Exercise

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

```
setwd("C:\\Users\\IT24103570\\Desktop\\IT24103570")

#Q1
branch_data <- read.csv("Exercise.txt", header = TRUE)
```

```
> setwd("C:\\Users\\IT24103570\\Desktop\\IT24103570")
> branch_data <- read.csv("Exercise.txt", header = TRUE)
```

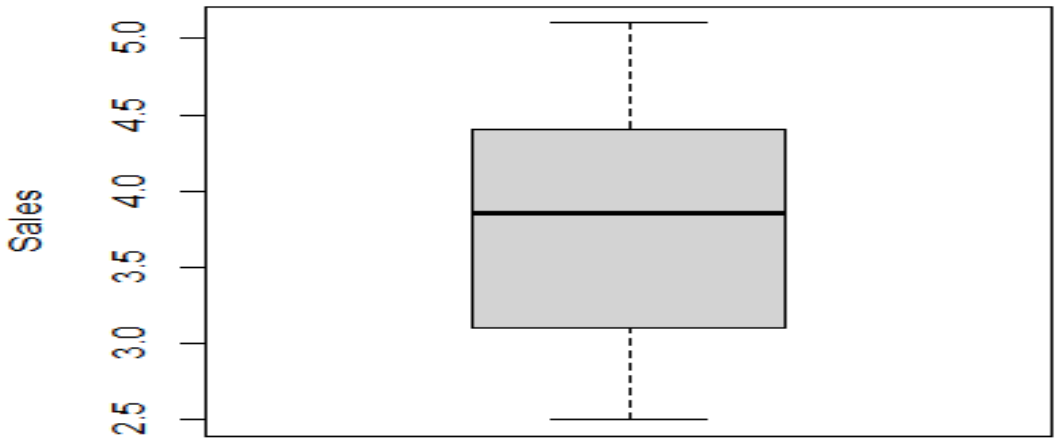
2. Identify the variable type and scale of measurement for each variable.

```
#Q2
str(branch_data)
names(branch_data)
```

```
> str(branch_data)
'data.frame':   30 obs. of  4 variables:
 $ Branch        : int  1 2 3 4 5 6 7 8 9 10 ...
 $ Sales_X1      : num  3.4 4.1 2.8 5 3.7 4.5 3 4.9 3.2 2.5 ...
 $ Advertising_X2: int  120 150 90 200 110 175 95 185 105 80 ...
 $ Years_X3      : int  4 7 3 10 5 6 2 9 4 1 ...
> names(branch_data)
[1] "Branch"         "Sales_X1"       "Advertising_X2" "Years_X3"
```

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

```
#Q3
boxplot(branch_data$Sales,main="Boxplot of sales",YLAB="Sales")
```



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

```
#Q4
summary(branch_data$Advertising_X2)
iqr_advertising <- IQR(branch_data$Advertising_X2)
print(paste("IQR of advertising:", iqr_advertising))
```

```
> boxplot(branch_data$Sales_X1,MAIN="Boxplot of sales",ylab="Sales")
> summary(branch_data$Advertising_X2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   80.0   101.2   132.5   134.8   158.8   210.0
> iqr_advertising <- IQR(branch_data$Advertising_X2)
> print(paste("IQR of advertising:", iqr_advertising))
[1] "IQR of advertising: 57.5"
```

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

```
#Q5
find_outliers <- function(x) {
  Q1 <- quantile(x, 0.25,na.rm =TRUE)
  Q3 <- quantile(x,o.75,na.rm=TRUE)
  IQR_val <- Q3 - Q1
  lower_bound <- Q1-1.5*IQR_val
  upper_bound <- Q3+1.5*IQR_val
  outliers <- x[x < lower_bound | x > upper_bound]
  return(outliers)
}
outliers_years <- find_outliers(branch_data$Years_X3)
print("outliers in 'years' variable:")
print(outliers_years)
```

```
> find_outliers<-function(x){
+   Q1<-quantile(x,0.25,na.rm=TRUE)
+   Q3<-quantile(x,0.75,na.rm=TRUE)
+   IQR_val<-Q3-Q1
+   lower_bound<-Q1-1.5*IQR_val
+   upper_bound<-Q3+1.5*IQR_val
+   outliers<-x[x<lower_bound | x>upper_bound]
+   return(outliers)
+ }
>
> outliers_years<-find_outliers(branch_data$Years_X3)
> print("Outliers in 'years' variable:")
[1] "Outliers in 'years' variable:"
> print(outliers_years)
integer(0)
```