

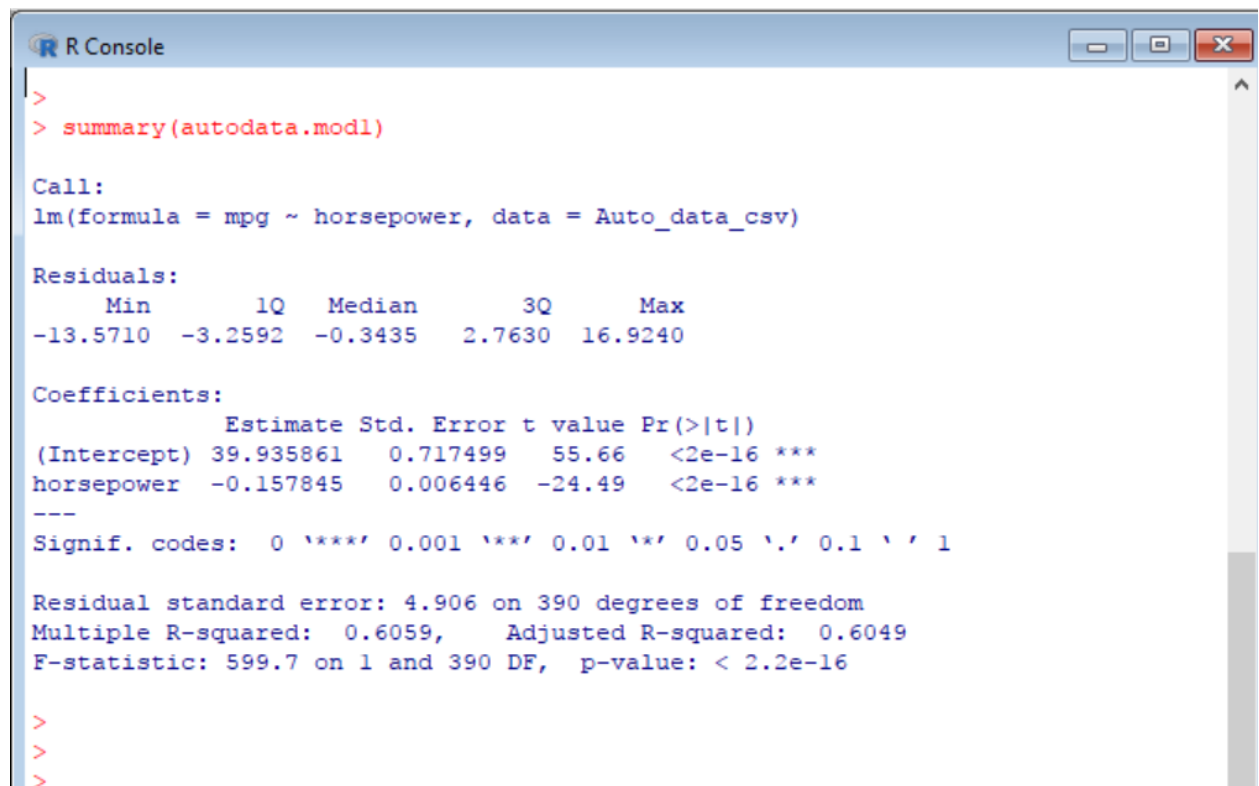
Q-1. For this question, use of simple linear regression on the 'Auto' data set.

- (a) Use the `lm()` function to perform a simple linear regression with **mpg** as the response and **horsepower** as the predictor. Use the `summary()` function to print the results to include in your submission. Comment on the output.

Reading the Auto Data.

```
>
> setwd("C:/Users/putha/Desktop/ISL")
> Auto_data_csv = read.csv("Auto_rev.csv", header = T, na.strings = "?")
> autodata.mod1 = lm(mpg ~ horsepower, data = Auto_data_csv)
```

Summary the Auto Data for mpg and horsepower

The image shows a screenshot of an R Console window. The title bar says "R Console". The console contains the following text:

```
>
> summary(autodata.mod1)

Call:
lm(formula = mpg ~ horsepower, data = Auto_data_csv)

Residuals:
    Min       1Q   Median       3Q      Max
-13.5710  -3.2592  -0.3435   2.7630  16.9240

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 39.935861   0.717499   55.66  <2e-16 ***
horsepower  -0.157845   0.006446  -24.49  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.906 on 390 degrees of freedom
Multiple R-squared:  0.6059,    Adjusted R-squared:  0.6049
F-statistic: 599.7 on 1 and 390 DF,  p-value: < 2.2e-16

>
>
>
```

i. Is there a relationship between the predictor and the response?

Yes. From the summary p-value $< 2.2e-16$ and F-statistics is 599.7 which is larger than 1. So, we can reject the null hypothesis. Thus, there is a clear relationship between the response mpg and predictor horsepower.

ii. How strong is the relationship between the predictor and the response?

From the summary, Multiple R-squared is 0.6059 which indicated to be approximately 61%. Thus, there is **61% variance** between the predictor horsepower with respect to the response mpg.

iii. Is the relationship between the predictor and the response positive or negative?

We can observe from summary that the coefficient of predictor i.e., horsepower is **negative**. The more horsepower an automobile has the linear regression indicates the less mpg fuel efficiency the automobile will have.

iv. What is the predicted **mpg** associated with a **horsepower** of 98? What are the associated 95% confidence and prediction intervals?

```
> predict(autodata.modl, data.frame(horsepower=c(98)), interval="prediction")
      fit      lwr      upr
1 24.46708 14.8094 34.12476
> |
```

```
> predict(autodata.modl, data.frame(horsepower=c(98)), interval="confidence")
      fit      lwr      upr
1 24.46708 23.97308 24.96108
> |
```

(b) Plot the response and the predictor. Use the *abline()* function to display the least squares regression line. Include this graph.

```
> attach(Auto_data_csv)
The following objects are masked from Auto_data_csv (pos = 3):

  acceleration, cylinders, displacement, horsepower, mpg, name,
  origin, weight, year

The following objects are masked from Auto_data_csv (pos = 4):

  acceleration, cylinders, displacement, horsepower, mpg, name,
  origin, weight, year

> plot(horsepower, mpg)
> abline(autodata.modl, lwd=5, col="orange")
```

R Graphics: Device 2 (ACTIVE)

