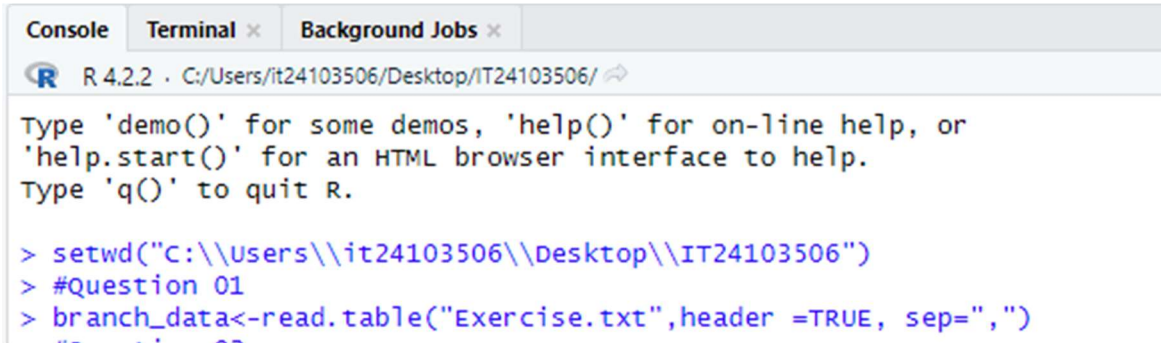


IT2120 - Probability and Statistics

01.

```
1 setwd("C:\\Users\\it24103506\\Desktop\\IT24103506")
2
3 #Question 01
4 branch_data<-read.table("Exercise.txt",header =TRUE, sep=",")
5 head(branch_data)
```



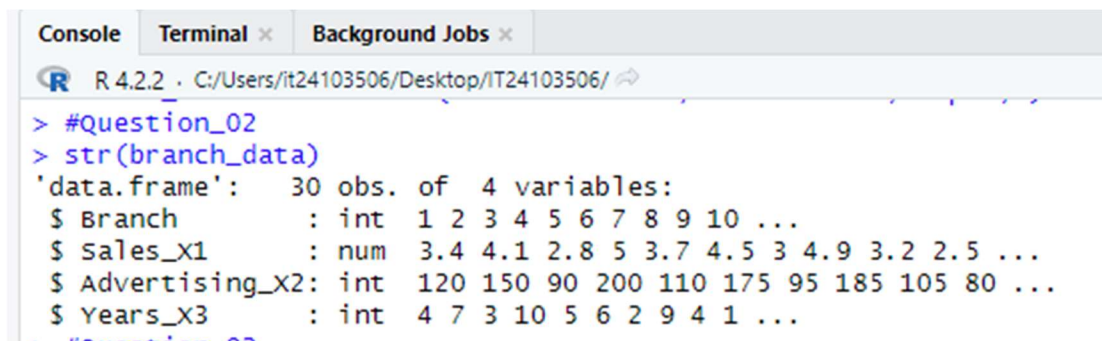
The screenshot shows the R console interface with tabs for Console, Terminal, and Background Jobs. The console displays the R version (4.2.2) and the current working directory (C:/Users/it24103506/Desktop/IT24103506/). It shows the execution of the first script, including the directory change, the comment '#Question 01', and the loading of 'Exercise.txt' into 'branch_data'. The output shows the first few rows of the data frame.

```
R 4.2.2 · C:/Users/it24103506/Desktop/IT24103506/
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("C:\\Users\\it24103506\\Desktop\\IT24103506")
> #Question 01
> branch_data<-read.table("Exercise.txt",header =TRUE, sep=",")
#>      Branch Sales_X1 Advertising_X2 Years_X3
```

