Name: Jordon Tourville

UmichID: 79911480

What dataset are you working with: US Births 2000-2014, Senate polls

List 3 questions that you can ask with your dataset.

Q1: Is there a significant difference between months of the year in terms of total births?

Q2: Are more babies being born in later years than earlier years?

Q3: Is there a correlation between polling average and election result for each election year?

List the associated null hypothesis for each question:

Q1: There is no significant difference between the months of the year and the number of babies born in each month.

Q2: There is no significant increase in the number of babies being born in later years than in earlier years.  
Q3: There is no significant correlation between polling average and election result for each of the years.

What statistical test(s) will you use to answer each of the questions:

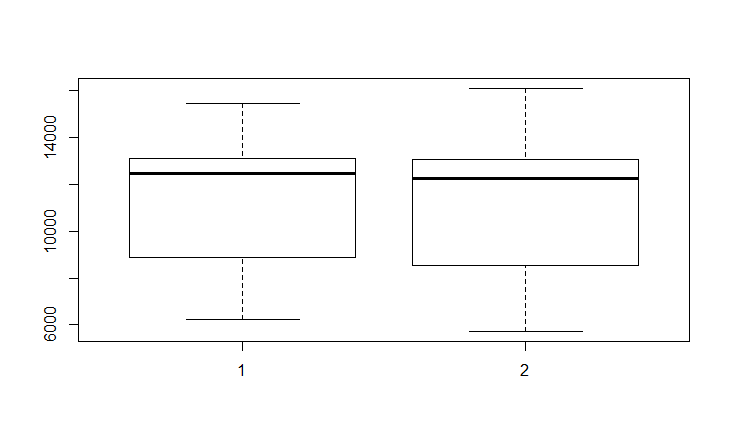
Q1: One-way ANOVA

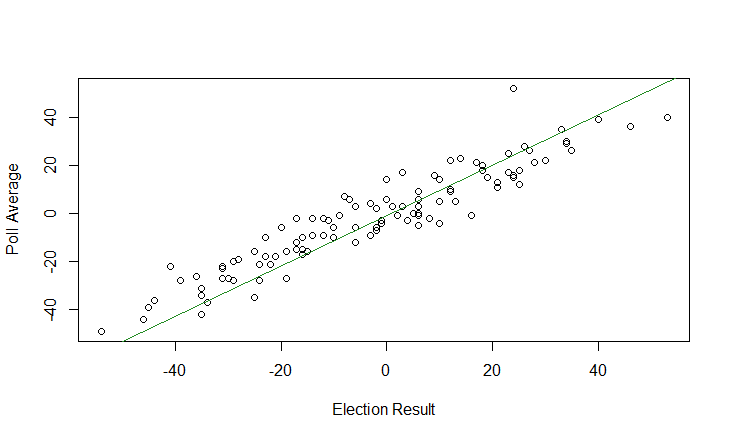
Q2: One sample t-test

Q3: Correlation analysis, linear regression

Make a visual plot showing the relationship that you will analyze statistically (e.g. boxplot for t-test or ANOVA; scatterplot for regression; table for chi-square).

Q1:

Q2: 

Q3: 

Do your data meet the assumptions required for the statistical test you want to run? Please state the assumptions you examined and whether or not your data meet those assumptions:

Q1:

Q2: Normality: yes, fits qqplot.

Q3: Normality: yes, fits qqplot.

Run the statistical test! Put your results here:

Q1:

Q2: p=0.003

Q3: r=0.9377175, p<0.0001

Interpret your results!

Q1:

Q2: There were significantly less babies born in the later years (2007-2014) than in the earlier years (2000-2006).

Q3: There is a strong positive correlation between the polling average of senators, and the results of the associated election.