Sri Lanka Institute of Information Technology

Lab Submission
Lab Sheet 05

IT24101219
Piyarathna S.G.D.V
Probability and Statistics| IT2120

B.Sc.(Hons) in Information Technology

Question

Exercise

Instructions: Create a folder in your desktop with your registration number (Eg: "IT......"). You need to save the R script file and take screenshots of the command prompt with answers and save it in a word document inside the folder. Save both R script file and word document with your registration number (Eg: "IT....."). After you finish the exercise, zip the folder and upload the zip file to the submission link.

1. Import the dataset ('Exercise – Lab 05.txt') into R and store it in a data frame called "Delivery_Times".

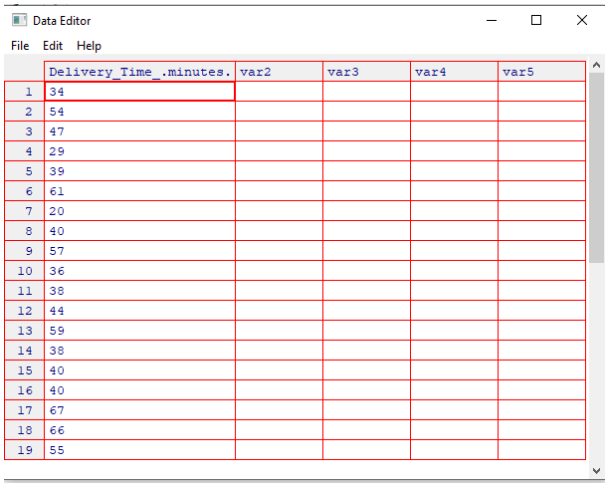
Answer (Code)

```
setwd("C:\\Users\\it24101219\\Desktop\\IT24101219")

Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE, sep = "")

fix(Delivery_Times)

names(Delivery_Times) = "Time_in_mins"

attach(Delivery_Times)

summary(Delivery_Times)
```

Output (Terminal)

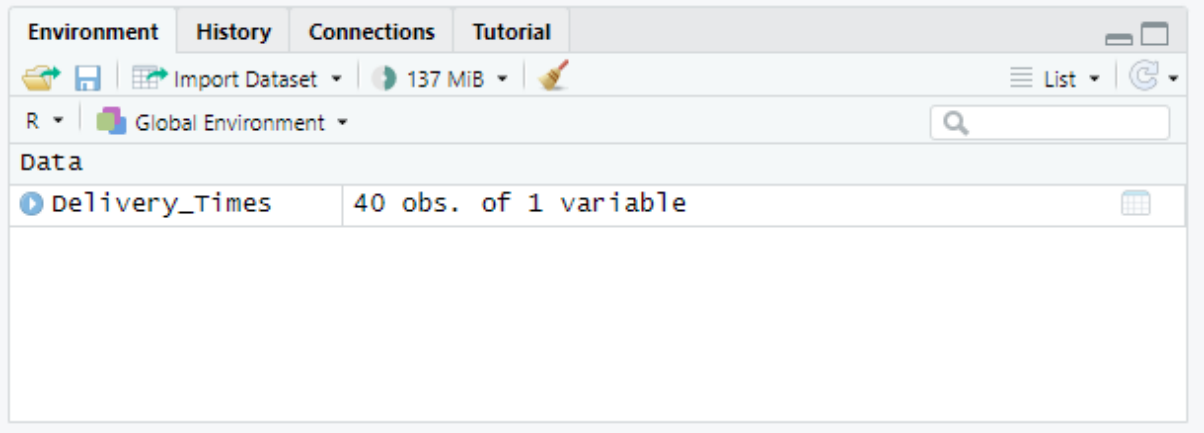
```
Console   Terminal ×   Background Jobs ×
R 4.2.2 · C:/Users/it24101219/Desktop/IT24101219/
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

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\\it24101219\\Desktop\\IT24101219")
>
> Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE, sep = "")
>
> fix(Delivery_Times)
>
> names(Delivery_Times) = "Time_in_mins"
>
> attach(Delivery_Times)
>
> summary(Delivery_Times)
  Time_in_mins
 Min.   :20.00
 1st Qu.:36.00
 Median :42.50
 Mean   :43.75
 3rd Qu.:54.00
 Max.   :67.00
>
```

Output (Fix View)



	Delivery_Time_.minutes.	var2	var3	var4	var5
1	34				
2	54				
3	47				
4	29				
5	39				
6	61				
7	20				
8	40				
9	57				
10	36				
11	38				
12	44				
13	59				
14	38				
15	40				
16	40				
17	67				
18	66				
19	55				

Output(Variable Editor)



Question

2. Draw a histogram for deliver times using nine class intervals where the lower limit is 20 and upper limit is 70. Use right open intervals.