02.

```
#Question 02
str(branch_data)
```



The screenshot shows the R console interface with tabs for Console, Terminal, and Background Jobs. The console displays the R version (4.2.2) and the current working directory (C:/Users/it24103506/Desktop/IT24103506/). It shows the execution of the second script, which runs '#Question_02' and 'str(branch_data)'. The output shows the structure of the 'branch_data' data frame, including the number of observations (30) and the data types and values for the four variables: Branch (integer), Sales_X1 (numeric), Advertising_X2 (integer), and Years_X3 (integer).

```
R 4.2.2 · C:/Users/it24103506/Desktop/IT24103506/
> #Question_02
> 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 ...
```

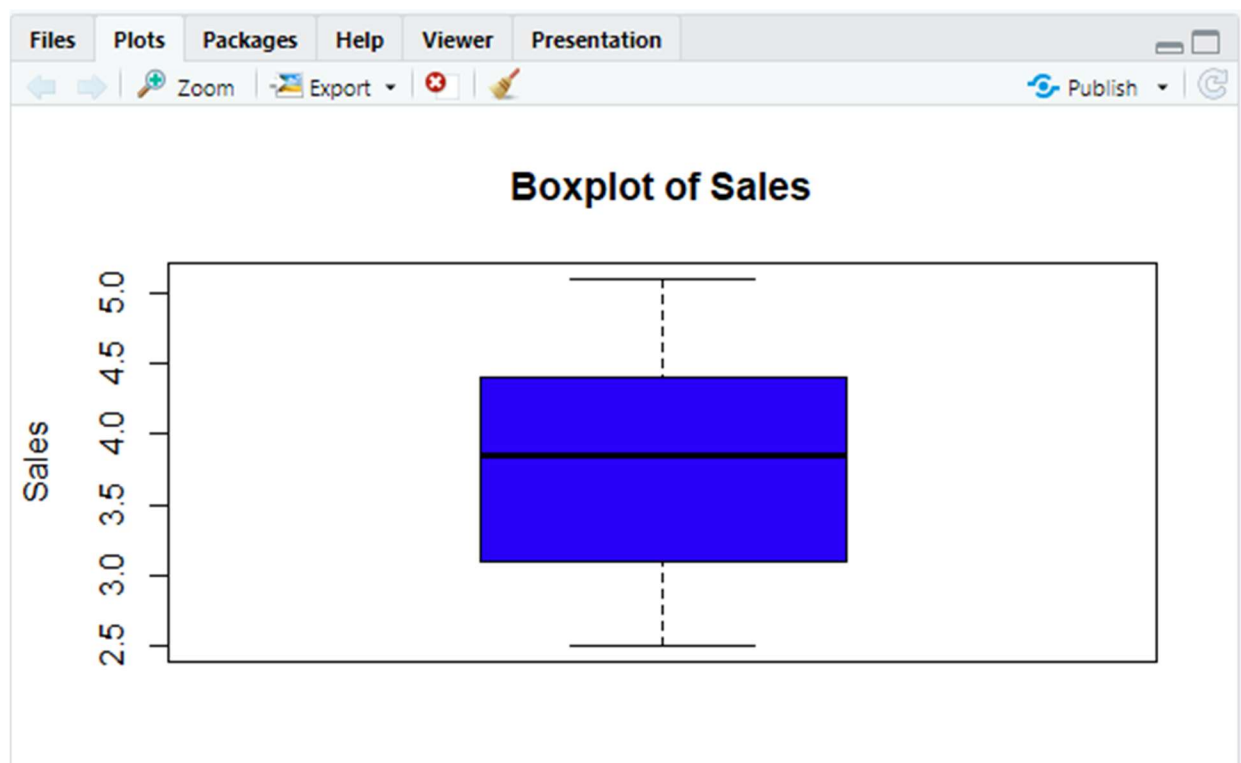
03.

```
10 #Question 03
11 boxplot(branch_data$Sales,
12         main = "Boxplot of sales",
13         ylab = "Sales",
14         col = "blue")
15
```

