DMPM Assignment 2 part 2

Name: Rushikesh Jyoti

Division: A

Roll no: 27

SRN: 201901139

***Question**: Lab Assign-02- Linear Regression Model for Toyota Used-Car Price*

Code

```
library(Metrics)

library(caret)

library(dplyr)

library(scales)


# 1. Read the dataset

cars<-read.csv("ToyotaCorolla.csv")


# 2. Create a model

model<-lm(Price ~ Age + KM + FuelType + HP + MetColor + Automatic + CC +
Doors + Weight,

        data = cars)

print(model)

print(summary(model))
```

```
# 3. Filter out the parameters with less significance
model<-lm(Price ~ Age + KM + FuelType + HP + CC + Weight, data = cars)
print(model)
print(summary(model))
pred1<-predict(model)
resd1<-residuals(model)

predict(model, data.frame(Age=5, KM=2000, FuelType="Diesel", HP=90,
CC=2000, Weight=1200))

# 4. Scatter and Residual Plots
par(mfrow = c(2,1))
plot(cars$Age,resd1,main = "Residual Plot(Age and Price)",abline(0,0,col =
"red"),ylab =
      "Residuals",xlab = "Price in $")
plot(cars$KM,resd1,main = "Residual Plot(KM and Price)",abline(0,0,col =
"red"),ylab =
      "Residuals",xlab = "Price in $")
plot(cars$HP,resd1,main = "Residual Plot(HP and Price)",abline(0,0,col =
"red"),ylab =
      "Residuals",xlab = "Price in $")
plot(cars$CC,resd1,main = "Residual Plot(CC and Price)",abline(0,0,col =
"red"),ylab =
```

```
        "Residuals",xlab = "Price in $")

# 5. Metrics and Evaluation
x<-cbind(cars$Price,pred1)
x<-data.matrix(x)
x<-rescale(x)
x<-as.data.frame(x)
mae<-MAE(x$V1,x$pred1)
mse<-mse(x$V1,x$pred1)
rmse<-RMSE(x$V1,x$pred1)
r2<-R2(x$V1,x$pred1)
cat("\nMean Absolute Error:",mae,"\n\nMean Squared Error:",mse)
cat("\n\nRoot Mean Squared Error:",rmse,"\n\nR-squared:",r2,"\n\n")

# 6. Predictions
x=1:length(pred1)

plot(x, cars$Price,
    pch=19, col = "yellow",main = "Model Evaluation",
    xlab = "Count", ylab = "Price")

lines(x, pred1,col="red")
```

```
legend("topright", legend = c("y-original", "y-predicted"),
       col = c("yellow", "red"),
       pch = c(19,NA), lty = c(NA,1))
```

