

1 Using SAS outside of FITCs

You can access SAS from off campus by using the FITC Virtual Desktop. Go to

itc.uiowa.edu

and click on Virtual Desktop. You will be prompted to assent to an agreement and to enter your HawkID and password. Select the SAS 9 folder, and from it, SAS 9.1.3. SAS will come up on your screen. When you are finished, you must both exit from SAS and log off from the Virtual Desktop. When you run SAS this way, although you can interact with it on your screen, it is actually running on a computer on campus. Consequently, SAS can't access files that are on your hard drive unless you specify drive letter mapping. In this lab, I will show you the most reliable way to get data into SAS when you are running it on the Virtual Desktop.

2 Getting started in the FITC

FITC computers will display a login screen. Users should enter their HawkID and password in the spots provided. Students can find their HawkID and default passwords on the SIS system.

3 Downloading files from the course web page

Bring up a web browser, either Firefox or Internet Explorer. Enter the address of my web page in the location box.

www.stat.uiowa.edu/~kcowles

Then click on "Course homepages" and "22S:30/106."

Click on "Datasets," and when the next screen appears, click on the underlined link "Datasets." Three types of files may be accessed:

- Files ending in .dat are data files for your use with the software package SAS.
- Files ending in .info contain descriptive information about datasets.
- Files ending in .txt are datasets for use in a different class with a different software package.

Note that this list is case-sensitive, with all files with names beginning with capital letters appearing before all other files.

If you are going to work with SAS on the Virtual Desktop, it is easiest to copy and paste the data from the web page directly into the SAS program. If you are running SAS on the computer at which you are sitting, it is preferable to download the dataset into a file and tell SAS where to find the file. Below, I will show you how to have SAS access data both ways.

To download a data file for use in this lab:

- Right click the file name (use the right, not left, mouse button)
- In the dialog box that opens, left click "Save target as" (use the left mouse button)
- In the "save as" dialog box chose My Computer/Local Disk C:/Temp. Click "Save."
- (If you preferred to save the file on your own disk in drive A: so that you could use it on a different computer later, you would move to drive A: in the dialog box before saving.)

Left click the "billion.info" file to read a description of the "billion" dataset. Click "Back" to get back to the list of datasets. Then download the file "billion.dat" according to the directions above.

4 Other useful features on the course web page

Return to the main course web page and click "Web resources." Note that there is a directory of the locations and hours of all the campus FITCs.

Lecture notes, homework assignments, and lab handouts are posted under "Handouts." These are in a format that may be read and printed in most FITCs.

5 Accessing SAS

To access SAS on your own laptop, go to

<https://virtualdesktop.uiowa.edu>

Log in with your HawkID and password; then choose SAS, and SAS 9.2.

To access SAS in the FITCs, click on Start/All Programs/SAS/SAS 9.2. This actually takes you to the Virtual Desktop as described above.

In either case, you will get a screen that shows:

- a menu bar
- a log window
- a program editor window

6 Entering commands and programs

Click in the program editor window. You may now type commands and programs in this window.

7 How SAS programs and commands are organized

Use a *DATA* step to organize your data by creating a SAS dataset. Then use *PROC* steps or automated features to analyze your data. Once you have created a SAS dataset, you may apply any SAS procedures or automated features to it during the SAS session without recreating the dataset.

DATA and PROC steps consist of SAS statements. Each statement must end with a semi-colon. Most statements include one or more keywords that must be spelled exactly as shown.

8 The DATA step: Creating a SAS dataset

Before it can process data, SAS must read in the data in the form of a table with

- a row for each observation
- a column for each variable

You must choose a name for the entire dataset and a name for each variable. SAS has the following rules for names:

- SAS names must begin with a letter or an underscore. The remaining characters in a SAS name can be letters, numbers, or underscores. There must be no embedded blanks.

SAS distinguishes between two types of variables: numeric variables, which contain only digits and decimal points and with which arithmetic operations may be done, and character variables (all other kinds of data).

9 Controlling print width

Put this line at the beginning of every SAS program if you want output to print correctly on 8-1/2 by 11 inch paper:

```
options linesize = 75 ;
```

10 Reading data in from an existing dataset

You have saved the file "billion.dat" in the "temp" directory. Use an "infile" statement to tell SAS to use it.

```
data billion ;          + gives dataset a name for SAS ;
infile 'c:\temp\billion.dat' ; + tells SAS where the data is ;
input with age region $ ; + names the variables in each row ;
                        + $ after region identifies character vbl ;
run ;                  + end of data step ;
```

11 Copying data directly into the data step

In this alternative method to get data into SAS, do not use an "infile" statement. Instead use the "datalines" statement, and copy and paste the data in immediately following it. Put a semicolon all by itself on the line after the last line of data.

```
data billion ;          + gives dataset a name for SAS ;
input with age region $ ; + names the variables in each row ;
                        + $ after region identifies character vbl ;

datalines ;
37 50 M
24 88 U
14 64 A
-
-
-
1 50 E
1 . E
1 . 0
;
run ;                  + end of data step;
```

Type one of these data steps into the program editor window. To make SAS run these statements and create the dataset, use the mouse to highlight the block of statements and then click on the lion of the running man.

SAS will use the log window to tell you what it has done. Be sure to check the log window for any error messages. If any errors are reported, click in the program editor window to make it active. Correct the errors in the code and then rerun the block of code.

Note: If you wanted to read in the file from your own disk in the A: drive the infile statement would be:

```
infile 'a:billion.dat' ;
```

12 Using SAS procedures to list and tabulate the dataset

Once the dataset is created, you may run SAS procedures to analyse it. To list the entire dataset:

```
proc print data = billion ;
run ;
```

To get a frequency distribution of the regions in which billionaires lived:

```
proc freq data = billion ;
tables region ;          * tables is a keyword; region is the name of
                        * the variable for which you want counts ;
run ;
```

The output is:

The FREQ Procedure

region	Frequency	Percent	Cumulative Frequency	Cumulative Percent
A	38	16.31	38	16.31
E	80	34.33	118	50.64
M	22	9.44	140	60.09
D	29	12.45	169	72.53
U	64	27.47	233	100.00

13 Proc univariate: SAS workhorse of descriptive statistics

Use `proc univariate` for quantitative variables when you want the following:

- means
- medians
- quartiles
- 5-number summary
- stem plots (for small datasets) or histograms (large datasets)
- boxplots

```
proc univariate plot data = billion ;
var with ;
run ;
```

The output is:

The UNIVARIATE Procedure
Variable: with

Moments			
N	233	Sum Weights	233
Mean	2.68154506	Sum Observations	624.8
Std Deviation	3.31884032	Variance	11.0147011
Skewness	6.57544276	Kurtosis	56.9655987
Uncorrected SS	4230.84	Corrected SS	2555.41064
Coeff Variation	123.785972	Std Error Mean	0.21742446

Basic Statistical Measures

Location		Variability	
Mean	2.681545	Std Deviation	3.31884
Median	1.800000	Variance	11.01470
Mode	1.000000	Range	36.00000
		Interquartile Range	1.70000

Tests for Location: Mu0=0

Test	-Statistic-	-----p Value-----	
Student's t	t 12.33323	Pr > t	<.0001
Sign	M 116.5	Pr >= M	<.0001
Signed Rank	S 13630.5	Pr >= S	<.0001