STAT 43000/STAT 53001 Applied Statistics Spring 2023

Lab 3

Suppose we wish to predict the Price of a used sports car, using its Mileage and the type of Car as the predictor variables. The *PorscheJaguar* data set in *Stat2Data* package includes the information for a sample of cars.

a. Read the data in R and determine its dimension

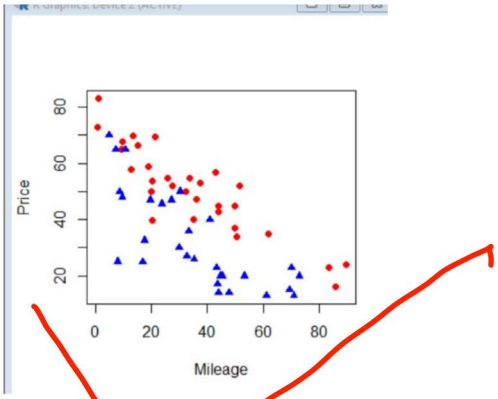
```
> library(Stat2Data)
> data(PorscheJaguar)
> dim(PorscheJaguar)
[1] 60  5
```

b. How many Jaguar cars are included in this datasest?

```
> sum(PorscheJaguar$Porsche=="0")
[1] 30
```

c. Display the Price of the cars using mileage as a predictor. Be sure to choose different color or poin characters for different brand of cars

plot(Price ~ Mileage, data = PorscheJaguar, col = ifelse(Porsche == "1", "red", "blue"), pch = ifelse(Porsche == "1", 16, 17), xlab = "Mileage", ylab = "Price")
>

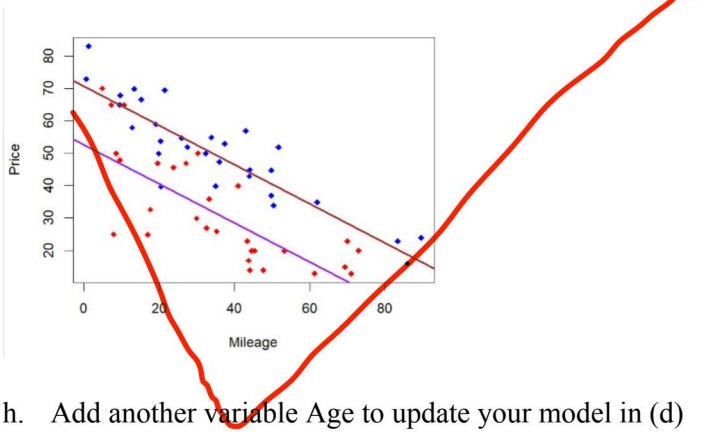


d. Develop a multiple linear regression model and superimposed the models on the scatterplot.

e. `> plot(Mileage,Price,col=ifelse(Porsche=="0","red","blue"),pch=18)`

f. `> abline(52.606,-0.603,col="purple",lwd=2)`

g. `> abline(52.606+17.958,-0.603,col="brown",lwd=2)`



h. Add another variable Age to update your model in (d)

```
> newmodel=lm(Price~Mileage+Porsche+Age)
> newmodel

Call:
lm(formula = Price ~ Mileage + Porsche + Age)

Coefficients:
(Intercept)      Mileage      Porsche          Age
    53.3616      -0.5522      18.4030      -0.3181
```

i. Predict the price of a 10 years old Porsche if it has 33,500 mileage.

j.

```
> predict(newmodel,data.frame(Mileage=33.5,Porsche=1,Age=10))
       1
50.08563
```