1. Load and attach the “mtcars” dataset. Conduct an exploration of what variables it contains and what the values look like (**if you don’t know about any variables, find out what they are. If you don’t know about a car, use Google to find out what it is. Have a comprehensive understanding of what the cars are in your dataset**). (0.5 Points)

It contains 32 observations and 11 variables:

1. Use the variables “mpg” and “cyl” to perform K-means clustering on the cars (get a subset of data including only “mpg” and “cyl” variables for the k-mean clustering). Run kmeans function with K=3 and 4 respectively. Which one is better? Use the ratio of betweenness and within sum of squared distance (=betweenness / tot.withinss) (1.5 Points)
2. Compute the mean of each variables by clusters using original data for each model

K= 3

K = 4

1. Drawing a plot to visually represent the clustering analysis results

K= 3

K = 4

1. Get a subset of data including “disp” and “hp” variables. Using the subset of data, run kmeans clustering analysis with K= 3, 4, 5, and 6 respectively. Which value of K is better now among 4 models? **Use the Elbow method for this one**.(2 Points)
2. You have now used two sets of inputs (“mpg” and “cyl”; “disp” and “hp”). Which of the two sets of inputs gives better clusters? (**Hint : You need to use the same evaluation metric to compare them. You can restrict to comparing for K=3 and 4**) (2 Points)