

Sri Lanka Institute of Information Technology



Lab Submission
Lab sheet No 05

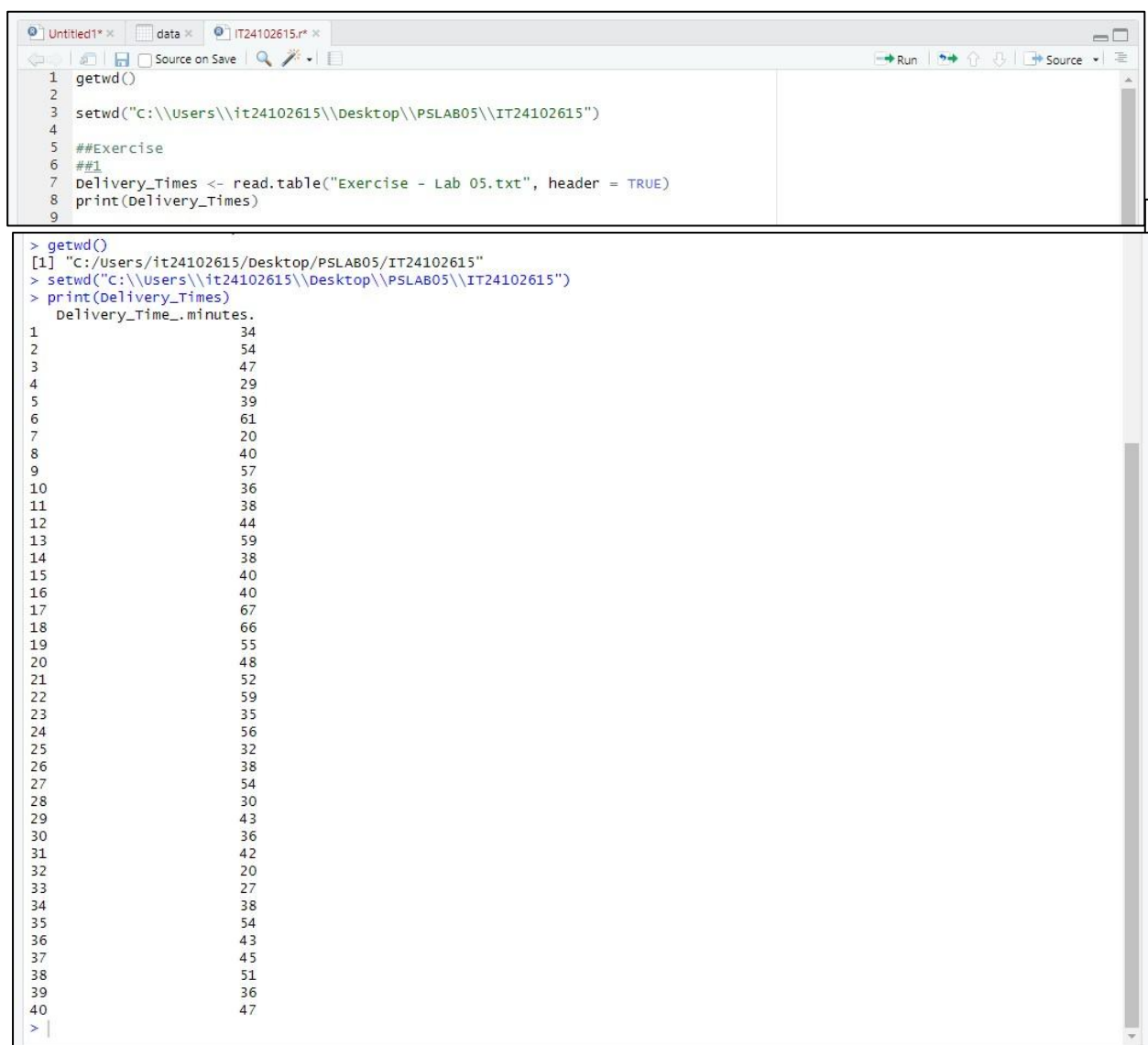
IT24102615
Mamalgaha I.G.W.S

Probability & Statistics | IT2120
B.Sc. (Hons) in Information Technology

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".