Answer (Code)

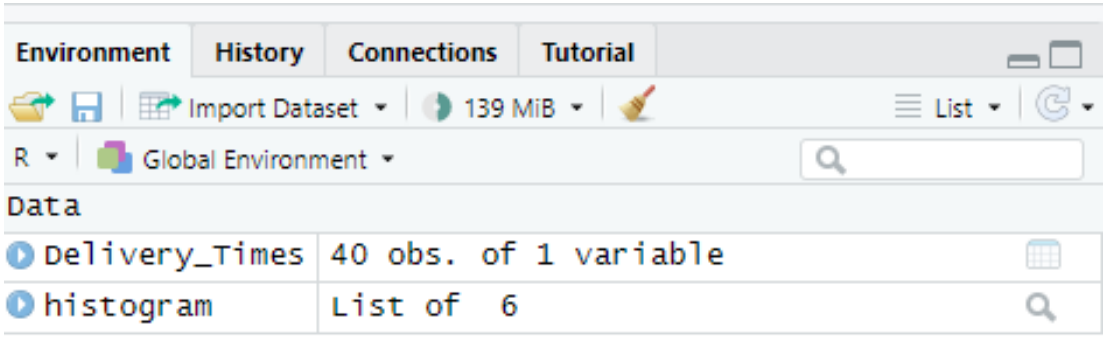
```
histogram <- hist(Time_in_mins, main = "Histogram for Delivery Minutes", breaks = seq(20,70, length = 9), right = FALSE)
#Most of the delivery times are in between 30 - 60 minutes
```

Output (Terminal)

```
> histogram <- hist(Time_in_mins, main = "Histogram for Delivery Minutes", breaks = seq(20,70, length = 9), right = FALSE)
> #Most of the delivery times are in between 30 - 60 minutes
>
```

Output (Plot)

Output (Variable Editor)



3. Comment on the shape of the distribution.

Most of the delivery times are in between 30 - 60 minutes

Question

4. Draw a cumulative frequency polygon (ogive) for the data in a separate plot.

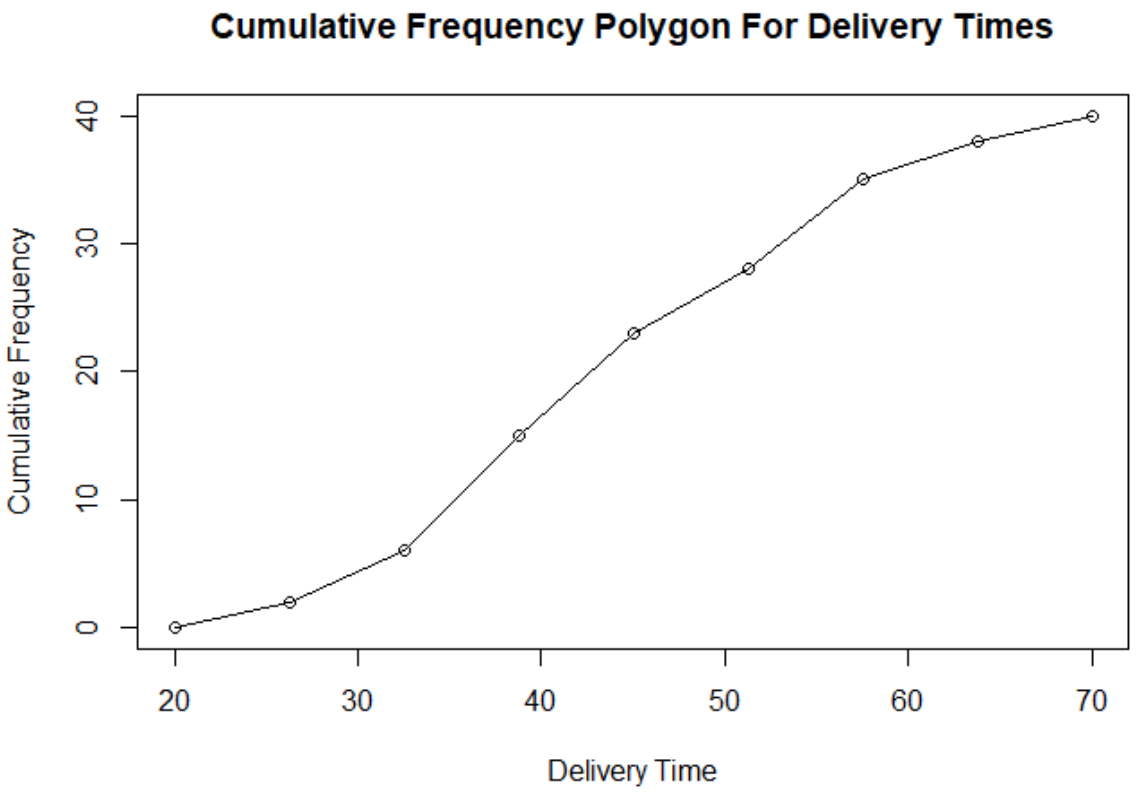
Answer (Code)

```
16
17  freq <- histogram$counts
18  breaks <- histogram$breaks
19
20  cum.freq <- cumsum(freq)
21
22  new <- c()
23
24  for (i in 1:length(breaks)) {
25    if(i == 1){
26      new[i] = 0
27    } else{
28      new[i] = cum.freq[i-1]
29    }
30  }
31
32  new
33
34  plot(breaks, new, type = "o", main = "Cumulative Frequency Polygon For Delivery Times",
35       ylab = "Cumulative Frequency", xlab = "Delivery Time",
36       ylim=c(0, max(cum.freq)))
```

Output (Terminal)

```
> freq <- histogram$counts
> breaks <- histogram$breaks
>
> cum.freq <- cumsum(freq)
>
> new <- c()
>
> for (i in 1:length(breaks)) {
+   if(i == 1){
+     new[i] = 0
+   } else{
+     new[i] = cum.freq[i-1]
+   }
+ }
>
> new
[1]  0  2  6 15 23 28 35 38 40
>
> plot(breaks, new, type = "o", main = "Cumulative Frequency Polygon For Delivery Times",
+      ylab = "Cumulative Frequency", xlab = "Delivery Time",
+      ylim=c(0, max(cum.freq)))
>
```

Output (Plot)



Output (Variable Editor)