Name: Sophia Paul

UmichID: sophiakp

What dataset are you working with: congress\_age

List 3 questions that you can ask with your dataset.

Q1: Are the congresspeople from one chamber older than those from another?

Q2: Is there a difference in the ages of incumbents and non-incumbents in congress?

Q3: Are Republicans usually older than Democrats? (I took a subset to remove all the other parties from the data for this test)

List the associated null hypothesis for each question:

Q1: There is no significant difference between the ages of senators and representatives.

Q2: There is no significant difference in the ages of incumbents and non-incumbents in congress  
Q3: There is no difference in between the ages of Democrats and Republicans in congress

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

Q1: A t-test

Q2: a t-test

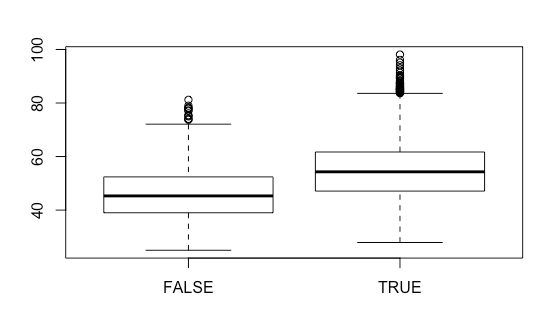
Q3: a t-test

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).

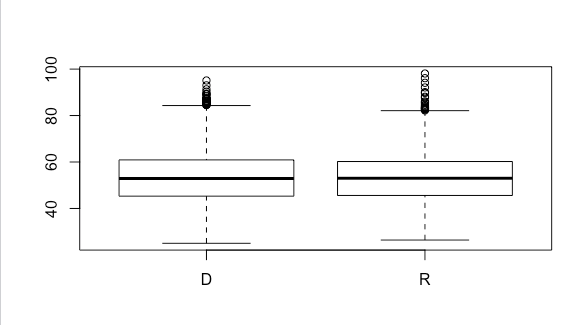
Q1: boxplot(age~chamber,data=congress\_age)



Q2: boxplot(age~incumbent, data=congress\_age)



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:

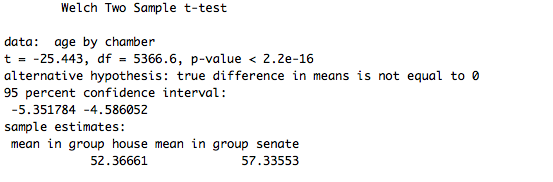
Q1: Since the data set is very large, I’m not worried about distribution. I did test for variance and it is not standard

Q2: I tested for standard variance and the variance is standard

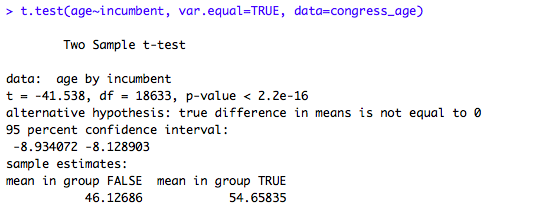
Q3: The variance is standard

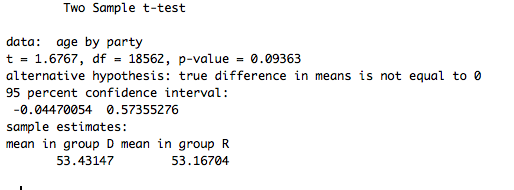
Run the statistical test! Put your results here:

Q1:



Q2:



Q3:

Interpret your results!

Q1: Senators are, on average, older than representatives

Q2: Incumbents are older than non-incumbents

Q3: There is a slight difference in between the ages of Democrats and Republicans in congress.