Student Number:- IT24103418

Name:- Fernando Pulle I.S.

Labsheet -04

1. Import the dataset ('Exercise.txt') into R and store it in a data frame called "branch data".

```
#Question 01
setwd("C:\\Users\\IT24103418\\Desktop\\IT24103418")
branch_data<-read.table("Exercise.txt",header =TRUE, sep=",")
head(branch_data)
```

2. Identify the variable type and scale of measurement for each variable.

```
6
7  #Question_02
8  str(branch_data)
9
```

3. Obtain boxplot for sales and interpret the shape of the sales distribution.

```
11 #Question_03
12 boxplot(branch_data$Sales,
13         main = "Boxplot of Sales",
14         ylab = "Sales",
15         col = "red")
16
```

4. Calculate the five number summary and IQR for advertising variable.

```
#Question_04
fivenum(branch_data$Advertising)

summary(branch_data$Advertising)


IQR(branch_data$Advertising)
```

5. Write an R function to find the outliers in a numeric vector and check for outliers in years variables.

```r
#Question_05
find_outliers <- function(x) {
  Q1 <- quantile(x, 0.25)
  Q3 <- quantile(x, 0.75)
  IQR <- Q3 - Q1
  lower <- Q1 - 1.5 * IQR
  upper <- Q3 + 1.5 * IQR
  outliers <- x[x < lower | x > upper]
  return(outliers)
}
find_outliers(branch_data$Years)
```