



Faculty of Computing

Year 2 Semester 1 (2025)

IT2120 - Probability and Statistics

Lab Sheet 05

Exercise

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

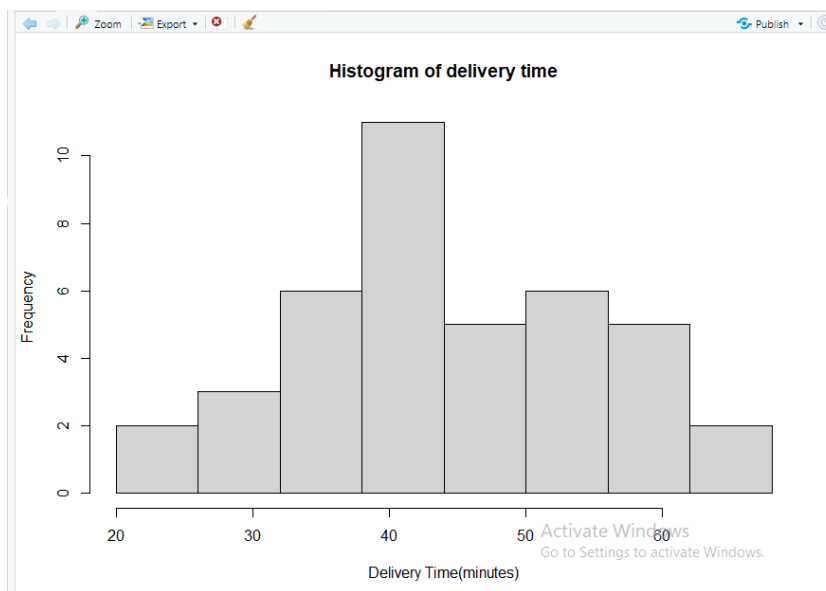
```
#1
#set directory
setwd("C:\\Users\\it24100015\\Desktop\\IT24100015LAB5")
getwd()
#importing data set
Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE, sep = ",")
print(Delivery_Times)
#view the file in a separate window
fix(data)
```

```
> #1
> #set directory
> setwd("C:\\Users\\it24100015\\Desktop\\IT24100015LAB5")
> getwd()
[1] "C:/Users/it24100015/Desktop/IT24100015LAB5"
> #importing data set
> Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE, sep = ",")
> 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
```

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.

```
13 #2
14 hist(Delivery_Times$Delivery_Time,
15 main = "Histogram of delivery time",
16 xlab = "Delivery Time(minutes)",
17 ylab = "Frequency",
18 breaks=seq (20,70, by=6),
19 right = FALSE)
20
```

```
>
> #2
> hist(Delivery_Times$Delivery_Time,
+ main = "Histogram of delivery time",
+ xlab = "Delivery Time(minutes)",
+ ylab = "Frequency",
+ breaks=seq (20,70, by=6),
+ right = FALSE)
>
```



3. Comment on the shape of the distribution.

The distribution of delivery times is right-skewed, with most of delivery times falling between 30 and 45 minutes. A few higher values (above 50 minutes) extend the tail to the right.

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

```
#4
cum_freq <- cumsum(table(cut(Delivery_Times$Delivery_Time, breaks=seq(20,70, by=5), right = FALSE)))
plot(seq(20,65, by=5),cum_freq, type = 'o',
     main = "cumulative frequency polygon(ogive) for Delivery Times",
     xlab="Delivery Time(minutes)",
     ylab="cumulative frequency",
     ylim=c(0,max(cum_freq)),
     pch=16)

+       pch=16)
> #4
> cum_freq <- cumsum(table(cut(Delivery_Times$Delivery_Time, breaks=seq(20,70, by=5), right = FALSE)))
> plot(seq(20,65, by=5),cum_freq, type = 'o',
+       main = "cumulative frequency polygon(ogive) for Delivery Times",
+       xlab="Delivery Time(minutes)",
+       ylab="cumulative frequency",
+       ylim=c(0,max(cum_freq)),
+       pch=16)
> |
```

