Stat 404
Lab 8

All files and data will be published in BbLearn to download and save or on the shared (S:) drive (you need to have mapped the uidaho drives to your computer). The address is:

When using SAS: S:\Courses\stat-renaes\Stat404\
When using R: S:/Courses/stat-renaes/Stat404/
Something to remember is that I will have example code for you to use to help you write your programs in labs.

Lab collection:

When I decide to collect a lab, I will let you know at the beginning of the lab and it will be due within 1-2 class periods of the lab.

What to copy and paste into the document for submission:

All code from both programs and the log window from SAS (please clean it up a bit and get rid of errors if there are any). Do not paste the results from either program unless I specify in the exercises.

Submission:

Is through BbLearn. Go to the Labs link, click on the lab and follow instructions to attach the file. The file MUST be in PDF format. No other formats will be allowed for submission. The easiest way I find to create a PDF is to do all the work in Word or Pages then "save as" PDF format.

R:

1. Read in the data set called 'shoes_trackers.sas7bdat' using the sas7bdat package. Name your new data set shoes. Look at the dataset and watch for missing values.

2. Use the command fix( ) to bring up the spreadsheet to fix values.

3. Change the following:
   a. Product_Category change or input to Shoes
   b. Supplier_Country must be either GB or US.
   c. Supplier_Name must be 3TOP Sports, or Greenline Sports Ltd.
   d. Supplier_ID must be 2963 14682

4. Close the spreadsheet from fix( ) and it will save your changes. Print the data to the console to make sure that the changes stuck.

5. Read in the data set called 'staff.sas7bdat' using the sas7bdat package. Name your new data set increase.

6. Create two new variables called Increase and NewSalary.
    * Increase is Salary multiplied by 0.07
    * NewSalary is the sum of Salary and Increase

7. Create a new dataset called increase2 that includes only the following variables: Employee_ID, Salary, Increase and NewSalary.

8. Run the following command in the console:
    data(faithful)
    attach(faithful)

9. Graph both eruptions and waiting in:
    a. Boxplots (one boxplot per variable)
    b. Histograms (one histogram per variable)
    c. Histograms with normal density curves overlayed on them (your color of choice)

SAS:

1. Use and INFILE and INPUT statements to read in the csv file called "old faithful.csv" that is on the S: drive. Look at the file before reading in to see what needs to go in your input statement. You will also need to specify on your INFILE line that the delimiter is a comma and that the first observation starts on line two (firstobs=2).

2. Use a PROC PRINT statement to print the data to look at it. You should have 3 variables and 272 observations.

3. Use PROC SGPLOT to create histograms of the variables current and interval. You can choose to put a density curve over the histograms or not (I have no preference either way).

4. Create a 2D scatterplot (with either GPLOT or SGPLOT) with x=current and y=interval.

5. Next, create two time series (line plot) with Index as your x-axis variable and each of the other variables (one at a time). Use either GPLOT or SGPLOT.

6. For submission: from R, copy the code only; for SAS, copy code and the log (please clear it and run the program one last time before you copy the log results (ctrl-E will clear the selected window)). Paste code and log into your document and upload to BbLearn. No output results (there are WAY too many observations)! ☺