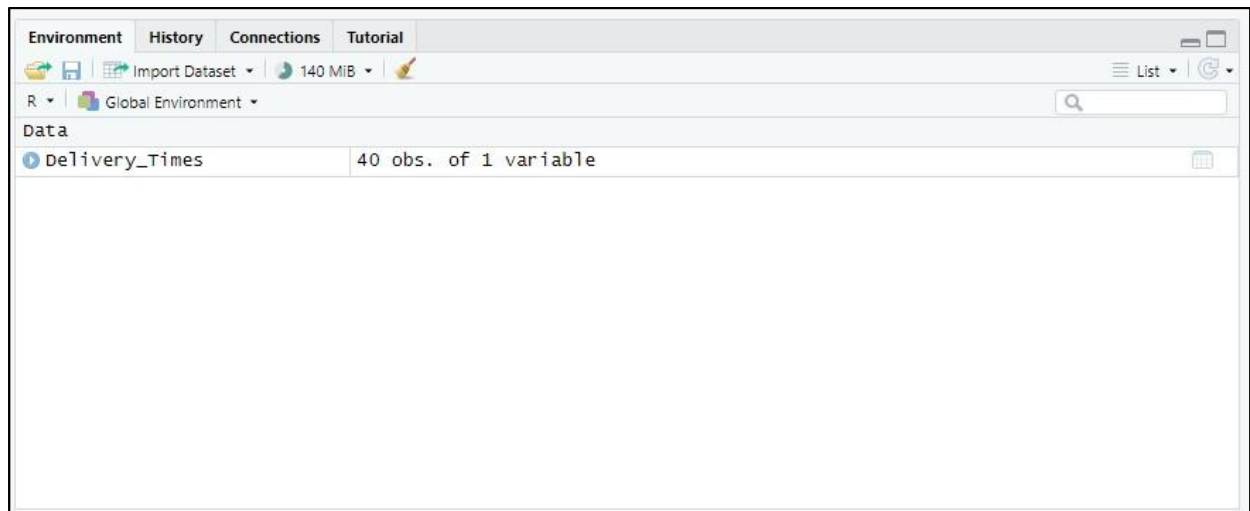


The screenshot shows an RStudio window with two panes. The top pane contains an R script with the following code:

```
1 getwd()
2
3 setwd("C:\\Users\\it24102615\\Desktop\\PSLAB05\\IT24102615")
4
5 ##Exercise
6 ##1
7 Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)
8 print(Delivery_Times)
9
```

The bottom pane shows the output of the script:

```
> getwd()
[1] "C:/Users/it24102615/Desktop/PSLAB05/IT24102615"
> setwd("C:\\Users\\it24102615\\Desktop\\PSLAB05\\IT24102615")
> print(Delivery_Times)
  Delivery_Time_.minutes.
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
20                     48
21                     52
22                     59
23                     35
24                     56
25                     32
26                     38
27                     54
28                     30
29                     43
30                     36
31                     42
32                     20
33                     27
34                     38
35                     54
36                     43
37                     45
38                     51
39                     36
40                     47
>
```



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.

```
10 ##2
11 hist(Delivery_Times$Delivery,
12       breaks = seq(20, 70, by = 5),
13       right = FALSE,
14       main = "Histogram of Delivery Times",
15       xlab = "Delivery Times Minutes",
16       ylab = "Frequency",
17       col = "lightblue",
18       border = "black")
19
```

```
> ##2
> hist(Delivery_Times$Delivery,
+       breaks = seq(20, 70, by = 5),
+       right = FALSE,
+       main = "Histogram of Delivery Times",
+       xlab = "Delivery Times Minutes",
+       ylab = "Frequency",
+       col = "lightblue",
+       border = "black")
```



3. Comment on the shape of the distribution.

```
> ##The distribution of delivery times is approximately symmetric and bell-shaped, resembling a normal distribution centered around 40-45 minutes. Most delivery times fall near the center, with fewer very short or very long times
> |
```

