hw2_sol

TA

1/9/2021

Q.1

Source: The file contains data extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models). The dataset contains 32 observations on 11 variables.

```
mtcars <- read.table("mtcars.csv",sep = ",", header=TRUE)

head(mtcars)
```

```
##                 model  mpg cyl disp  hp drat    wt  qsec vs am gear carb
## 1          Mazda RX4 21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## 2      Mazda RX4 Wag 21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## 3         Datsun 710 22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## 4     Hornet 4 Drive 21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## 5  Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## 6            Valiant 18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```
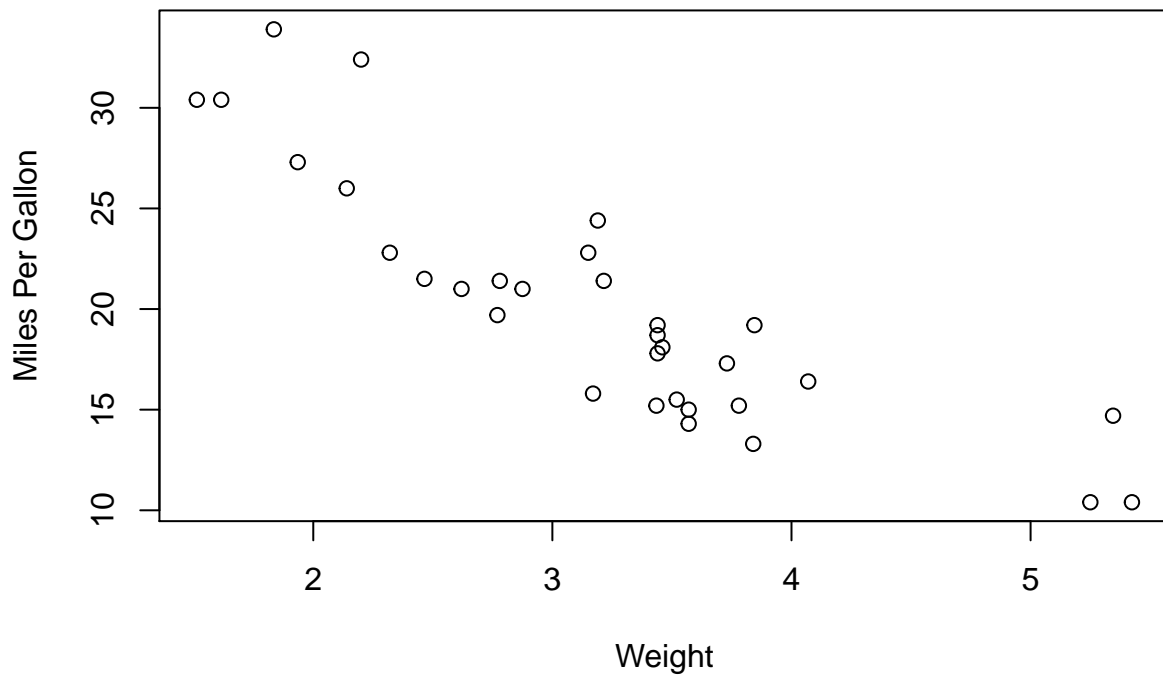
```
# colMeans(mtcars[,2:11])


# apply(mtcars[,2:11], 2, var)

typeof(mtcars)
```

```
## [1] "list"
```

```
plot(x = mtcars$wt, y = mtcars$mpg,
     xlab = "Weight", ylab = "Miles Per Gallon",
     main = "Weight vs. miles per gallon")
```
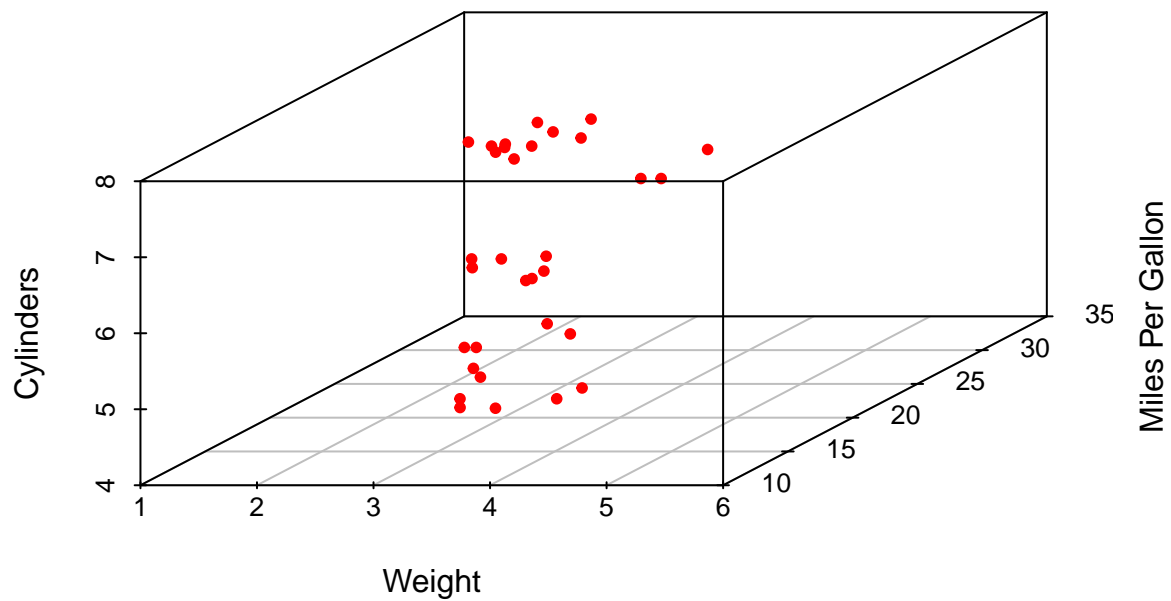
Weight vs. miles per gallon

The scatter plot shows a decreasing linear relationship between mpg and weight which can be investigated further if the relationship is statistically significant.

Q. 2

