Lab_7_Assignment

Raghul V_21122048

17/03/2022

Importing mtcars data from R package

```
df<-data("mtcars")
```

Getting the first 6 rows of the data (head of the data)

```
head(mtcars, 6)
```

```
##                    mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

Getting the summary of the dataset

```
summary(mtcars)
```

```
##       mpg             cyl             disp             hp
##  Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0
##  1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5
##  Median :19.20   Median :6.000   Median :196.3   Median :123.0
##  Mean   :20.09   Mean   :6.188   Mean   :230.7   Mean   :146.7
##  3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0
##  Max.   :33.90   Max.   :8.000   Max.   :472.0   Max.   :335.0
##       drat             wt             qsec             vs
##  Min.   :2.760   Min.   :1.513   Min.   :14.50   Min.   :0.0000
##  1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000
##  Median :3.695   Median :3.325   Median :17.71   Median :0.0000
##  Mean   :3.597   Mean   :3.217   Mean   :17.85   Mean   :0.4375
##  3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000
##  Max.   :4.930   Max.   :5.424   Max.   :22.90   Max.   :1.0000
##       am             gear             carb
##  Min.   :0.0000   Min.   :3.000   Min.   :1.000
```

```
##  1st Qu.:0.0000   1st Qu.:3.000   1st Qu.:2.000
##  Median :0.0000   Median :4.000   Median :2.000
##  Mean   :0.4062   Mean   :3.688   Mean   :2.812
##  3rd Qu.:1.0000   3rd Qu.:4.000   3rd Qu.:4.000
##  Max.   :1.0000   Max.   :5.000   Max.   :8.000
```

Getting the total number of rows

