

## MATH/STAT 3930: Intro to Stats, Fall 2016

### Project 2.

Due Friday, March 15 during class.

---

I need a print-out of your project. 20% of your grade will be solely based on appearance of the printed out version. Make sure that all of your graphs have meaningful axes and labels. Your grade will primarily be determined by the hard copy you turn in. There will be a 3% penalty per day, or part of day, for late assignments.

This project counts for 10% of your final grade in this class (for comparison, each homework counts for about 1%).

---

Consider the 2015 world data sheet that was sent to you.

1. Clean up the data set in the following ways, at a minimum.
  - Provide meaningful column names.
  - Convert all numerical data to numeric, and all categorical data to factors.
  - Convert the NA symbol from the data to the NA symbol recognized by R.
  - Create a categorical variable that encodes the continent that each country is in.
  - Create a categorical variable that encodes the region that each country is in.
  - Pull out the summary data for each continent, and save in a second data frame. Delete the corresponding rows from the original data frame.
  - Pull out the summary data for each region, and save it in a third data frame. Be sure to delete the corresponding rows from the original data frame.
2. I have emailed you the column that you are supposed to focus on for the purposes of this assignment.
  - Examine your data column. Do you see any outliers? Does it appear normal? Did you expect it to be approximately normal before you started? Provide much information here.
  - Plot your column versus the other columns in the data set. Interpret.
  - Give examples of data columns that are correlated with the data in your column.
  - Give examples of data columns that are uncorrelated with the data in your column.
3. Decide on data that you think is related to the column that you have been assigned. Find data of that type and add it to your data frame. Plot the data and determine whether the data is correlated.