SLIIT UNI
THE KNOWLEDGE UNIVERSITY

Faculty of Computing

Year 2 Semester 1 (2025)

IT2120 - Probability and Statistics Lab Sheet 05

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".

```
setwd("C:\\Users\\IT24610818\\Desktop\\Lab05_IT24610818")

#Question 01
Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)

fix(Delivery_Times)

names(Delivery_Times)<-c("X1")

attach(Delivery_Times)
```

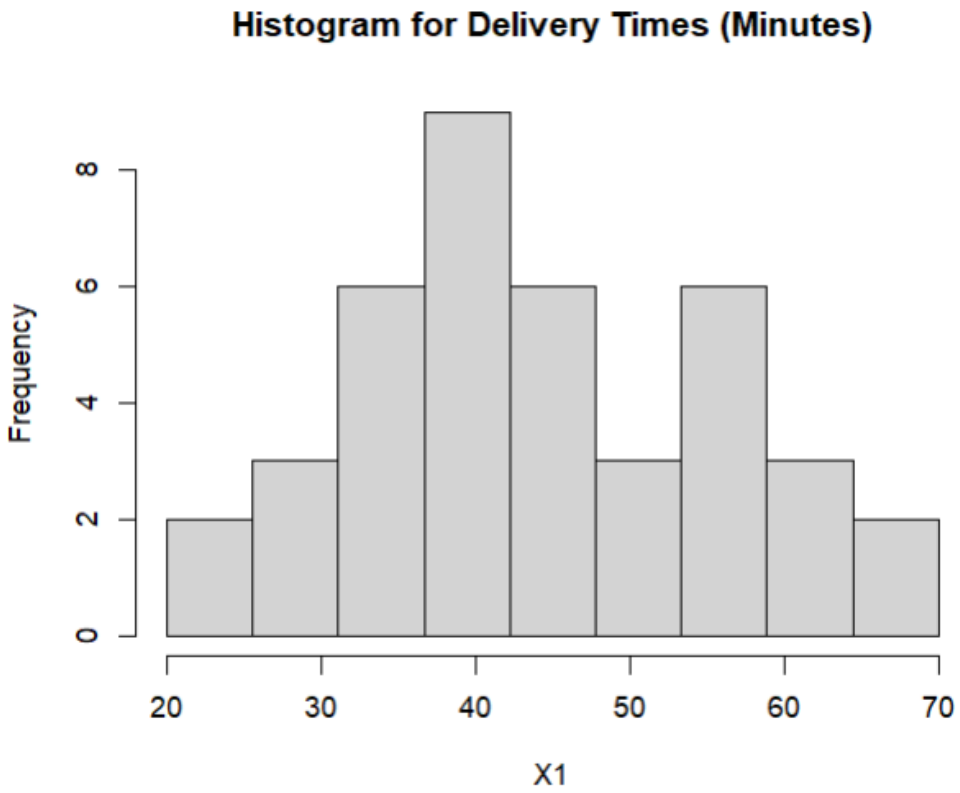
Data Editor

File Edit Help

	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				

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.

```
#Question 02
histogram<-hist(X1,main="Histogram for Delivery Times (Minutes)",breaks = seq(20,70,length = 10),right = FALSE)
```



3. Comment on the shape of the distribution.

```
# ---Question 03---
# The histogram shows that delivery times are approximately symmetric.
# Most delivery times fall between 35 and 45 minutes.
# The shape is bell-shaped, resembling a normal distribution.
# There are fewer observations at both the lower and upper ends.
```

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

```
#Question04
breaks<-round(histogram$breaks)
freq <- histogram$counts

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]
  }
}

plot(breaks,new,type='l',main = 'Cumalative Frequency Polygon for Delivery Times',
     xlab="Delivery Times",ylab="Cumalative Frequency",ylim=c(0,max(cum.freq)))
```

```
> cbind(Upper = breaks, CumFreq = new)
      Upper CumFreq
 [1,]    20       0
 [2,]    26       2
 [3,]    31       5
 [4,]    37      11
 [5,]    42      20
 [6,]    48      26
 [7,]    53      29
 [8,]    59      35
 [9,]    64      38
[10,]    70      40
```