Assigment 2

Luis Hinostroza

2023-10-06

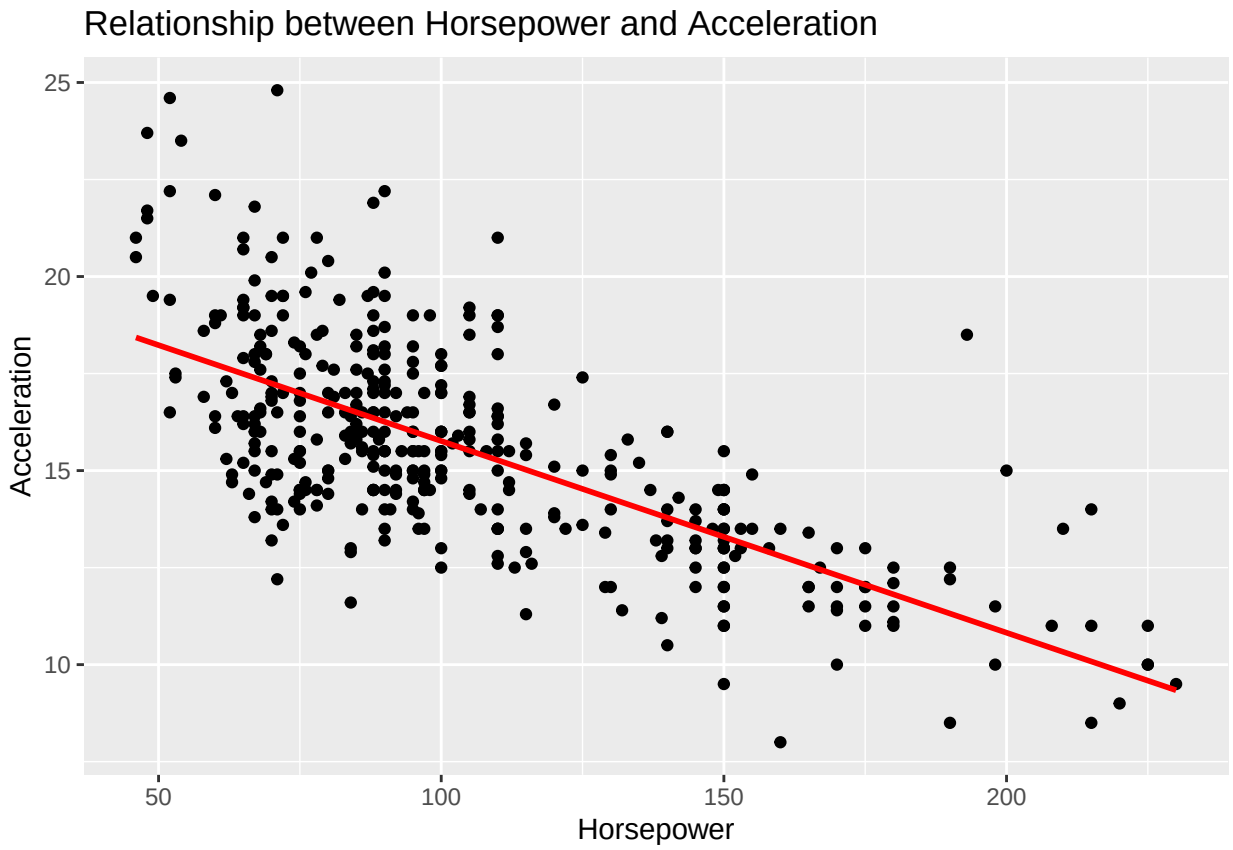
Questions about Cars from a Sample Data.

In this I used ...

Q#01 - Does more horsepower decreases the acceleration time?