```
nrow(mtcars)
```

```
## [1] 32
```

Getting the total number of columns

```
ncol(mtcars)
```

```
## [1] 11
```

Checking for null values in mpg column

```
is.na(mtcars$mpg)
```

```
##  [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

Checking the null values in the dataset

```
is.na(mtcars)
```

```
##                       mpg   cyl  disp    hp  drat    wt  qsec    vs    am  gear
## Mazda RX4           FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Mazda RX4 Wag       FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Datsun 710          FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Hornet 4 Drive      FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Hornet Sportabout   FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Valiant             FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Duster 360          FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Merc 240D           FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Merc 230            FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Merc 280            FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Merc 280C           FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Merc 450SE          FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Merc 450SL          FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Merc 450SLC         FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
## Cadillac Fleetwood   FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Lincoln Continental FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Chrysler Imperial    FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Fiat 128             FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Honda Civic          FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Toyota Corolla       FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Toyota Corona        FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Dodge Challenger     FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## AMC Javelin          FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Camaro Z28           FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Pontiac Firebird     FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Fiat X1-9            FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Porsche 914-2        FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Lotus Europa         FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Ford Pantera L       FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Ferrari Dino         FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Maserati Bora        FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## Volvo 142E           FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##                       carb
## Mazda RX4            FALSE
## Mazda RX4 Wag        FALSE
## Datsun 710           FALSE
## Hornet 4 Drive       FALSE
## Hornet Sportabout    FALSE
## Valiant              FALSE
## Duster 360           FALSE
## Merc 240D            FALSE
## Merc 230             FALSE
## Merc 280             FALSE
## Merc 280C            FALSE
## Merc 450SE           FALSE
## Merc 450SL           FALSE
## Merc 450SLC          FALSE
## Cadillac Fleetwood   FALSE
## Lincoln Continental  FALSE
## Chrysler Imperial    FALSE
## Fiat 128             FALSE
## Honda Civic          FALSE
## Toyota Corolla       FALSE
## Toyota Corona        FALSE
## Dodge Challenger     FALSE
## AMC Javelin          FALSE
## Camaro Z28           FALSE
## Pontiac Firebird     FALSE
## Fiat X1-9            FALSE
## Porsche 914-2        FALSE
## Lotus Europa         FALSE
## Ford Pantera L       FALSE
## Ferrari Dino         FALSE
## Maserati Bora        FALSE
## Volvo 142E           FALSE
```
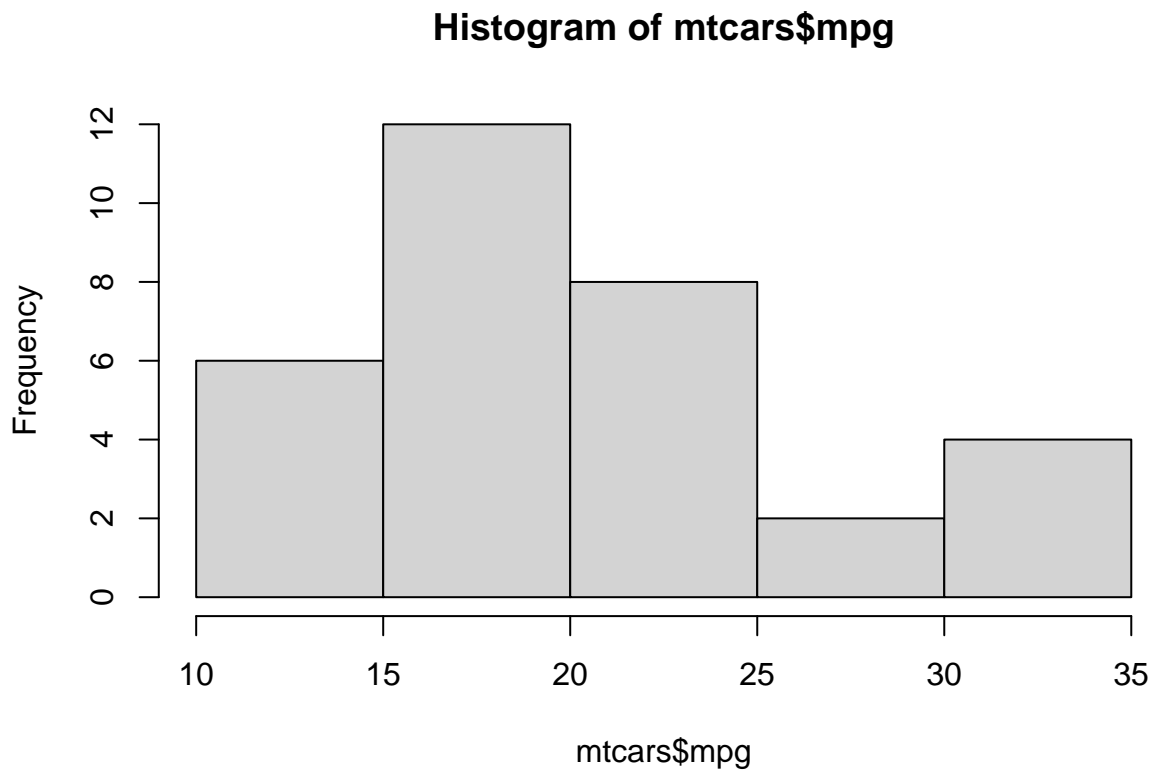
Getting the sum of null values in the total dataset
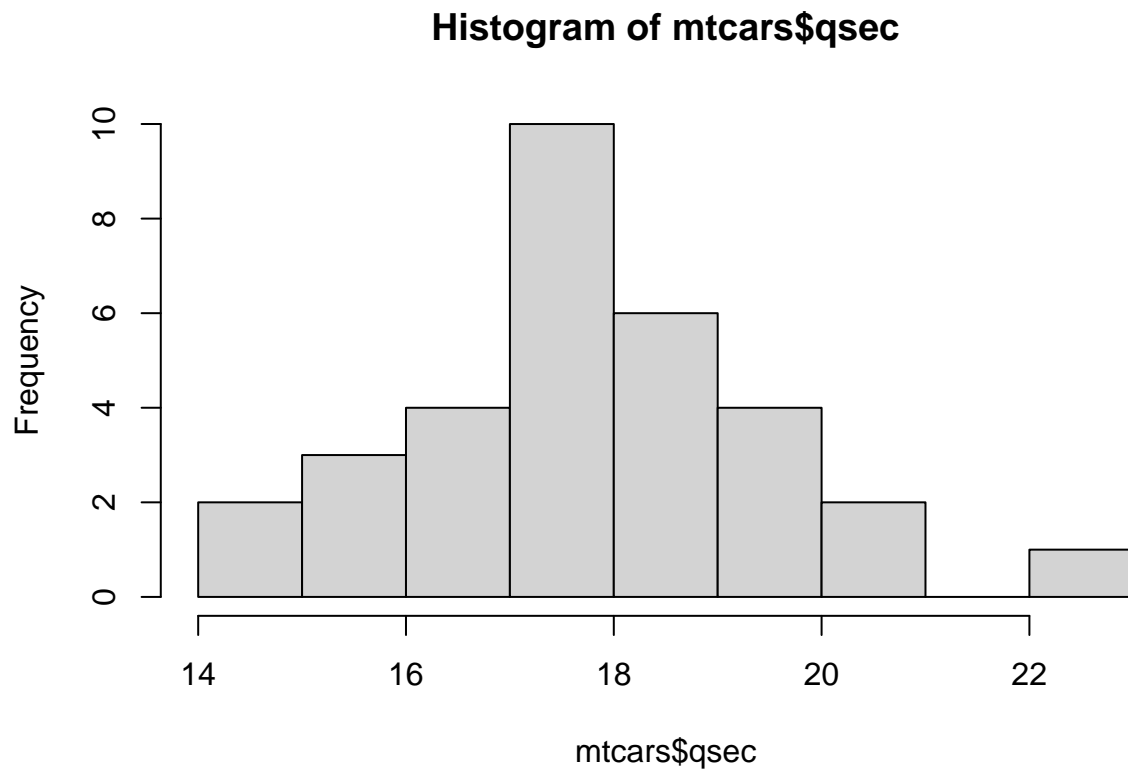
```
sum(is.na(mtcars))
```

```
## [1] 0
```

