IT2120 - Probability and Statistics

Lab Sheet 05

IT24103512

SENEVIRATHNA M.M.C.M

Exercise

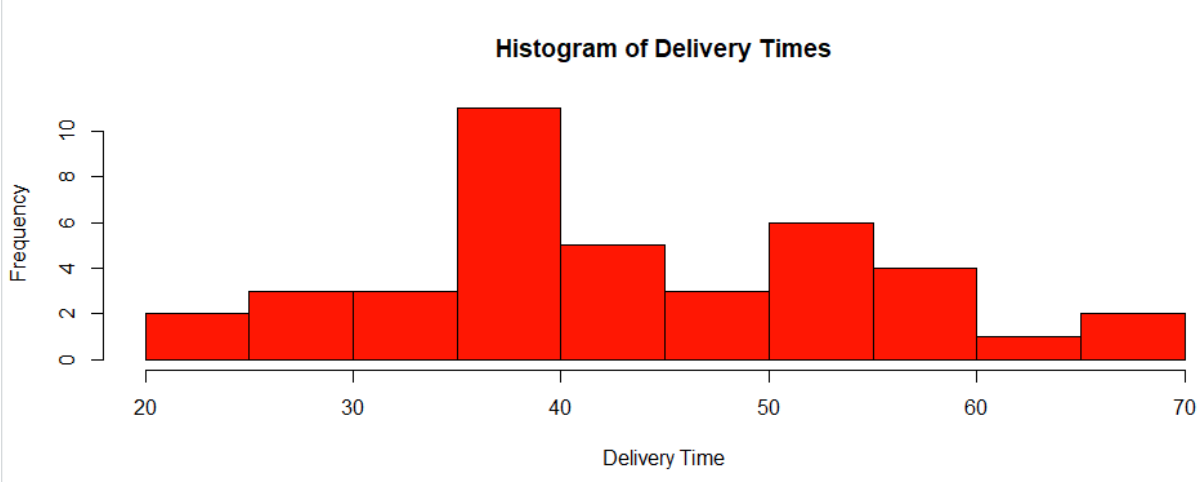
Instructions: Create a folder in your desktop with your registration number (Eg: ”IT.....”). You need to save the R script file and take screenshots of the command prompt with answers and save it in a word document inside the folder. Save both R script file and word document with your registration number (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”.

```
1  setwd("C:\\Users\\it24103512\\Desktop\\IT24103512")
2  Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)
> Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)
```

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.

```
Delivery_Times$Delivery_Time_.minutes. <- as.numeric(Delivery_Times$Delivery_Time_.minutes.)
hist(Delivery_Times$Delivery_Time_.minutes.,
     breaks = seq(20, 70, by = 5),
     right = TRUE,
     col = "RED",
     main = "Histogram of Delivery Times",
     xlab = "Delivery Time",
     ylab = "Frequency")
> Delivery_Times$Delivery_Time_.minutes. <- as.numeric(Delivery_Times$Delivery_Time_.minutes.)
> hist(Delivery_Times$Delivery_Time_.minutes.,
+      breaks = seq(20, 70, by = 5),
+      right = TRUE,
+      col = "RED",
+      main = "Histogram of Delivery Times",
+      xlab = "Delivery Time",
+      ylab = "Frequency")
>
```



3. Comment on the shape of the distribution

   The distribution of delivery times appears to be unimodal, with a peak around 40 days, indicating the most frequent delivery time. It is slightly skewed to the right (positively skewed), as the tail extends more toward the higher delivery times (up to 70 days) compared to the lower end (around 20 days).

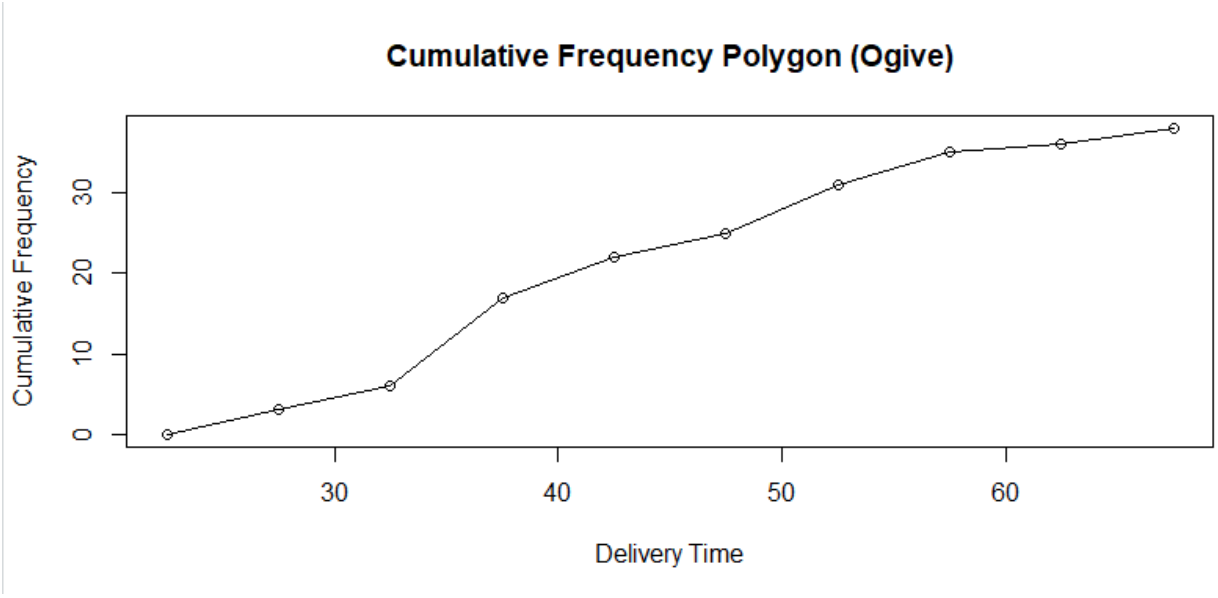
4. Draw a cumulative frequency polygon (ogive) for the data in a separate plot.

```
cf <- cumsum(table(cut(Delivery_Times$Delivery_Time_.minutes.,breaks = seq(20, 70, by = 5), right = TRUE)))

plot(seq(22.5, 67.5, by = 5), cf, type = "o", col = "black",
     xlab = "Delivery Time", ylab = "Cumulative Frequency",
     main = "Cumulative Frequency Polygon (Ogive)")
```


```
cf <- cumsum(table(cut(Delivery_Times$Delivery_Time_.minutes.,breaks = seq(20, 70, by = 5), right = TRUE)))

plot(seq(22.5, 67.5, by = 5), cf, type = "o", col = "black",
     xlab = "Delivery Time", ylab = "Cumulative Frequency",
     main = "Cumulative Frequency Polygon (Ogive)")
```


Cumulative Frequency Polygon (Ogive)
Cumulative Frequency
0
10
20
30
30
40
50
60
Delivery Time