

# YESWANTH SIRIPURAPU R Notebook

```
#Install the ISLR PACKAGE
```

```
library("ISLR")
```

With this line of code, a summary of the “Carseats” dataset is produced.

#The summary() method offers a number of statistics and details about the #variables in the dataset. For numerical variables, it will show measurements #like mean, median, minimum, maximum, and quartiles; for categorical variables, #it will show counts. Understanding the properties of the dataset in general #is made easier by reading this summary. The number of rows in the “Carseats” #dataset are determined by the statement nrow(Carseats). The entire number of #observations or data points in the dataset is returned, in other words.

```
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      ShelfLoc      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
##
##
##
##
```

```
nrow(Carseats)
```

```
## [1] 400
```

The greatest value in the “Advertising” column of the “Carseats” dataset will

#be returned by this line. It will specifically provide you the observation in #the dataset with the highest average advertising expenditure.

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

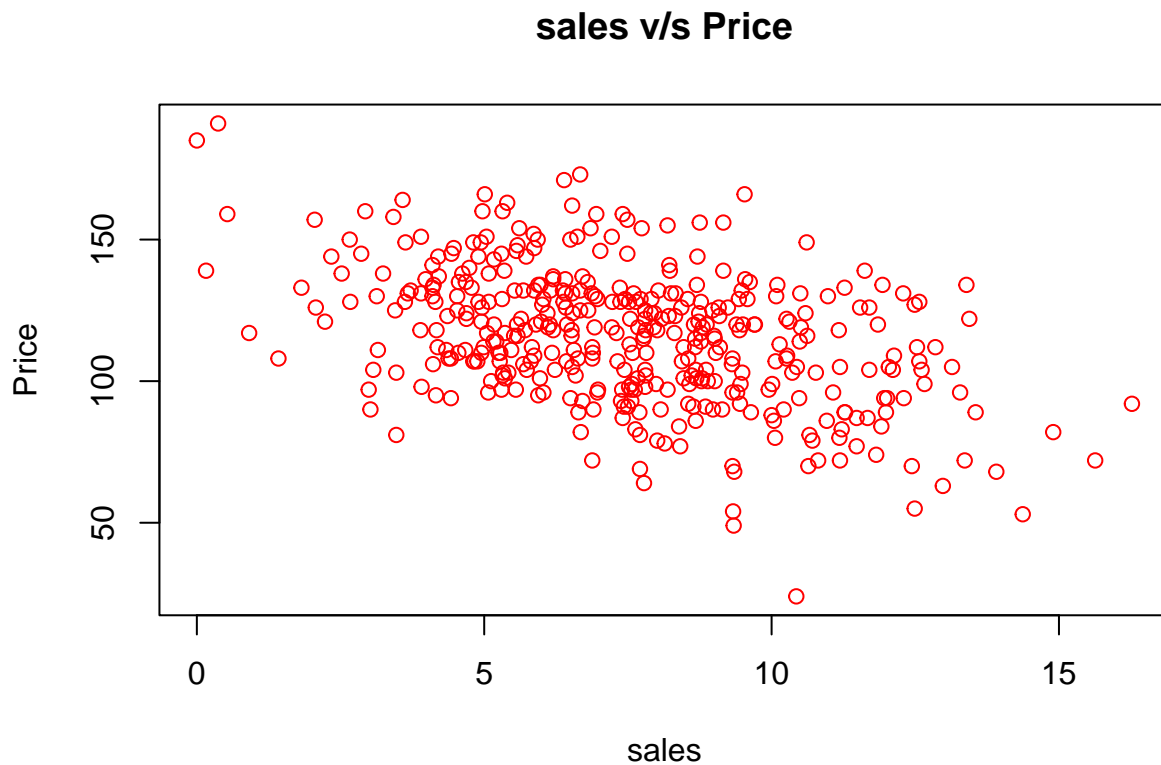
```
## [1] 29
```

```
IQR(Carseats $ Price)
```

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

#The code plots the sales of carseats over time, and the price. #The main is “sales v/s Price”. #The x-axis shows the number of cars seats sold in a given year. #The y-axis shows how much each car seat cost at that time. #The code creates a scatter plot of the data from Carseats. #The code also includes a title, main, and x-axis label.

```
plot(Carseats$`Sales`,Carseats$`Price`,  
     main = "sales v/s Price",  
     xlab = "sales",  
     ylab = "Price",  
     col= "red")
```



#The code starts by creating a list of cars that are available for sale. #The code then creates a new

variable called “Carseats” which is the name of the #data set and contains all the information about each car. #Next, it creates another new variable called “Sales” which is an array #containing all the sales prices for each car. # Finally, it calculates the correlation between Sales and Price using cor() #function in R programming language. # The first if statement checks to see if there is a positive correlation between #Sales and Price (correlation > 0). # If there was no correlation greater than zero, then this would mean that there #No linear relationship between these two variables (no linear relationship) #The code will print “The correlation is close to zero (no linear relationship). #If the correlation between CarseatsSalesandCarseatsPrice is less than 0.5.

```
correlation <- cor(Carseats$Sales, Carseats$Price)
correlation
```

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

```
if (correlation > 0) {
  cat("The correlation is positive.\n")
} else if (correlation < 0) {
  cat("The correlation is negative.\n")
} else {
  cat("The correlation is close to zero (no linear relationship).\n")
}
```

```
## The correlation is negative.
```