Periandri_Anthony_Assignment1_05-30-2025

Anthony Periandri

2025-05-31

```
## load the csv file
df<-read.csv("C:\\Users\\Antho\\Desktop\\KSU\\Machine Learning\\Assignment 1\\automobile.csv")
## show the data
head(df)
```

```
##   symboling normalized.losses        make fuel.type aspiration num.of.doors
## 1         3                 NA alfa-romero       gas        std          two
## 2         3                 NA alfa-romero       gas        std          two
## 3         1                 NA alfa-romero       gas        std          two
## 4         2                164        audi       gas        std         four
## 5         2                164        audi       gas        std         four
## 6         2                 NA        audi       gas        std          two
##    body.style drive.wheels engine.location wheel.base length width height
## 1 convertible          rwd           front       88.6  168.8  64.1   48.8
## 2 convertible          rwd           front       88.6  168.8  64.1   48.8
## 3   hatchback          rwd           front       94.5  171.2  65.5   52.4
## 4       sedan          fwd           front       99.8  176.6  66.2   54.3
## 5       sedan          4wd           front       99.4  176.6  66.4   54.3
## 6       sedan          fwd           front       99.8  177.3  66.3   53.1
##   curb.weight engine.type num.of.cylinders engine.size fuel.system bore stroke
## 1        2548        dohc             four         130        mpfi 3.47   2.68
## 2        2548        dohc             four         130        mpfi 3.47   2.68
## 3        2823        ohcv              six         152        mpfi 2.68   3.47
## 4        2337         ohc             four         109        mpfi 3.19    3.4
## 5        2824         ohc             five         136        mpfi 3.19    3.4
## 6        2507         ohc             five         136        mpfi 3.19    3.4
##   compression.ratio horsepower peak.rpm city.mpg highway.mpg price
## 1               9.0        111     5000       21          27 13495
## 2               9.0        111     5000       21          27 16500
## 3               9.0        154     5000       19          26 16500
## 4              10.0        102     5500       24          30 13950
## 5               8.0        115     5500       18          22 17450
## 6               8.5        110     5500       19          25 15250
```

```
## Descriptive table showing values for wheel base and length
summary(df[,c("wheel.base","length")])
```

```
##    wheel.base         length
##  Min.   : 86.60   Min.   :141.1
##  1st Qu.: 94.50   1st Qu.:166.8
##  Median : 97.00   Median :173.2
```

```
##  Mean   : 98.85   Mean   :174.3
##  3rd Qu.:102.40   3rd Qu.:183.5
##  Max.   :120.90   Max.   :208.1
```

```
table(df$engine.location)
```

```
##
## front  rear
##   199     3
```

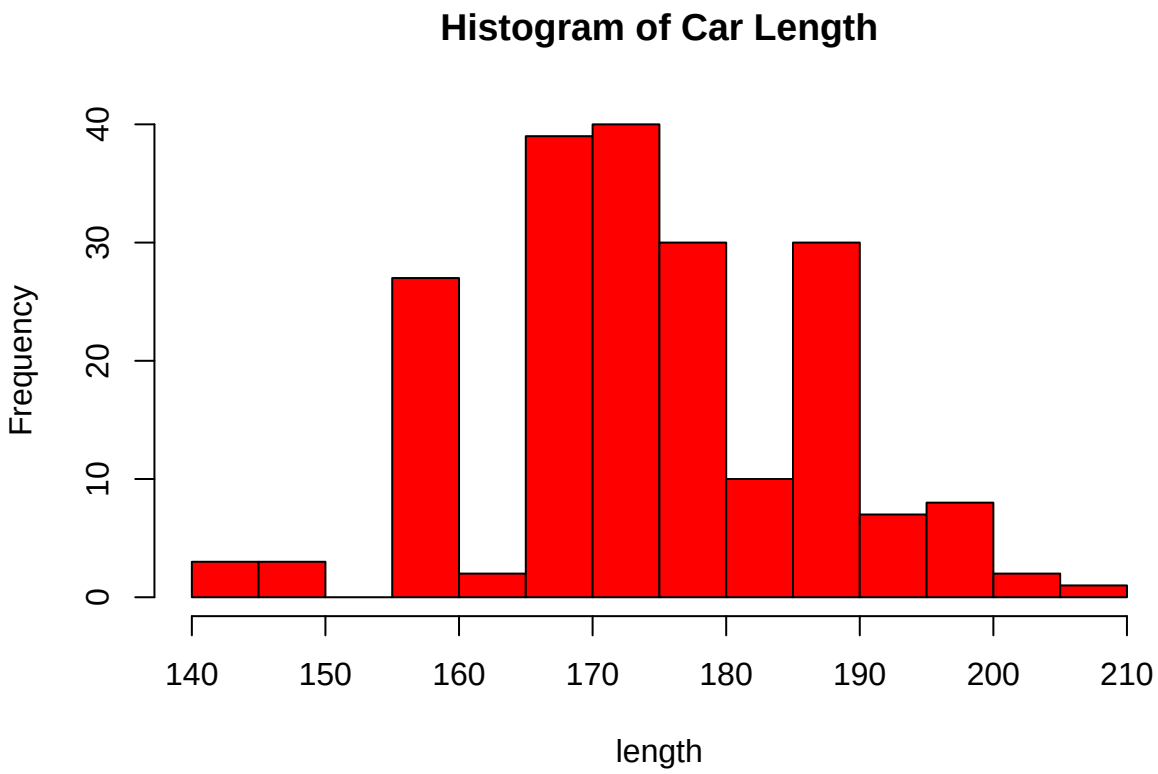
```
table(df$fuel.type)
```

```
##
## diesel    gas
##     20    182
```

