Problem Set #1

1. GSS Web Question
   1. 2.22
   2. Agree 44.6%
   3. Descriptive. There is no analysis, it shows the different responses and the percentages of the responses are present with visual representation, but again, the is not analysis and the summary covers all responses not a sample of the responses.
2. Trump Poll Question
   1. 2,808,605. The people who voted in Missouri
   2. 1,941 is the sample taken from the overall population
   3. 57.23% of people in the sample who voted for Trump
   4. 56.8% is the parameter of the percentage of the entire population who voted for Trump
3. Scale of Measurement
   1. Ordinal
   2. Nominal
   3. Ordinal
   4. Interval
   5. Interval
   6. Nominal
   7. Interval
   8. Nominal
   9. Nominal
   10. Nominal
   11. Ordinal
   12. Interval
4. Box Plot
   1. boxplot(AvgAge , ylab="Ages" , main="Average Ages Between Men and Women" , col = 4)
   2. It is not likely to be a random sample. The median and mean values are not the same suggesting this is not a normal distribution.
5. Discrete vs Continuous
   1. Discrete
   2. Discrete
   3. Discrete
   4. Discrete
   5. Continuous
   6. Discrete if we’re only using 0, 1, 2, 3 etc., but if we’re going for exact like how many years at a given point in time, it’s continuous.
   7. Discrete
   8. Discrete
   9. Discrete
   10. Discrete
   11. Discrete
   12. Argument can be made for both. It works on a decimal system so not just 2.00 and 3.00, but 2.37 and 3.44. With that being said, that means there are still only 400 possible GPAs.
6. No Response
7. <https://www.cnn.com/election/2016/results/exit-polls/california/president>
   1. This is a visual breakdown of the exit polls in 2016 from California’s presidential election. These are represented in almost all discrete variables. They are yes and no questions about Trump and Clinton’s decorum as well as the ages of the different people voting blocked off into set ranges. Income is a variable which would be continuous, but again it’s blocked off into set numbers. Is your income between 50k and 100k etc. There are inferential and descriptive variables here. One question is that of age, but there is no analysis, and another is a question of whether or not people are bothered by Clinton’s email scandal. The response implies more information that just what are the people answering the poll like.