Console Terminal x Background Jobs x

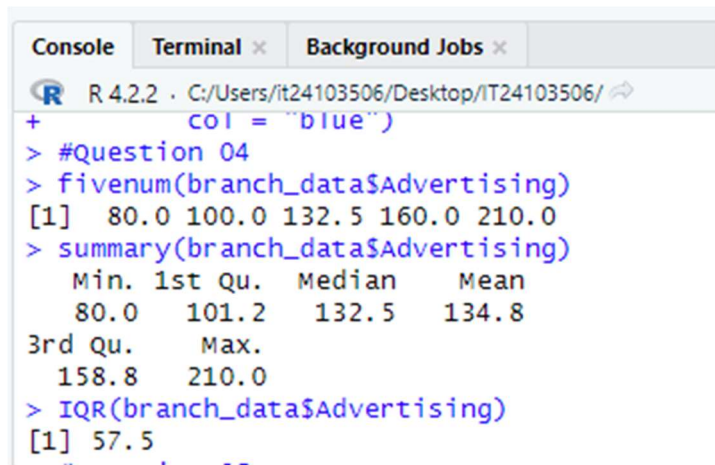
R 4.2.2 · C:/Users/it24103506/Desktop/IT24103506/

```
> #Question 03
> boxplot(branch_data$Sales,
+         main = "Boxplot of sales",
+         ylab = "Sales",
+         col = "blue")
```



04.

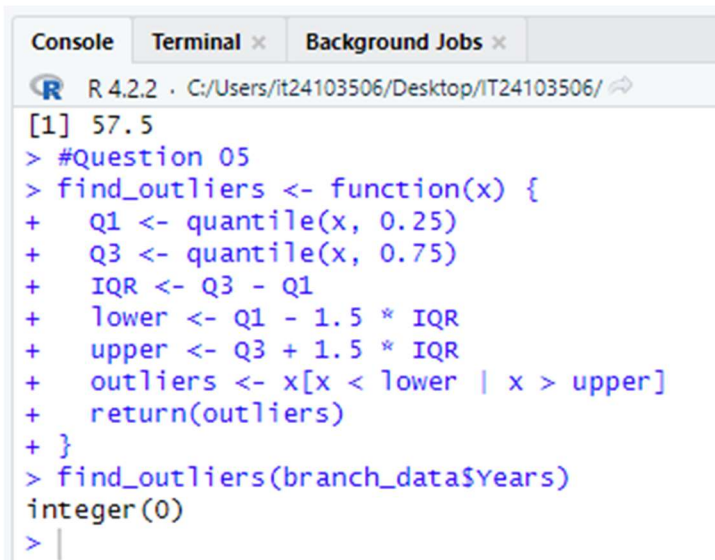
```
16 #Question 04
17 fivenum(branch_data$Advertising)
18
19 summary(branch_data$Advertising)
20 IQR(branch_data$Advertising)
21
```



```
R 4.2.2 - C:/Users/it24103506/Desktop/IT24103506/
+ col = "blue")
> #Question 04
> fivenum(branch_data$Advertising)
[1] 80.0 100.0 132.5 160.0 210.0
> summary(branch_data$Advertising)
   Min. 1st Qu.  Median    Mean
   80.0  101.2   132.5   134.8
3rd Qu.    Max.
 158.8   210.0
> IQR(branch_data$Advertising)
[1] 57.5
```

05.

```
23 #Question 05
24 find_outliers <- function(x) {
25   Q1 <- quantile(x, 0.25)
26   Q3 <- quantile(x, 0.75)
27   IQR <- Q3 - Q1
28   lower <- Q1 - 1.5 * IQR
29   upper <- Q3 + 1.5 * IQR
30   outliers <- x[x < lower | x > upper]
31   return(outliers)
32 }
33 find_outliers(branch_data$Years)
34
```



```
R 4.2.2 - C:/Users/it24103506/Desktop/IT24103506/
[1] 57.5
> #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)
integer(0)
>
```