Sri Lanka Institute of Information Technology

Lab Submission 05

IT24104128

Palapagama T.H.P

IT2120 - Probability and Statistics

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".

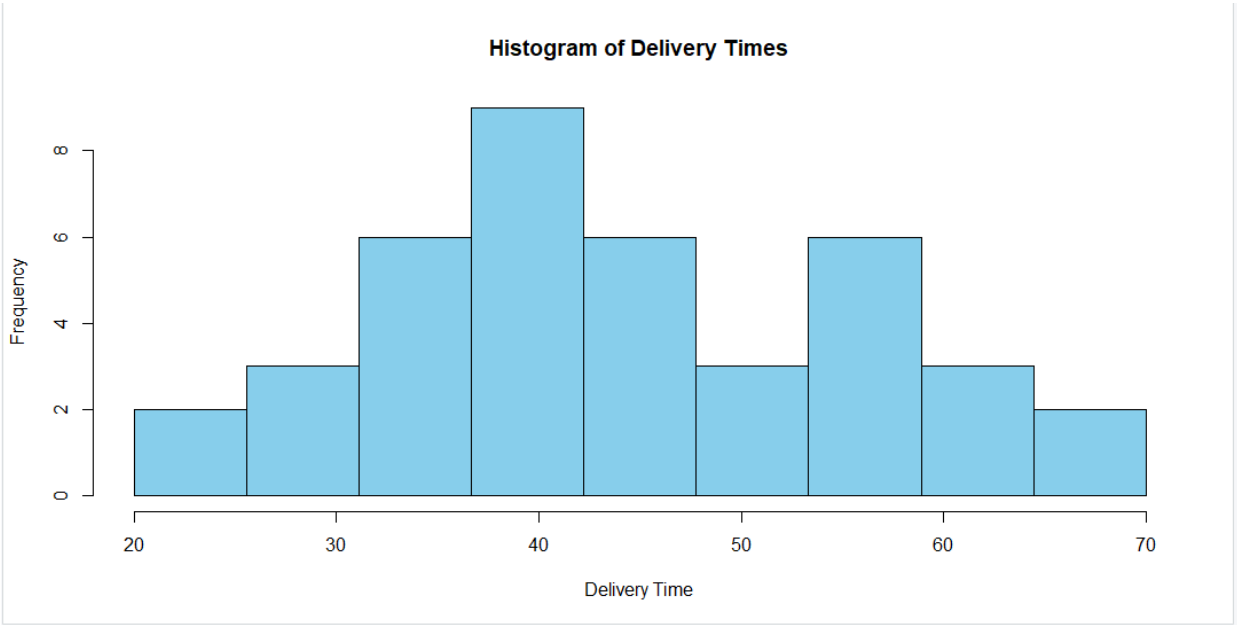
```
setwd("C:\\Users\\it24104128\\Desktop\\Lab 05-20250828 (1)")
Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)

colnames(Delivery_Times) <- "Delivery_Time"
head(Delivery_Times)
str(Delivery_Times)

breaks <- seq(20, 70, length.out = 10)
hist(Delivery_Times$Delivery_Time, right = FALSE, breaks = breaks,
     main = "Histogram of Delivery Times", xlab = "Delivery Time", ylab = "Frequency",
     col = "skyblue")
```

```
> setwd("C:\\Users\\it24104128\\Desktop\\Lab 05-20250828 (1)")
> Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)
>
> colnames(Delivery_Times) <- "Delivery_Time"
> head(Delivery_Times)
  Delivery_Time
1            34
2            54
3            47
4            29
5            39
6            61
> str(Delivery_Times)
'data.frame':   40 obs. of  1 variable:
 $ Delivery_Time: int  34 54 47 29 39 61 20 40 57 36 ...
>
```

```
hist_data <- hist(Delivery_Times$Delivery_Time, breaks = breaks, right = FALSE, plot = FALSE)
frequencies <- hist_data$counts
cum_freq <- cumsum(frequencies)
print(frequencies)
print(cum_freq)
```

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.

```
> hist_data <- hist(Delivery_Times$Delivery_Time, breaks = breaks, right = FALSE, plot = FALSE)
> frequencies <- hist_data$counts
> cum_freq <- cumsum(frequencies)
> print(frequencies)
[1] 2 3 6 9 6 3 6 3 2
> print(cum_freq)
[1]  2  5 11 20 26 29 35 38 40
>
```



