Stat 404
Lab 5

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.

One recommendation I have is to create a few folders on your computer for this class. I would create R, SAS, Data, Labs. That way you can save all the files and data you need for labs and it will make things easier for later labs.

SAS:

1. Create a library (your name of choice up to 8 characters) that refers to the following address:
   S:\Courses\stat-renaes\stat404\

2. Look at the following data files in your library called sales and nonsales using PROC CONTENTS. Example to help is (with me naming my library herc – you'll put your library name in place of mine):
   PROC CONTENTS data=herc.sales;
   run;
   PROC CONTENTS data=herc.nonsales;
   run;
   What are the names of the two variables that are different in the two data sets?

3. Create a new data set called allemployees by concatenating these un-like structures and using the RENAME option so that the variables that are different in nonsales are renamed to match the ones in sales.

4. Use PROC PRINT to view the new data set to verify it has 400 observations.

R:

1. The address to access the data is: S:/Courses/stat-renaes/Stat404. Read in all of the following data files: mnth72007.csv, mnth82007.csv and mnth92007.csv.
   While I showed examples and you worked with read.table( ), I usually use read.csv( ) whenever I have .csv files (which is almost always). Type ?read.csv into the console and see the help file for the syntax; it is really similar to read.table( ) but no sep= option is necessary.

2. Create a new dataset by merging the three datasets you read in in step 1 by the Custom ID variable. Do two different datasets, one that is sorted and one that is not sorted.

3. Look at the merged dataset (either the whole thing or use head( ) to just see the first 6 observations).