Output

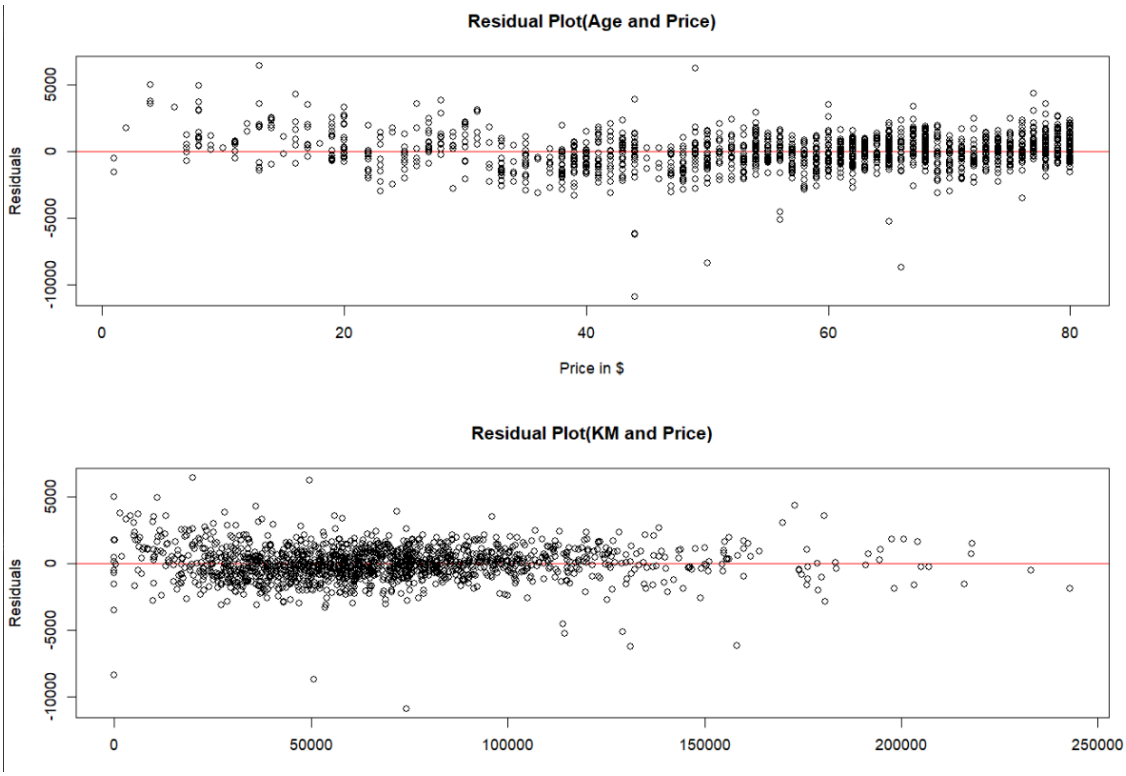
Creating the model

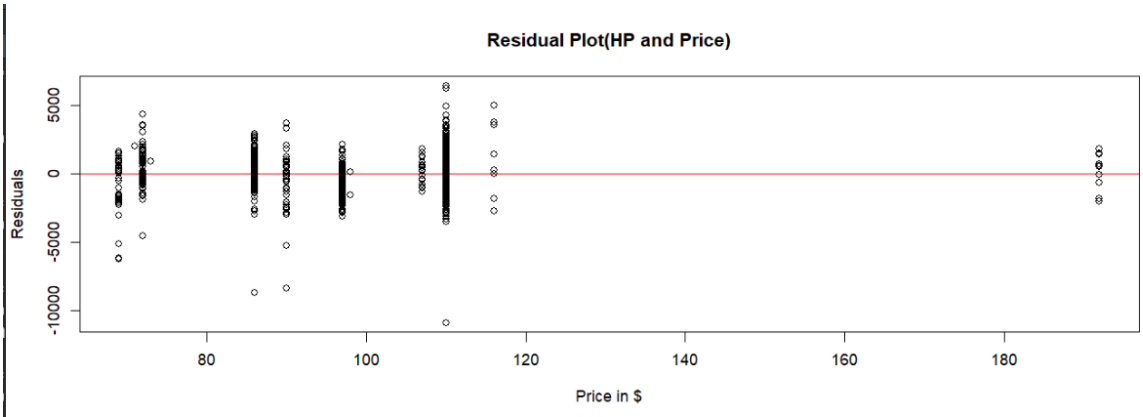
```
> print(model)

Call:
lm(formula = Price ~ Age + KM + FuelType + HP + MetColor + Automatic +
    CC + Doors + Weight, data = cars)

Coefficients:
   (Intercept)             Age              KM  FuelTypeDiesel  FuelTypePetrol
    -3.801e+03      -1.220e+02      -1.621e-02       3.390e+03       1.121e+03
            HP        MetColor       Automatic              CC           Doors
     6.081e+01       5.716e+01       3.303e+02      -4.174e+00      -7.776e+00
        Weight
     2.001e+01
```

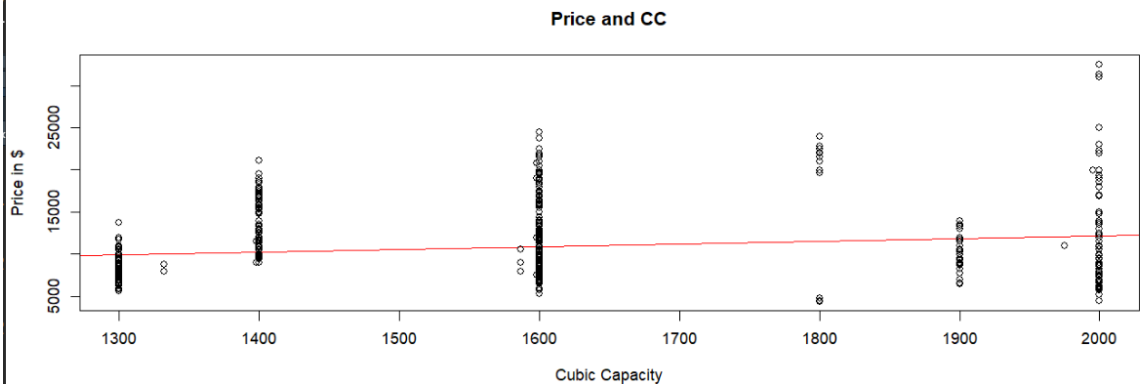
Summary of model

```
> print(summary(model))

Call:
lm(formula = Price ~ Age + KM + FuelType + HP + MetColor + Automatic +
    CC + Doors + Weight, data = cars)

Residuals:
     Min       1Q   Median       3Q      Max
-10642.3   -737.7      3.1    731.3   6451.5

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)    -3.801e+03  1.304e+03  -2.915 0.003613 **
Age            -1.220e+02  2.602e+00 -46.889  < 2e-16 ***
KM             -1.621e-02  1.313e-03 -12.347  < 2e-16 ***
FuelTypeDiesel  3.390e+03  5.188e+02   6.535 8.86e-11 ***
FuelTypePetrol  1.121e+03  3.324e+02   3.372 0.000767 ***
HP              6.081e+01  5.756e+00  10.565  < 2e-16 ***
MetColor        5.716e+01  7.494e+01   0.763 0.445738
Automatic       3.303e+02  1.571e+02   2.102 0.035708 *
CC             -4.174e+00  5.453e-01  -7.656 3.53e-14 ***
Doors          -7.776e+00  4.006e+01  -0.194 0.846129
Weight          2.001e+01  1.203e+00  16.629  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1316 on 1425 degrees of freedom
Multiple R-squared:  0.8693,    Adjusted R-squared:  0.8684
F-statistic:   948 on 10 and 1425 DF,  p-value: < 2.2e-16
```

Plotting the residual values







Metrics

```
> cat("\nMean Absolute Error:",mae,"\n\nMean Squared Error:",mse)

Mean Absolute Error: 0.03382402

Mean Squared Error: 0.00217556
> cat("\n\nRoot Mean Squared Error:",rmse,"\n\nR-squared:",r2,"\n\n")


Root Mean Squared Error: 0.0466429

R-squared: 0.8688574
```

The observed errors are very small, so the accuracy of our model is good

Prediction

```
> predict(model, data.frame(Age=5, KM=2000, FuelType="Diesel", HP=90, CC=2000, Weight
=1200))
       1
20091.74
```

Prediction of a 5 year old car is 20 thousand bucks