R- Assignment 4

Text-

1. What value for the “position” argument in a geom\_point() function enables you to see all the observations in a dataset, even if they would otherwise overlap?

position=”jitter”

1. In ggplot, how do color, size, and shape aesthetics behave differently for categorical data than for continuous data?

When you map a color, size, or shape to continuous data all of the data points will be changed to match. When you map a color, size, or shape to categorical data, the data points that align with the categories will be changed.

R Script/Images-

1. Create a graph showing the frequency distribution of the categories in eye\_color of

characters in the starwars dataset. Which eye color is most common?

* 1. Create a graph showing the frequency distribution of the height of characters in the starwars dataset. Use 8 bins.
  2. Create a graph of boxplots to summarize the distribution of height, by species. Which species has the widest variance in height (the largest bar in the bar plot)?
  3. Using the data in the *mtcars*, create a scatterplot of the relationship between *hp* (horsepower) and *mpg*.
  4. Create another scatterplot of the relationship between *hp* (horsepower) and *mpg* with an added linear trendline.
  5. Create scatterplots between *hp* (horsepower) and *mpg*; differentiate the two categories of the variable *am* (transmission type) by color. Each of the two values in *am* should have its own trendline.
  6. Create scatterplots (with linear trendlines) between *hp* (horsepower) and *mpg*; differentiate the two categories of *am* (transmission type) by using facets with the facet\_wrap() function. Don’t forget to add trend lines within each facet.
  7. Which of the graphs in parts c) or d) are easier to interpret, if any?