Looking at the graph we can notice there is a relation between hp and acceleration. The more hp the less time accelerating. Also the correlation coefficient represent a pretty strong linear relation.

```
library(ggplot2)
library(ISLR)
library(dplyr)
hp_acce <- select(Auto, horsepower, acceleration)
ggplot(hp_acce, aes(x = horsepower, y = acceleration)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  labs(title = "Relationship between Horsepower and Acceleration",
       x = "Horsepower",
       y = "Acceleration")
```



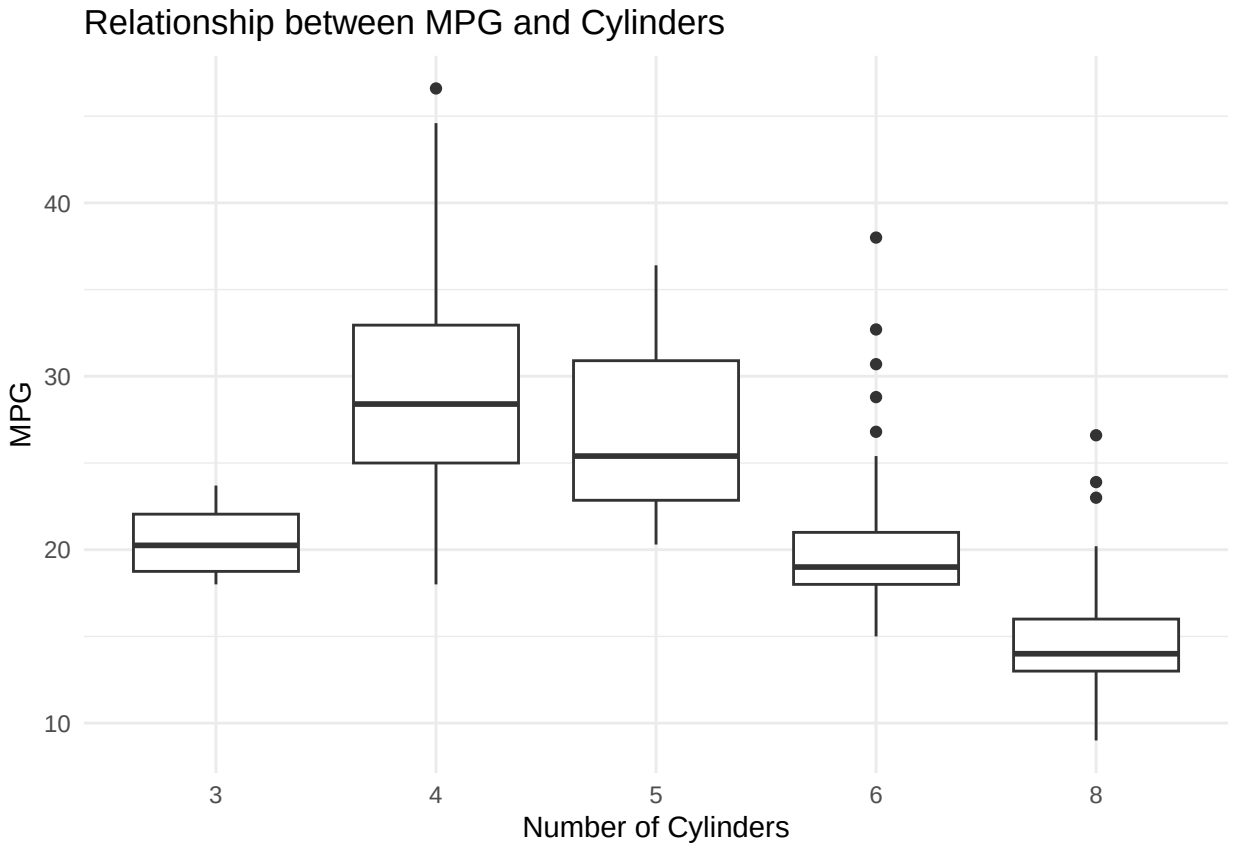

```
cor(Auto$horsepower, Auto$acceleration)
```

```
## [1] -0.6891955
```

Q#02 - Does the mpg depends by the amount of cylinders?

