9. ***This question involves the 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. Comment on the output. For example :

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

- auto <- read.csv("Auto (1).csv", na.strings = "?")
- auto <- na.omit(auto)
- fit <- lm(mpg ~ horsepower, data = auto)
- summary(fit)

```
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
```

The p-value of F-statistic is smaller than 2.2e-16. So we can reject the hypothesis which states the coefficient between "mpg" and "horsepower" are zero. There has a clear evidence of a relationship between "mpg" and "horsepower".

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

We can note that as the R-squared is equal to 0.6059, almost 60.59% of the variability in "mpg" can be explained using "horsepower".

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

The relationship is negative. As shown above, the coefficient of "horsepower" is negative, therefore the relationship is also negative. Which means, if the automobile equipped with more horsepower, then the less mpg fuel efficiency it will have.

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

```
> predict(fit, data.frame(horsepower = 98), interval = "confidence")
       fit      lwr      upr
1 24.46708 23.97308 24.96108
> predict(fit, data.frame(horsepower = 98), interval = "prediction")
       fit     lwr      upr
1 24.46708 14.8094 34.12476
```