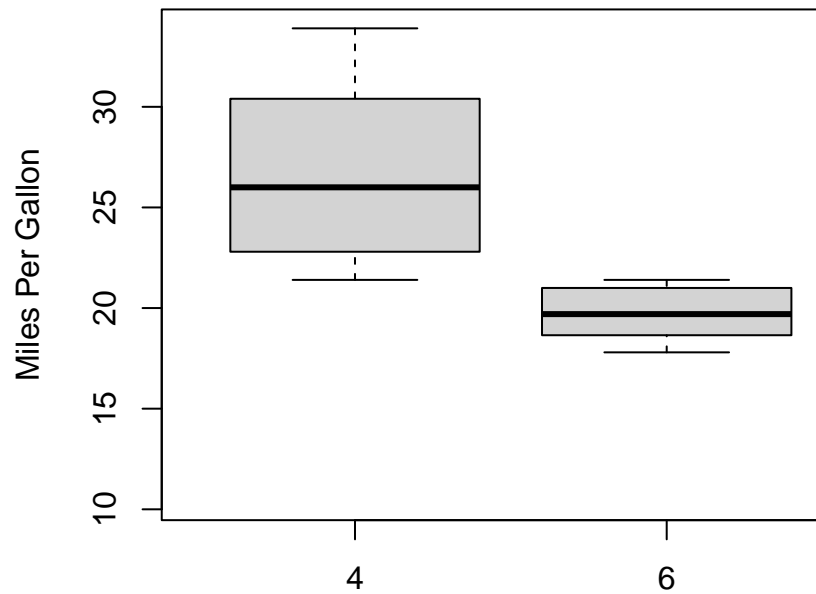Creating a histogram diplaying the mpg

```
hist(mtcars$mpg)
```

Histogram of mtcars$mpg

Creating a histogram displaying the qsec

```
hist(mtcars$qsec)
```

Histogram of mtcars$qsec


```
boxplot(mpg ~ cyl, data = mtcars, xlab = "Number of Cylinders",
    ylab = "Miles Per Gallon", main = "Mileage Data")
```

