Sri Lanka Institute of Information Technology

Lab Submission
Lab sheet No5

IT24100861

Fernando M.S.F.

Probability and Statistics | IT2120

B.Sc. (Hons) in Information Technology

Exercise

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

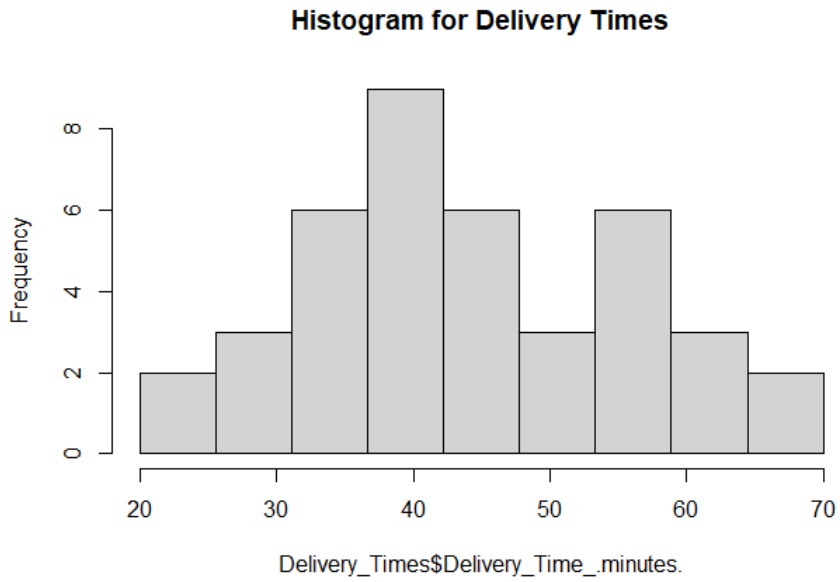
```
setwd("C:\\Users\\it24100861\\Desktop\\IT_24100861")

#1
Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)
names(Delivery_Times)
```

```
> setwd("C:\\Users\\it24100861\\Desktop\\IT_24100861")
> Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)
> names(Delivery_Times)
[1] "Delivery_Time_.minutes."
```

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.

```
#2
histogram1 <- hist(
  Delivery_Times$Delivery_Time_.minutes.,
  main = "Histogram for Delivery Times",
  breaks = seq(20, 70, length = 10),
  right = FALSE)
```

```
> histogram1 <- hist(
+   Delivery_Times$Delivery_Time_.minutes.,
+   main = "Histogram for Delivery Times",
+   breaks = seq(20, 70, length = 10),
+   right = FALSE)
```



3. Comment on the shape of the distribution.

   The histogram is mostly symmetric but slightly skewed 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)
```

```
> 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)
```