Assignment-1 (BA)

Dev

2023-09-19

Q1: Install the ISLR library using the install.packages() command. Call the library using the library(ISLR) command to ensure that the library is correctly installed.

```
#Installing ISLR package and using it
library(ISLR)
```

Q2: Create a new R-Notebook (.Rmd) file. In the first code chunk, call the ISLR library and then print the summary of the Carseats dataset. How many observations (rows) this dataset contains?

```
#Displaying summary of carseats data:
library(ISLR)
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
```

```
#Counting number of rows:
nrow(Carseats)
```

```
[1] 400
```

Q3:Using the summary statistics shown above, what is maximum value of the advertising attribute?

As per summary above, maximum in Advertising attribute is 29.00. We can verify this observation from following code:

```
max(Carseats$Advertising)
```

```
[1] 29
```

Hence, our observation is correct that maximum of Advertising attribute is 29.

Q4:Calculate the IQR of the Price attribute.

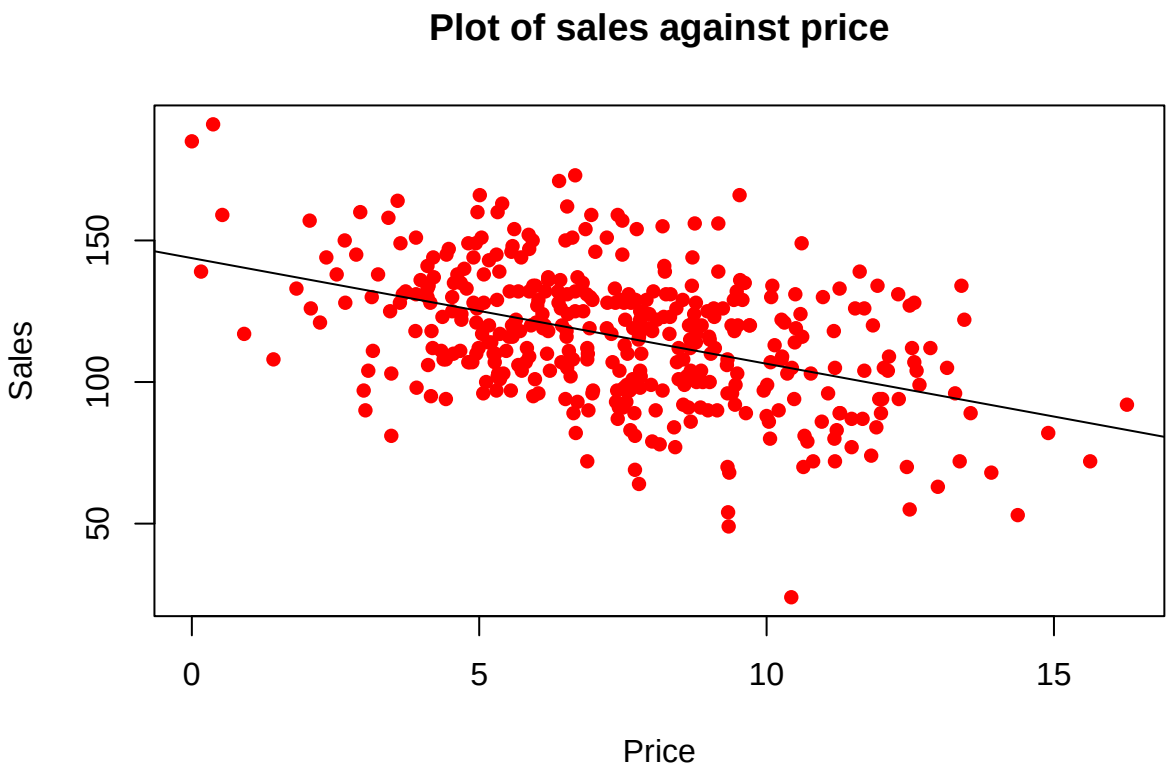
IQR of Price attribute is :

```
#Using IQR functin to calculate IQR of price attribute
IQR(Carseats$Price)
```

```
[1] 31
```

Q5: Plot the Sales against Price. What do you see in there? Calculate the correlation of the two attributes. What does the sign of the correlation coefficient suggest?

```
#Plotting sales against price
plot(Carseats$Sales,
     Carseats$Price,
     main="Plot of sales against price",
     xlab="Price",
     ylab="Sales",
     pch=16,
     col="red")
abline(lm(Carseats$Price~Carseats$Sales,data=Carseats),col="black")
```



We see that the graph is most crowded in the price range of 5 to 10 and the sales are maximum in this range as well. Also, most of the points are not close to the regression line which shows us that the relation between attributes is not strong. Besides this, we have a declining regression line which means that the attributes have a inverse linear relation, which can be verified by negative value of Pearson's correlation coefficient.

Calculating co-relation between sales and price:

```
#Calculating relation between sales and price
print(cor(Carseats$Sales,Carseats$Price))
```

```
[1] -0.4449507
```

The value of correlation coefficient hence turns out to be -0.44. The negative value of coefficient confirms our observation that the relation between attributes is inverse and it's small magnitude confirms that the inverse linear relation between them is not strong.