Mileage Data

Miles Per Gallon

Number of Cylinders

Creating a boxplot to display the mileage data

Creating a boxplot using notch to display the mileage data

```r
boxplot(mpg ~ cyl, data = mtcars,
   xlab = "Number of Cylinders",
   ylab = "Miles Per Gallon",
   main = "Mileage Data",
   notch = TRUE,
   varwidth = TRUE,
   col = c("green","yellow","purple"),
   names = c("High","Medium","Low")
)
```

```
## Warning in (function (z, notch = FALSE, width = NULL, varwidth = FALSE, : some
## notches went outside hinges ('box'): maybe set notch=FALSE
```
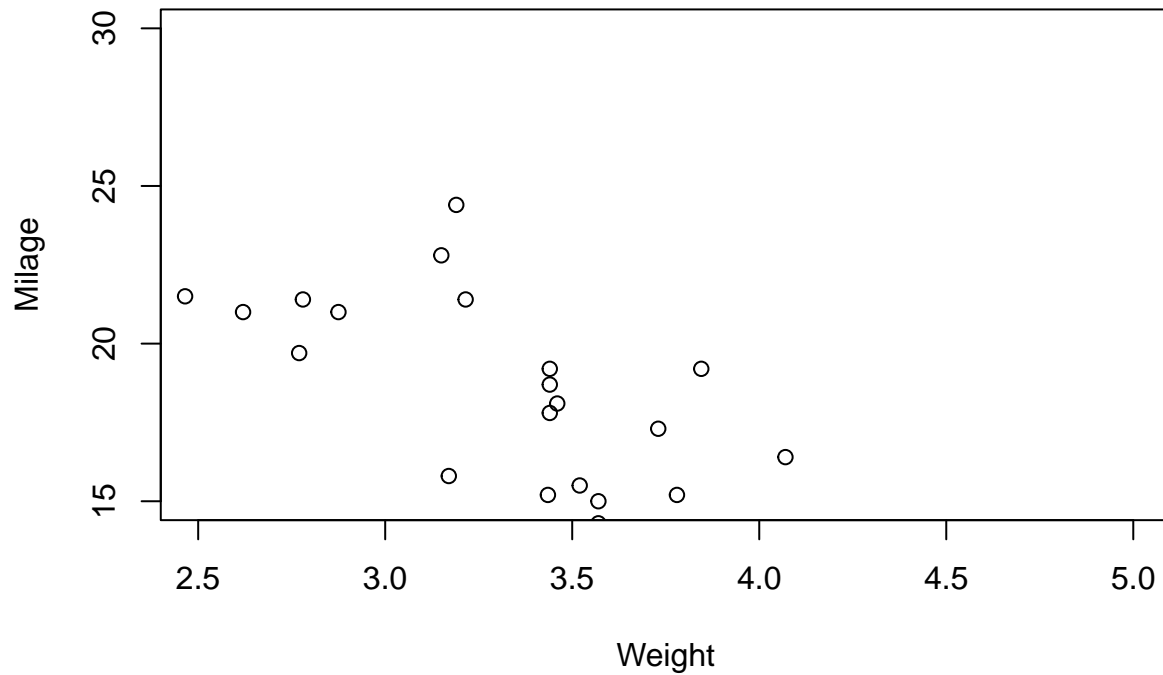
Mileage Data

Number of Cylinders
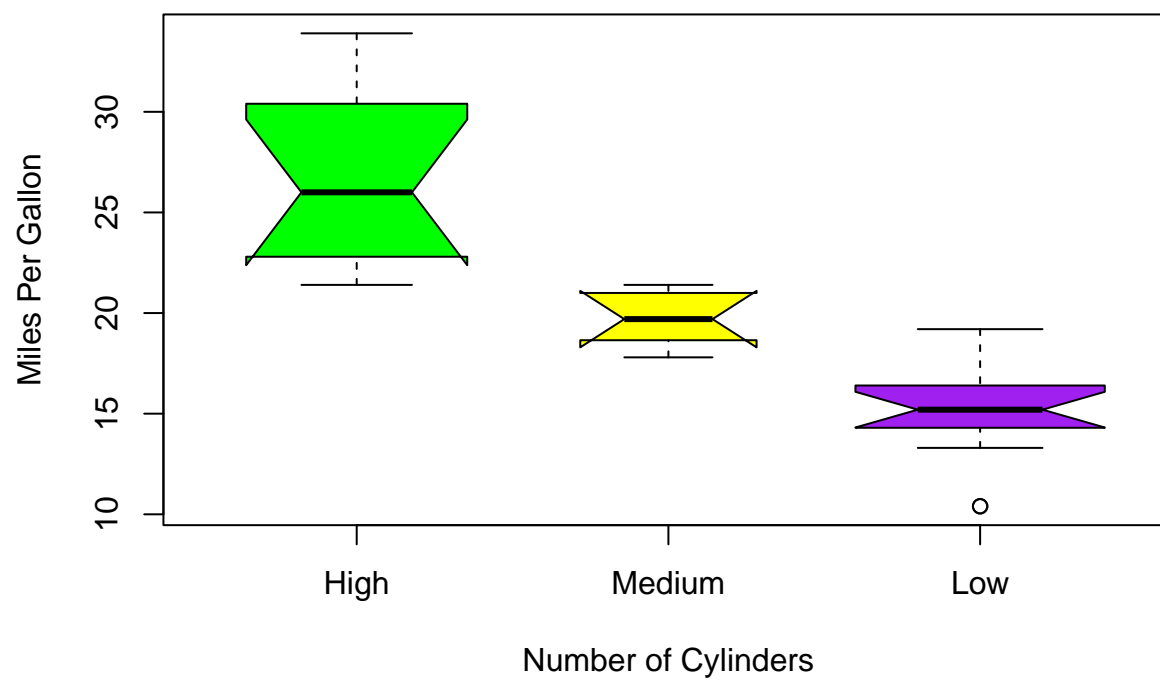
Creating a scatterplot

```
plot(x = mtcars$wt,y = mtcars$mpg,
    xlab = "Weight",
    ylab = "Milage",
    xlim = c(2.5,5),
    ylim = c(15,30),
    main = "Weight vs Milage"
)
```

