

Name \_\_\_\_\_

Thursday, September 24<sup>th</sup>, 2015

The data file is **countries.csv** and it is located at: <http://www.cs.uni.edu/~jacobson/4772/q> for your downloading an R processing use for this quiz. That is q as in Quiz for the URL.

Produce a scatter plot of the relationship between births per 1000 population in 1992 and female life expectancy. This is output #1 from the quiz. Put it into a Word document.

Add a funky line to your scatter plot. This is output #2 from the quiz.

Add a linear model or regression line to your scatter plot. This will be output #3 from the quiz.

Grouped scatterplot problem will be a scatterplot of the birthrate and the lifeexpf variables again, but this time split by the develop variable, separating **developed** countries from **underdeveloped** (developing) countries. This will be output #4 from the quiz.

Create a data frame that consists of just the AFRICAN countries. The R statement you used to do this is:

Create a data frame that consists of just the European countries. The R statement you used to do this is:

Create a data frame that selects just all those countries that have a life expectancy female of strictly less than 55 years. The R statement you used to do this is:

Create a dataframe that consists of just the lifeexpf, birthrate, region, and develop variables. The R statement you used to do this is:

Do an analysis of variance comparing **lifeexpf** for Europe, for North America, Africa, and Asia. You will have **FOUR** different groups, so you will have to recode region into fewer categories. Show all of the R code necessary to do this. You can create a new variable, such as continent if you like. You do NOT have to recode region into different values, but can recode it into a new variable such as continent or regionRecoded or regionRec.

***This homework assignment will be due on Thursday, October 1<sup>st</sup> at the beginning of class.***