtf file).

## **Exiting SAS**

To end your current *SAS* session, use the **File** menu and select **Exit**.