Introduction to SAS

Philip Dixon, 8 Jan 2018

Flavors of SAS

SAS has been around since the early 1970's. It is now available in many different flavors. Four versions are widely available now. I will describe how to use "classic" SAS . SAS Enterprise Guide and SAS Studio are graphical project organizers that interface to "classic" SAS. Enterprise Guide is old and probably being phased out in favor of SAS Studio. Some of the consulting students use SAS Studio and really like it. If you like graphical project organizers, I encourage you to explore what it can do. SAS Inc. has recently made SAS University Edition available. That is discussed briefly below.

On most installations, the '.sas' type of file is linked to EG. That means if you double click the file name, you get EG. If you want to recreate what I do, you have to start "classic" SAS then open the file of SAS commands.

Accessing SAS:

- SAS is installed on the computer you're using
  Most of the public computer labs on campus have SAS installed locally on the computer. Many research groups have SAS installed on one or more computers. Usually these computers have a SAS icon on the desktop. A list of public computer labs is available by going to `http://www.it.iastate.edu/labsdb/` and typing SAS into the Search box.

  If you want to install your own copy of SAS, visit `https://stat.iastate.edu/statistical-software-` and go down to the Student/Faculty/Staff Windows SAS Installations (New Licenses & Renewals) section.

- Access SAS on the stat terminal servers.
  The stat dept. maintains terminal servers that provide remote desktop access to SAS. Your computer becomes a display client; the actual computing is done on the server. Your cyfiles space is linked to the server when you log in, so you can move files back and forth. The computer labs in Gilman and Snedecor automatically link to these terminal servers. The terminal servers are accessible from off campus if you have a VPN connection.

  If you are off-campus:

  - Start a VPN (Virtual Private Network) connection. See
    `https://www.it.iastate.edu/howtos/vpn` for information on how to download and install AnyConnect VPN, then how to startup a VPN connection.
  - Start a Remote Desktop Connection (RDC) to `sas.stat.iastate.edu`. On a Windows PC, this is Accessories/Remote Desktop Connection. A Mac user can download and install the client. Instructions are in the Terminal Server SAS portion of `https://stat.iastate.edu/statistical-software-sas/`.

– You should see a login screen. This is to login into the server. You should see SAS as on icon on the desktop. If you don't use the Start Menu / All Programs to start SAS. You want either version 9.3 or version 9.4. You can also put the SAS icon on your desktop. Your desktop top is saved between RDC sessions.

Important: If you want the remote server to have access to a local printer, a flash drive, or the clipboard (for cut and paste), you need to set up access **before** you connect. You can't do it later. To do this, look at the RDC dialog box and click the Show Options button. Then click local resources and select what you want access to: printers, clipboard, or local drives (click More then Drives) then click Connect.

- Access SAS through SAS University Edition.
  SAS UE is a new offering from SAS. It is a web front-end that provides access to SAS. You write and edit code in a browser window. When you want to run that code, it is sent to a server at SAS Inc. The results are sent back to your browser. The browser front-end is available for Windows, Mac, and Unix boxes. See `www.sas.com/en_us/software/university-edition.html` for details and information on how to get access to UE. I don't have any experience with UE. I'm happy to help but I'll be learning too.

- Notes:
  More information about accessing SAS is available from the Stat Dept. IT group:
  `https://stat.iastate.edu/statistical-software-sas/`.

  This year we will be using SAS version 9.3 or 9.4. These have a couple of major chaanges from earlier versions. If you use SAS on one of the Statistics servers (what we will do in lab), you will be running 9.3 or 9.4. If you use SAS in a departmental computing room or office, you may be running an earlier version. If so, I strongly suggest you ask your IT folks to upgrade.

  There is almost always more than one way to do something in SAS. I will describe the approach that works for me. If you see a different PROC used or a different sort of DATA step in some other source, it is almost certainly just a different way to accomplish the same thing.

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 sort of analyses you want

3. Tell SAS to execute those commands

4. Look at the LOG window to see whether SAS understood your commands

5. If you made a mistake, you need to:

   (a) edit your program
   (b) resubmit all or part of it

6. Look at the output and interpret it.

After you start SAS, you will see a screen full of windows. Your commands go in the program editor window. The log window contains information about what SAS did and any errors or warnings. In version 9.3 and later, your results appear in a Results Viewer window.

