Assignment 1

Lokesh

2023-01-29

#Data Source : <https://www.kaggle.com/datasets/yashwanthkumarmn/motorcycles-in-india>

Motorcycles<-read.csv(file="C:/Users/jetan/Downloads/archive (1)/bike\_dataset.csv")  
  
  
#Displaying first six rows of data  
head(Motorcycles,6)

## model\_name price CC mileage type\_of\_bike weight\_in\_kg  
## 1 Gravton Motors Quanta 99000 N/A 320 Electric Bike 100  
## 2 Simple Energy One 109999 N/A 236 Electric Bike 110  
## 3 Okaya Classiq 69900 N/A 200 Electric Bike 95  
## 4 Oben Electric Rorr 102999 N/A 200 Electric Bike 120  
## 5 Ola Electric S1 85099 N/A 181 Electric Bike 121  
## 6 Ola Electric S1 Pro 120149 N/A 181 Electric Bike 125  
## links acceleration\_speed  
## 1 https://www.carandbike.com/gravton-motors-bikes/quanta 4.2  
## 2 https://www.carandbike.com/simple-energy-bikes/one 3.6  
## 3 https://www.carandbike.com/okaya-bikes/classiq N/A  
## 4 https://www.carandbike.com/oben-electric-bikes/rorr 3  
## 5 https://www.carandbike.com/ola-electric-bikes/s1 2.9  
## 6 https://www.carandbike.com/ola-electric-bikes/s1-pro 2.9  
## top\_speed  
## 1 70  
## 2 100  
## 3 25  
## 4 100  
## 5 116  
## 6 116

# Structure of Dataset  
str(Motorcycles)

## 'data.frame': 361 obs. of 9 variables:  
## $ model\_name : chr "Gravton Motors Quanta" "Simple Energy One" "Okaya Classiq" "Oben Electric Rorr" ...  
## $ price : int 99000 109999 69900 102999 85099 120149 114264 71990 67540 53000 ...  
## $ CC : chr "N/A" "N/A" "N/A" "N/A" ...  
## $ mileage : num 320 236 200 200 181 181 180 170 165 160 ...  
## $ type\_of\_bike : chr "Electric Bike" "Electric Bike" "Electric Bike" "Electric Bike" ...  
## $ weight\_in\_kg : int 100 110 95 120 121 125 101 96 73 100 ...  
## $ links : chr "https://www.carandbike.com/gravton-motors-bikes/quanta" "https://www.carandbike.com/simple-energy-bikes/one" "https://www.carandbike.com/okaya-bikes/classiq" "https://www.carandbike.com/oben-electric-bikes/rorr" ...  
## $ acceleration\_speed: chr "4.2" "3.6" "N/A" "3" ...  
## $ top\_speed : num 70 100 25 100 116 116 65 78 42 25 ...

#summary of my data  
summary(Motorcycles)

## model\_name price CC mileage   
## Length:361 Min. : 38000 Length:361 Min. : 8.00   
## Class :character 1st Qu.: 100000 Class :character 1st Qu.: 20.00   
## Mode :character Median : 242000 Mode :character Median : 30.00   
## Mean : 839908 Mean : 44.68   
## 3rd Qu.:1459000 3rd Qu.: 55.00   
## Max. :7990000 Max. :320.00   
##   
## type\_of\_bike weight\_in\_kg links acceleration\_speed  
## Length:361 Min. : 55.0 Length:361 Length:361   
## Class :character 1st Qu.:118.0 Class :character Class :character   
## Mode :character Median :169.0 Mode :character Mode :character   
## Mean :178.8   
## 3rd Qu.:216.0   
## Max. :433.0   
##   
## top\_speed   
## Min. : 25.00   
## 1st Qu.: 79.50   
## Median :100.00   
## Mean : 99.34   
## 3rd Qu.:129.12   
## Max. :200.00   
## NA's :161

#Minimum & Maximum price of Motorcycles from my data  
  
min(Motorcycles$price)

## [1] 38000

max(Motorcycles$price)

## [1] 7990000

#Find mean  
mean(Motorcycles$price)

## [1] 839907.9

#Find Median  
median(Motorcycles$price)

## [1] 242000

#type of Motorcycles electric bike-57 petrol-304  
m<-table(Motorcycles$type\_of\_bike)  
m

##   
## Electric Bike Petrol Bike   
## 57 304

#percentage of electric & petrol bikes  
mpercent<-prop.table(m)\*100  
mpercent

##   
## Electric Bike Petrol Bike   
## 15.78947 84.21053

#number of observations and variables  
dim(Motorcycles)

## [1] 361 9

#summary of particular variable  
summary(Motorcycles$mileage)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 8.00 20.00 30.00 44.68 55.00 320.00

summary(Motorcycles$price)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 38000 100000 242000 839908 1459000 7990000

summary(Motorcycles$model\_name)

## Length Class Mode   
## 361 character character

#glimpse  
dplyr:: glimpse(Motorcycles)

## Rows: 361  
## Columns: 9  
## $ model\_name <chr> "Gravton Motors Quanta", "Simple Energy One", "Okay…  
## $ price <int> 99000, 109999, 69900, 102999, 85099, 120149, 114264…  
## $ CC <chr> "N/A", "N/A", "N/A", "N/A", "N/A", "N/A", "N/A", "N…  
## $ mileage <dbl> 320, 236, 200, 200, 181, 181, 180, 170, 165, 160, 1…  
## $ type\_of\_bike <chr> "Electric Bike", "Electric Bike", "Electric Bike", …  
## $ weight\_in\_kg <int> 100, 110, 95, 120, 121, 125, 101, 96, 73, 100, 85, …  
## $ links <chr> "https://www.carandbike.com/gravton-motors-bikes/qu…  
## $ acceleration\_speed <chr> "4.2", "3.6", "N/A", "3", "2.9", "2.9", "4", "9.43"…  
## $ top\_speed <dbl> 70, 100, 25, 100, 116, 116, 65, 78, 42, 25, 60, 90,…

#transforming character to numeric variable  
x=as.factor(Motorcycles$type\_of\_bike)  
y<-unclass(x)  
y

## [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 2 2 2  
## [38] 1 1 1 2 1 2 1 1 1 2 1 2 2 2 1 1 1 2 2 2 2 1 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2  
## [75] 2 2 2 2 2 2 1 1 1 2 2 2 2 1 1 1 1 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 1 1 1  
## [112] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [149] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [186] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [223] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [260] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [297] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## [334] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  
## attr(,"levels")  
## [1] "Electric Bike" "Petrol Bike"

class(Motorcycles$type\_of\_bike)

## [1] "character"

# scatterplot  
x=c(Motorcycles$mileage)  
y=c(Motorcycles$weight\_in\_kg)  
  
plot(x,y, main = "scatter plot",  
 xlab = "mileage", ylab= "weight\_in\_kg")  
abline(lm(y ~ x), col = "blue")