```
library("scatterplot3d")
scatterplot3d(x = mtcars$wt, y = mtcars$mpg, z=mtcars$cyl,
              xlab = "Weight", ylab = "Miles Per Gallon",   zlab="Cylinders",
              color="red", pch=20)
```

One of the drawback of 3d scatter-plot is that the interpretation of relationships between the variables can be difficult to observe. An alternative can be to plot a 2d catter plot b/w mpg and weight and colorcode the points according to cylinder variable.

Q. 3

```
pairs(mtcars[,2:11], pch=20)
```
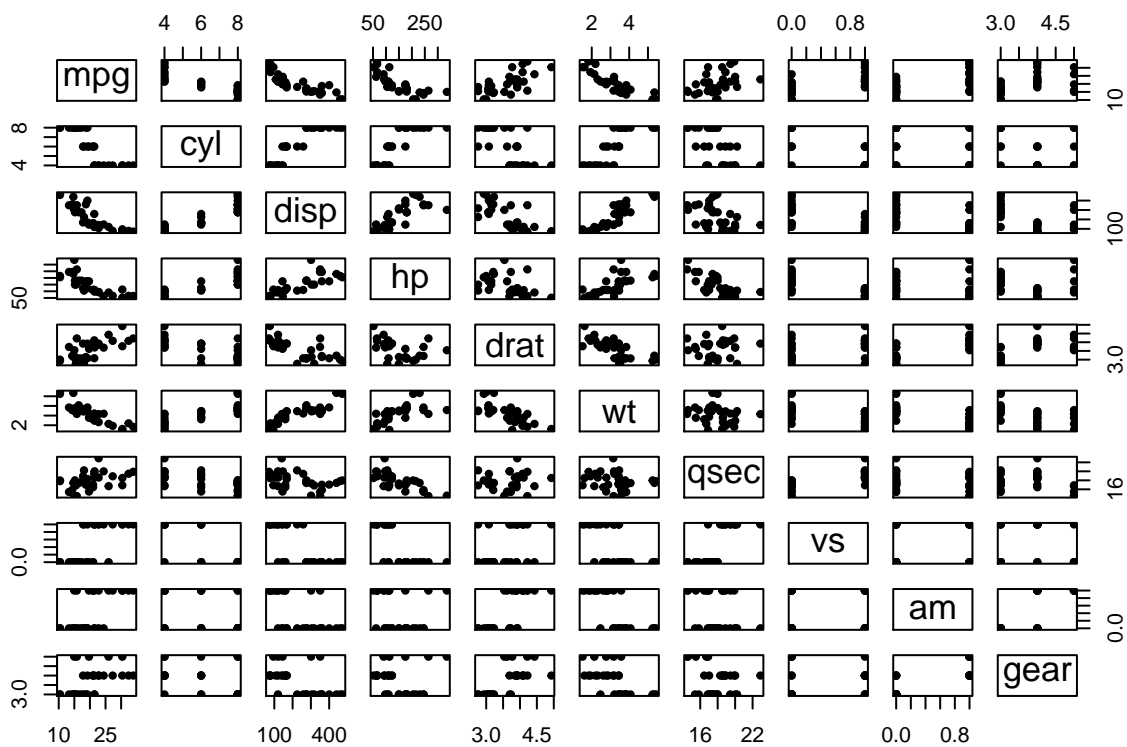
Q.4

If we look at the 2d scatter plot, and color code the points accoring to the number of cylinders, we can see that the points cluster into 3 groups according to the number of cylinders. Therefore, the engineer's suggestions reasonable.

```
library(ggplot2)
ggplot(mtcars, aes(x = wt, y = mpg)) +
    geom_point(aes(color = factor(cyl)))
```