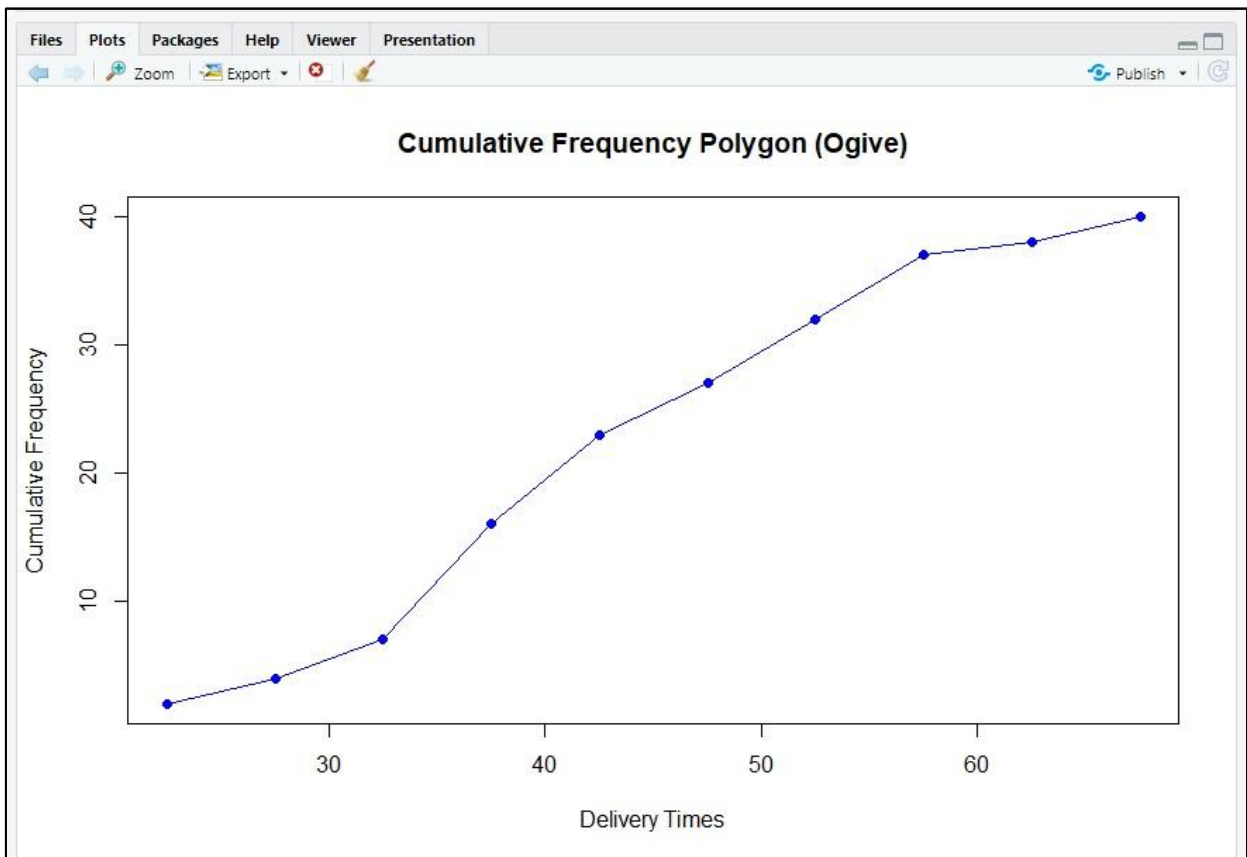
4. Draw a cumulative frequency polygon (ogive) for the data in a separate plot.

```
23 ##4
24 hist_data <- hist(Delivery_Times$Delivery,
25                   breaks = seq(20, 70, by = 5),
26                   right = FALSE,
27                   plot = FALSE)
28 cumulative_freq <- cumsum(hist_data$counts)
29
30
31 plot(hist_data$mids, cumulative_freq,
32      type = "o",
33      main = "Cumulative Frequency Polygon (ogive)",
34      xlab = "Delivery Times",
35      ylab = "Cumulative Frequency",
36      pch = 16,
37      col = "blue")
38
39
```

```

> ##4
> hist_data <- hist(Delivery_Times$Delivery,
+                   breaks = seq(20, 70, by = 5),
+                   right = FALSE,
+                   plot = FALSE)
> cumulative_freq <- cumsum(hist_data$counts)
> plot(hist_data$midpoints, cumulative_freq,
+      type = "o",
+      main = "Cumulative Frequency Polygon (Ogive)",
+      xlab = "Delivery Times",
+      ylab = "Cumulative Frequency",
+      pch = 16,
+      col = "blue")
>

```



EnvironmentHistoryConnectionsTutorial

Import Dataset

142 MiB

List

RGlobal Environment

Data

Delivery_Times

40 obs. of 1 variable

hist_data

List of 6

values

cumulative_freq

int [1:10] 2 4 7 16 23 27 32 37 38 40