**BIA 6311: Introduction to R for Data Analysis**

**Post Workshop Practice Assignment**

The Wine file is a set of wines that were judged and placed into three quality categories (A the best, C the least). The actual names of the wines were removed. However the chemical composition of each wine is listed. This data set is a great to data mine to see if you can predict quality category based on chemical composition. We will use it to practice some R functions.

Using the Wine.csv file:

1) Create a read.csv script to read in the file to R. Call the new dataframe called “Wine”.

2) What are the dimensions of the dataset?

3) Summarize all the numeric variables.

4) Create a new variable in the wine dataframe called log\_ alcohol. Make it the natural log of the alcohol variable.

5) Create a new categorical variable called high\_ Magnesium. If Magnesium is less than 100, the value should be “low”. If Magnesium is greater than or equal to 100, the value should be “high”.

6) Create tukey’s 5-numbers of the Magnesium variable**.**

7)Create a scatter plot of Malic\_Acid and Ash. Add a trend line. Add labels to the plots

8) Create 3 boxplots (on the same graph) of Proline by Type. Add labels to the plots. Make the boxes blue.

9) Create a new dataframe called Wine\_A, where the only data is Type “A” AND Magnesium > 100. Show the first 5 rows of this new dataframe.

10) Aggregate/summarize the data of the average of alcohol and proline BY type (you should get a table). Use Xtable function to make it look nicer.

11) Create a correlation matrix of the numeric variables.

12) Create a covariance matrix of the numeric variables.

13) Create a histogram and a qqplot of Malic\_Acid. Based on the results, is this variable approximately normally distributed? Put your conclusion in as a comment.

14) Use a t-test to determine if there is a significant difference in alcohol between “high” and “low” Magnesium levels. Put your conclusion in as a comment.

15) Please knit your R script into an .html or .doc output.