The graph shows that the right amount of cylinder can improve the mpg, Four cylinders been the most efficient miles per gallons. The p-value from the ANOVA test shows there's a statistically significant difference in the mean mpg across different cylinders groups.

```
library(ggplot2)
library(ISLR)
library(dplyr)
mpg_cyl <- select(Auto, mpg, cylinders)
ggplot(mpg_cyl, aes(x = as.factor(cylinders), y = mpg)) +
  geom_boxplot() +
  labs(title = "Relationship between MPG and Cylinders",
       x = "Number of Cylinders",
       y = "MPG") +
  theme_minimal()
```




```
anova_result <- aov(mpg ~ as.factor(cylinders), data = Auto)
summary(anova_result)
```

```
##                      Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(cylinders)   4  15275    3819     173 <2e-16 ***
## Residuals            387   8544      22
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

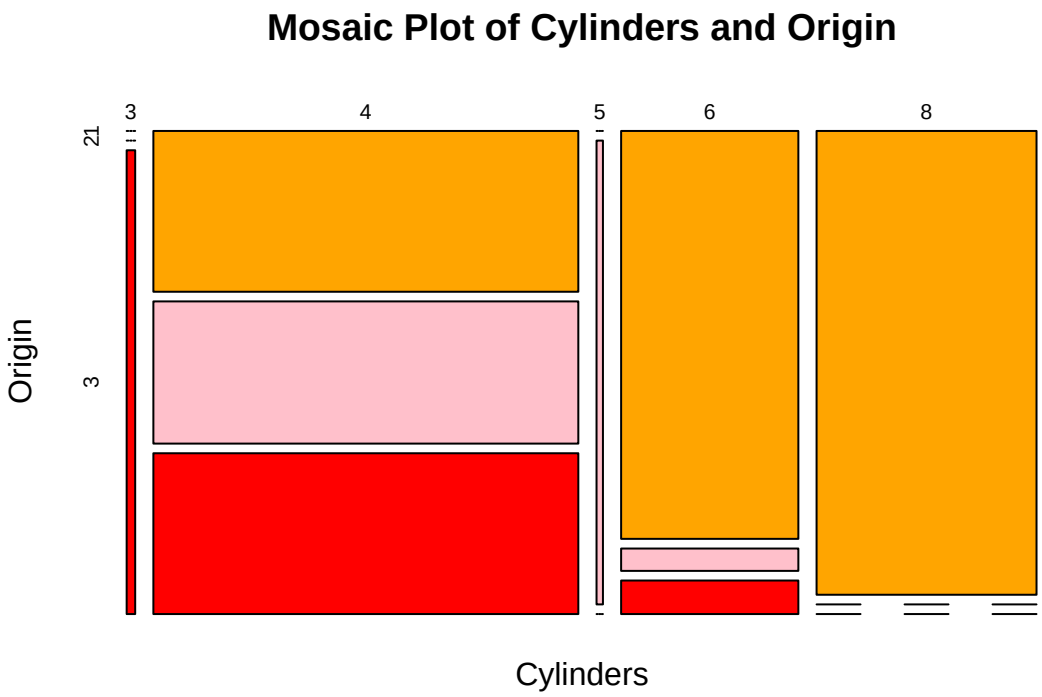
Q#03 - Is the amount of cylinders related to the origin?

We can notice from the graph the distribution of car by number of cylinders. Orange being form USA, pink from EU, and red from Japan. The p-value indicates that the two variables are not independent and that there is a significant association between cylinders and origin.

```
library(ggplot2)
library(ISLR)
library(dplyr)
library(vcd)
table_cyl_origin <- table(Auto$cylinders, Auto$origin)
print(table_cyl_origin)
```

```
##
##       1   2   3
##   3   0   0   4
##   4  69  61  69
##   5   0   3   0
##   6  73   4   6
##   8 103   0   0
```

```
mosaicplot(table_cyl_origin, main="Mosaic Plot of Cylinders and Origin", xlab="Cylinders", ylab="Origin
```




```
chi_sq_test <- chisq.test(table_cyl_origin)
print(chi_sq_test)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cyl_origin
## X-squared = 180.72, df = 8, p-value < 2.2e-16
```