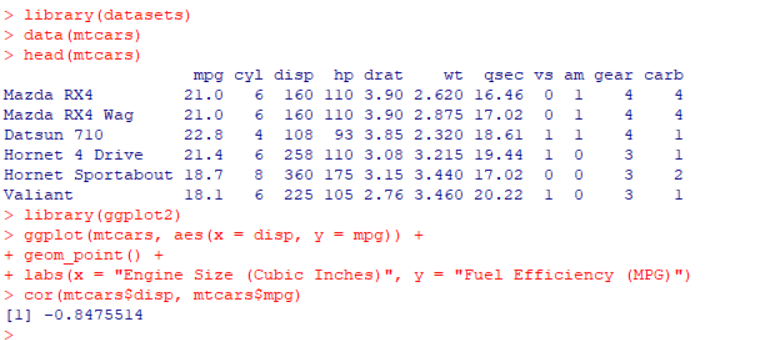
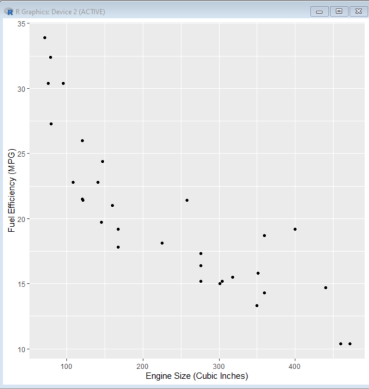
**ITA0448 – STATISTICS WITH R PROGRAMMING FOR VECTORIZED EXPRESSIONS**

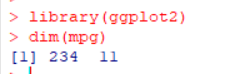
**SUBMITTED BY: DEVADARSHINI J REG NO: 192124092**

**DATE: 23/03/2023**

1. Use the built-in dataset mtcars for this assignment. Do cars with big engines use more fuel than cars with small engines? You probably already have an answer, but try to make your answer precise. What does the relationship between engine size and fuel efficiency look like? Is it positive? Negative?

1. How many rows are in mpg? How many columns?



1. Which variables in mpg are categorical? Which variables are continuous?

The mpg variable provided my the ggplot2 dataset contains 11 variables:

* ‘manufacturer’: manufacturer of the car (categorical)
* ‘model’: model of the car (categorical)
* ‘displ’: engine displacement in liters (continous)
* ‘year’: model year (categorical)
* ‘cyl’: no of cyl in the engine (categorical)
* ‘trans’: type of transmission (categorical)
* ‘drv’: drive train (categorical)
* ‘cty’: city miles per gallon (continous)
* ‘hwy’: highway miles per gallon (continous)
* ‘fl’: fuel type (categorical)
* ‘class’: type of car (categorical)

1. Take the first faceted plot in this section:

ggplot(data = mpg) +

geom\_point(mapping = aes(x = displ, y = hwy)) +

facet\_wrap(~ class, nrow = 2)

1. What are the advantages to using faceting instead of the colour aesthetic?
2. What are the disadvantages? How might the balance change if you had a larger dataset?

The advantages of using faceting instead of the color aesthetic include:

* Faceting creates separate plots for each group, which can make it easier to visually compare the relationships between the x and y variables within each group.
* Faceting allows for a larger number of groups to be displayed in a single plot, without becoming too cluttered or difficult to read.
* Faceting can help to avoid potential issues with color blindness or other visual impairments that may make it difficult to distinguish between colors.

The disadvantages of using faceting include:

* Faceting can sometimes make it more difficult to compare the relationships between the x and y variables across different groups, as each plot is scaled independently.
* Faceting can also lead to a loss of context, as the individual plots may not provide a clear sense of the overall trends or patterns in the data.
* With a larger dataset, faceting may become more challenging to implement effectively, as the number of groups may become too large to display clearly in a single plot, requiring more plots or different faceting strategies. Additionally, the plots may become too small to interpret easily, leading to issues with readability and visual clarity.

1. What geom would you use to draw a line chart? A boxplot? A histogram? An area chart?

In R programming, the commonlyused geoms to draw the specified types of charts:

* Line chart: geom\_line()
* Boxplot: geom\_boxplot()
* Histogram: geom\_histogram()
* Area chart: geom\_area()

