**AML 610 Fall 2014 Homework #3**

**Submit all files to** **smtowers@asu.edu****.**

**Due Wed Sep 17th, 2014 at noon.**

**Please submit with name format hwk3\_<first name>\_<initial of last name> Please provide your R file, and a Word file that gives the output to your R screen, plots, etc.**

**All code must conform to good coding practices, as described in** [**http://sherrytowers.com/2012/12/14/good-programming-practices-in-any-language/**](http://sherrytowers.com/2012/12/14/good-programming-practices-in-any-language/) **and all plots must conform to good plotting practices, as described in** [**http://sherrytowers.com/2013/01/04/good-practices-in-producing-plots/**](http://sherrytowers.com/2013/01/04/good-practices-in-producing-plots/)

**Question 1)**

Create a vector myarray with the following numbers in the given order:

73 110 92 93 109 128 141 106 70 151

Suppose that we originally forgot to type a value 50 into the vector, between the values 109 and 128. Read the R help file for the append() command. Find a way to insert this value into the vector, between the 109 and 128, using the append() command.

**Question 2)**

Solve question 1), but this time use the R combine function, c()

**Question 3)**

Read the R help file for R’s Binomial random number generator, rbinom(). Using set.seed() set the seed for the random number generator to 78123.

a) Generate a vector, mybinom, of 1000 Binomial random numbers, generated with n=120 and p=0.8. Print the mean, standard deviation, variance, and standard error of mybinom. Read the Wikipedia page on the Binomial distribution to remind yourself about its properties. What are the expected values of the mean, SD, variance, and SE? Do you get exactly those values? If not, why?

b) Print the first 4 elements of mybinom that have exactly three digits. Now print the indices of those elements.

c) Print the first 4 elements of mybinom that are divisible by 3. Now print the indices of those elements.

d) Print the **number** of elements of mybinom that are divisible by 5

e) Sort mybinom in descending order, and put the results in a new vector, mybinomb. Copy the first 100 elements of that to mybinomc. Print every 5th element of this vector to the screen, starting with the first element.

f) Read the R help file for R’s table() function.

Plot the table of mybinom. You should get a plot that looks like the one below. Ensure that your plot is appropriately labeled.



g)What does as.numeric(table(mybinom)) give you?

 What does as.numeric(names(table(mybinom))) give you?

 Based on the answers to the above two questions, print to the screen the most probable value in mybinom

**Question 4)**

Read in the file [www.sherrytowers.com/influenza\_summary\_ny\_nj.txt](http://www.sherrytowers.com/influenza_summary_by_region.txt) to a data frame.

The file contains the CDC confirmed influenza cases by week for each influenza season from 2006-07 to 2012-13 in NY and NJ, and for three influenza subtypes, B, AH3N2, and AH1N1

The “year” field contains the maximum of the year for each season, thus “year” for the 2006-07 season is 2007.

The field “day\_relative\_to\_janfirst” contains the date, expressed in days relative to Jan 1st of each season.

Loop over the years, and select the data corresponding to each year.

Plot the influenza AH1, AH3, and B weekly incidence for each year:

* Divide the plotting area up into 4 regions.
* In the first region you will plot the AH1N1, in the next AH3N2, and in the third B.
* In the fourth plotting region you will put a legend to show the different colored lines for each year.

Note that to make a blank plot over which you can overlay a legend use

 plot(1, type="n", axes=F, xlab="", ylab="")

Your plot should look like this:



Comment on how this plot can be improved. For 10 bonus points, improve it.

Subset the dataframe, selecting the 2012-13.

Using Google, look up how to remove a column from a data frame; remove the fluB column.

Then read the R help on the write.table().

Output this subset of the data frame to a file.

Open the file in an editor, and copy and paste the contents of the file into the doc file containing your homework solutions.