Default directory location:
I find it very helpful to organize my work in folders/directories. I store data, SAS programs, and output in that directory. Right after I start SAS, I change the default directory to that for the current project. The working directory is shown in the lower right part of the frame of the SAS window. For many computers, that will be something like `C:\ProgramFiles\SASHome\SASFoundation\9.4`. To change that, double left click on the directory name, navigate to the appropriate directory, then click OK.

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 401/402/500 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 'flat-file by rows' format, i. e.:

- Each row contains one observation.

- Data values are separated by spaces (one or more).

- 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 2 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 (or a CARDS line or a DATALINES line) and an INPUT line.

```
data tomato;
 infile 'tomato.txt';
 input group $ yield;
 run;
```

This creates a SAS data set called tomato. The data come from the file 'tomato.txt' in the default directory. If you don't change the default directory (procedure described above), you need to include the full path name to the file, e.g. `f:\tomato.txt` or `C:\Users\pdixon\Documents\Stat401\tomato.txt`.

Each row has two values. The first value will be stored with the variable name GROUP; the second will be called 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 $ **after** the variable name on the input line. If you omit the $, you get lots of errors (value 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. Some procs produce new SAS data sets. I will introduce appropriate commands as needed throughout the semester.

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

proc boxplot;
  plot yield*group;
```

```
  title 'High resolution boxplots for each group';
run;
```

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

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

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

HINT: What to do when 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, then click submit. You should get the output from your last proc; step.

Working with SAS

To start SAS: Click on the SAS icon.

The windows:
Four windows will appear on your screen.

- Results/Explorer: Results provides a point-and-click way to navigate directly to a specific result. I find it usually easier to page up and page down through the output window. Explorer provides a point-and-click interface to SAS data sets. This can be handy when you want quickly glance at a data set.

- 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 will appear here. Various informative messages will appear here.

- OUTPUT: This window contains the output from your SAS commands.

- Starting in version 9.3, something called ODS HTML is turned on by default. As a result, the Results Viewer window will be added to your screen and your output will appear there instead. This is also saved as an HTML (.htm) file.

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

If you use Word or other standard word processor to edit your file(s), you need to save the files in text format (.txt) using SaveAs. Or, cut and paste from the Word document to the program editor window. SAS can not read files saved as Word documents (.doc or .docx files).

Loading a class program The class web site will have examples of SAS programs. You can download them using your choice of browser, save them, then load them into SAS using FILE/OPEN.

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, rename the file (highlight the file name, right click, rename) and replace the '_sas.txt' with '.sas'. This is a windows/Internet Explorer problem.

Running SAS code Typing code into the editor window doesn't run the code. To run it, make the window with your code active by clicking somewhere in the window. Or, right click on the name of that window in the list at the bottom of the screen. Then click the running person icon on the main menu bar (submit), or select Run / Submit from the main menu. If you click the running person icon and a new program editor window opens up, your code window wasn't active.

To run only certain lines of code: highlight the lines you want by shift-clicking or mousing over code while holding down the left mouse button. Then submit (the running person icon).

When you submit code, the code is copied to the log window along with any feedback from SAS. If the code is good, you'll get information about what SAS did. If it wasn't, you'll get errors. The results will appear in the results viewer window. If you don't get any results, check the log for errors.

Saving a file To save the contents of any window to a file, under the *file* menu in the desired window, click on *save as...*,type in (or click on) where you want to save it, name the file, and click *OK*.

Reminder: always check the pathname to make sure it will save it where you want to save it. If you haven't changed the default directory, your work will be saved in the SAS program directory, which may not be easy to find. And, you may not have appropriate permissions to save something in that directory.

Saving graphs or other output:
This can be done various ways. I find the easiest is to store all output in both results viewer and in a `.rtf` file that is easily readable by WORD. You can then copy and paste any figure or table into another document. You create a `.rtf` file by including the command `ods rtf file='name.rtf';` before anything you want saved and the command `ods rtf close;` after everything you want saved. When you run the code including those commands, SAS will save the output in the named file. After executing the `ods rtf close;` command, SAS will open up WORD and display the document.

The creativity.sas file demonstrates this.

Note: If a file is open in WORD, no other program can write to that file. That means if you have already run SAS and the file has been opened in WORD, you need to close WORD before you can generate a new output file. If the file is still open, SAS will run but the log file will have an error

and the rtf file will not be updated.

Exiting SAS: File / Exit from the main menu. SAS will ask you if you really want to quit (you probably do) and then if you want to save any unsaved command files (you probably do).