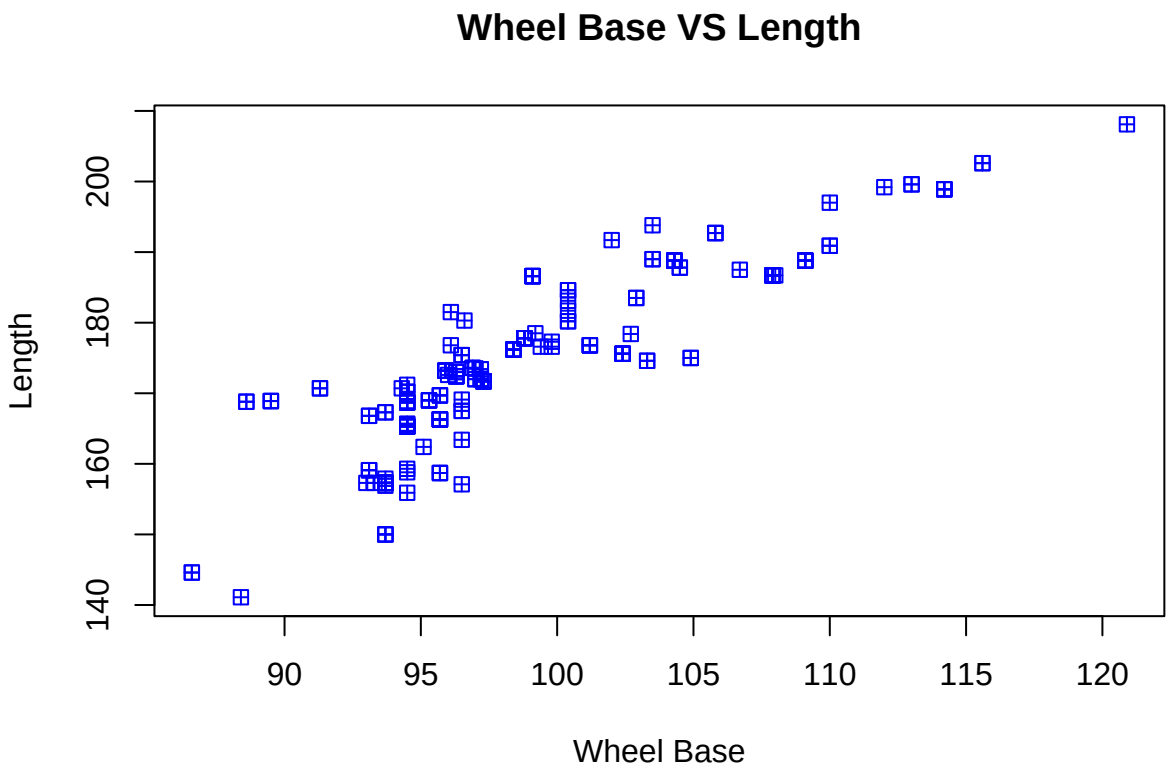
```
## Mutate variable to show the log of lengths of vehicles
df<-df %>%
  mutate(log_Lengthofcar= log(length))
summary(df[,c("log_Lengthofcar","length")])
```

```
##  log_Lengthofcar     length
##  Min.   :4.949   Min.   :141.1
##  1st Qu.:5.117   1st Qu.:166.8
##  Median :5.154   Median :173.2
##  Mean   :5.158   Mean   :174.3
##  3rd Qu.:5.212   3rd Qu.:183.5
##  Max.   :5.338   Max.   :208.1
```

```
## histogram showing quantity of different length vehicle
hist(df$length,
     main = "Histogram of Car Length",
     xlab = "length",
     col = "red",
     breaks = 10)
```




```
## Scatterplot comparing Wheel base and length
plot(df$wheel.base,df$length,
     main = "Wheel Base VS Length",
     xlab = "Wheel Base",
     ylab = "Length",
     col = "blue",
     pch = 12)
```


Wheel Base VS Length
Length
140
160
180
200
90
95
100
105
110
115
120
Wheel Base