Weight vs Milage

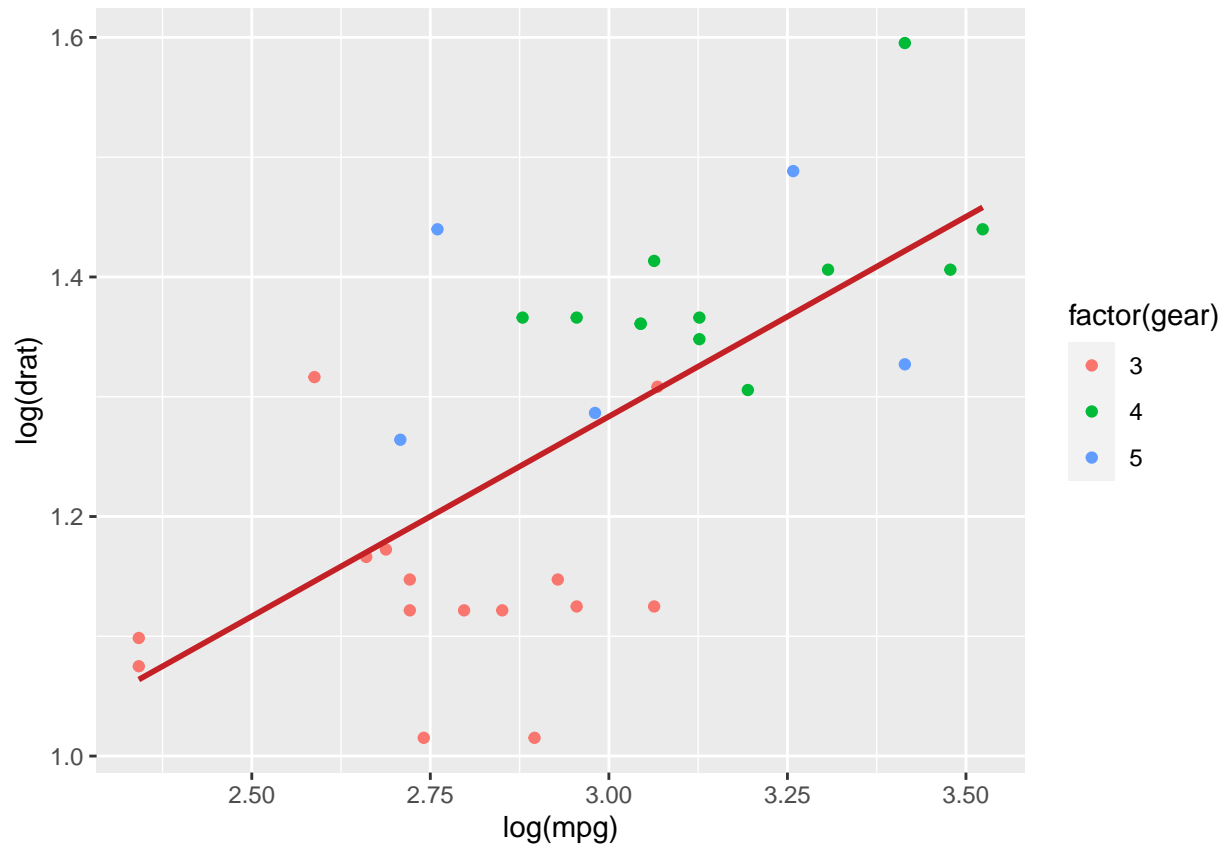importing the ggplot2 library for displaying scatter plot with fitted values

```
library(ggplot2)
```

```
ggplot(mtcars, aes(x = log(mpg), y = log(drat))) +
        geom_point(aes(color = factor(gear))) +
        stat_smooth(method = "lm",
        col = "#C42126", se = FALSE, size = 1
)
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

Summary on the dataset

- The selected data was mtcars that consisted of various car data such as mileage, weight, curb, displacement, horsepower, cylinders etc.
- A simple data exploration was conducted to summarize the data, to find the null values in the dataset.
- Simple histograms were created to visualize the mileage, number-of-cylinders etc.
- Using ggplot2 was able show a simple scatter plot graph with fixed values (i.e: Regression)

Learning

- Exploratory Data Analysis
- Visualizing the data

    - Boxplot
    - Scatterplot
    - Histogram

- ggplot2