BA ASSIGNMENT1 DIVYA CHANDRASEKARAN 811284790

2023-10-01

```r
# Load ISLR library
library(ISLR)
```

```r
# Load the Carseats dataset
data("Carseats")
```

```r
# Print summary of the Carseats dataset
summary(Carseats)
```

```
##      Sales          CompPrice       Income        Advertising
##  Min.   : 0.000   Min.   : 77   Min.   : 21.00   Min.   : 0.000
##  1st Qu.: 5.390   1st Qu.:115   1st Qu.: 42.75   1st Qu.: 0.000
##  Median : 7.490   Median :125   Median : 69.00   Median : 5.000
##  Mean   : 7.496   Mean   :125   Mean   : 68.66   Mean   : 6.635
##  3rd Qu.: 9.320   3rd Qu.:135   3rd Qu.: 91.00   3rd Qu.:12.000
##  Max.   :16.270   Max.   :175   Max.   :120.00   Max.   :29.000
##    Population        Price        ShelveLoc        Age          Education
##  Min.   : 10.0   Min.   : 24.0   Bad   : 96   Min.   :25.00   Min.   :10.0
##  1st Qu.:139.0   1st Qu.:100.0   Good  : 85   1st Qu.:39.75   1st Qu.:12.0
##  Median :272.0   Median :117.0   Medium:219   Median :54.50   Median :14.0
##  Mean   :264.8   Mean   :115.8                Mean   :53.32   Mean   :13.9
##  3rd Qu.:398.5   3rd Qu.:131.0                3rd Qu.:66.00   3rd Qu.:16.0
##  Max.   :509.0   Max.   :191.0                Max.   :80.00   Max.   :18.0
##  Urban        US
##  No :118   No :142
##  Yes:282   Yes:258
##
##
##
##
```

```r
#FROM THE ABOVE OUTPUT WE CAN SEE THAT THERE ARE 11 ROWS AND 400 OBSERVATIONS
##THE NAMES OF THESE 11 ROWS ARE: SALES, COMPRICE, INCOME, ADVERTISING, POPULATION, PRICE,
#SHELVELOC, AGE, EDUCATION, URBAN, US.
```

```r
# Find the maximum value of the "advertising" attribute
max_advertising <- max(Carseats$Advertising)
max_advertising
```

```
## [1] 29
```

```
# Calculate the IQR of the Price attribute
iqr_price <- IQR(Carseats$Price)
iqr_price
```

```
## [1] 31
```

```
#Plot Sales against Price
library(ggplot2) #FOR CREATING DATA VISUALIZATION
boxplot(Carseats$Sales ~ Carseats$Price, xlab = "Price", ylab = "Sales", main = "BOX PLOT OF SALES AGAI
```

BOX PLOT OF SALES AGAINST PRICE
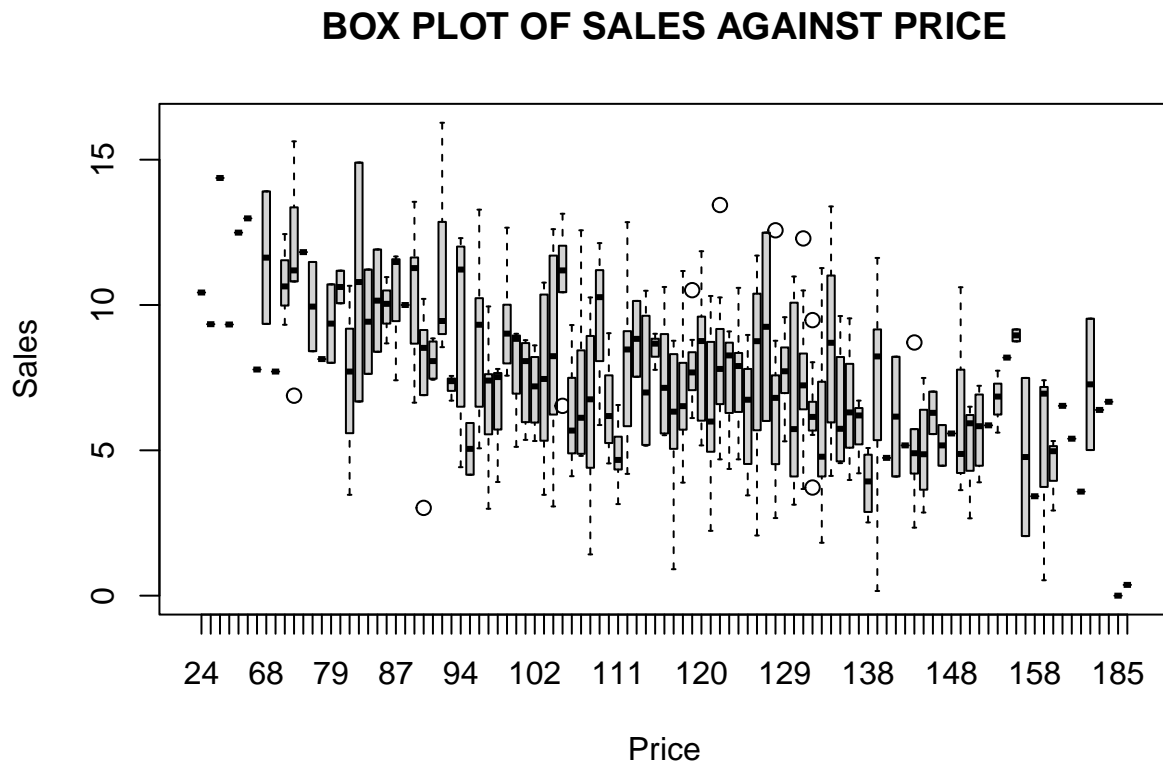

```
# Calculate the correlation coefficient
correlation <- cor(Carseats$Price, Carseats$Sales)
correlation
```

```
## [1] -0.4449507
```

```
#The correlation analysis will always result between -1 and 1. When the result is -1,
##the x value will increase and the y value will decrease; therefore, showing a negative significance.
##So, in the context of your Sales vs. Price plot and the negative correlation coefficient,
##it suggests that there is a negative linear relationship between the price of car seats,
##and the sales of car seats in the data set. As the price increases, sales tend to decrease,
##and as the price decreases, sales tend to increase.
```