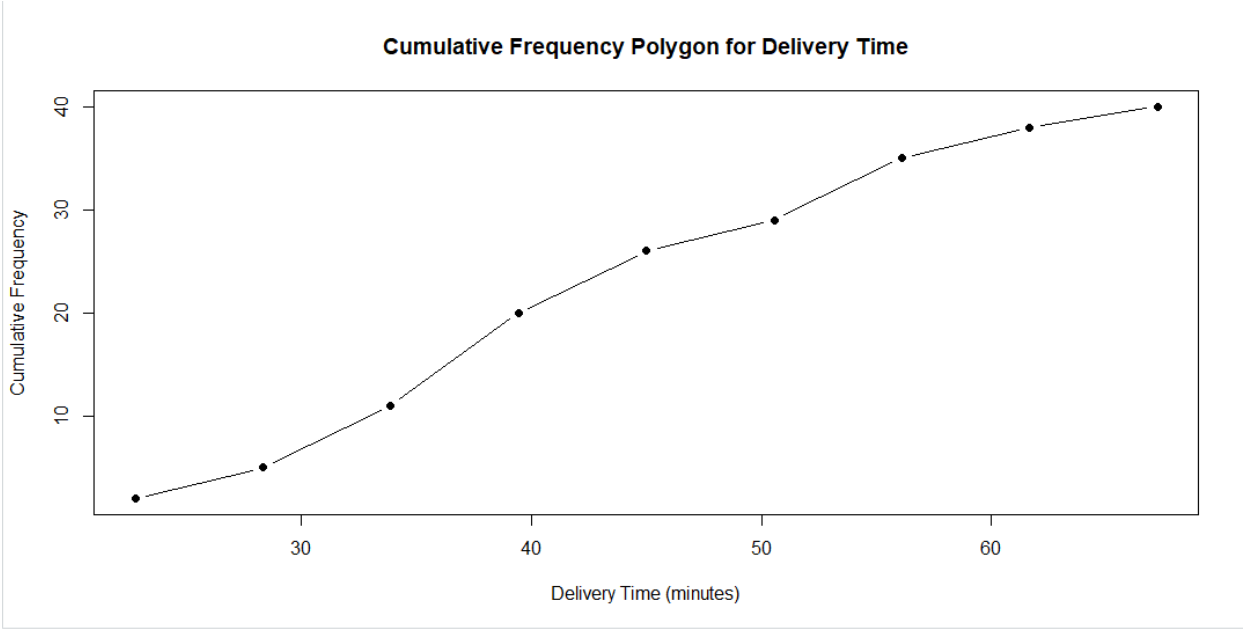
3. Comment on the shape of the distribution.

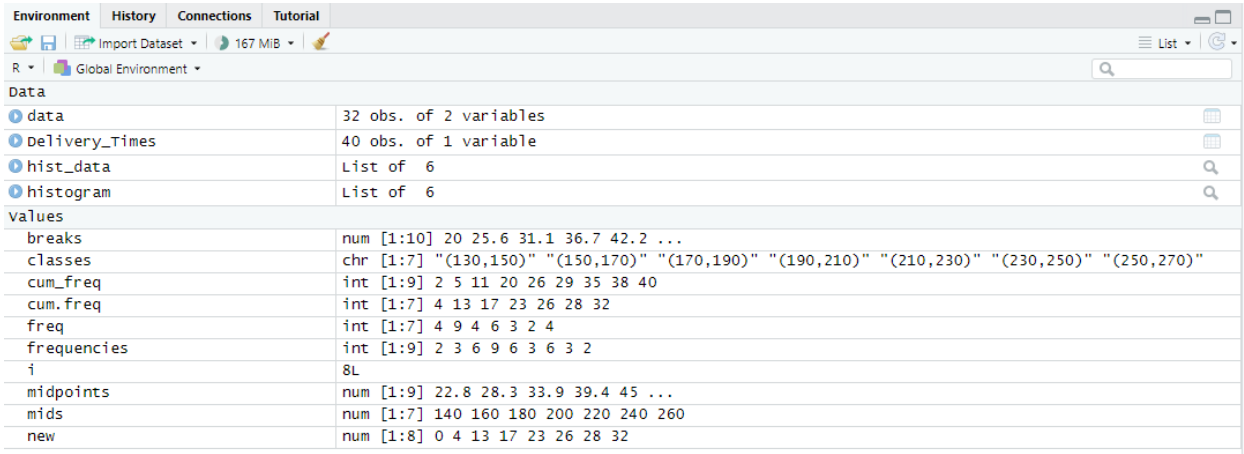
*The distribution is right-skewed (positively skewed)

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

```
midpoints <- hist_data$mids
plot(midpoints, cum_freq, type = "b",
     main = "Cumulative Frequency Polygon for Delivery Time",
     xlab = "Delivery Time (minutes)",
     ylab = "Cumulative Frequency",
     pch = 16)
```

```
> midpoints <- hist_data$mids
> plot(midpoints, cum_freq, type = "b",
+      main = "Cumulative Frequency Polygon for Delivery Time",
+      xlab = "Delivery Time (minutes)",
+      ylab = "Cumulative Frequency",
+      pch = 16)
>
```




Environment
History
Connections
Tutorial
Import Dataset
167 MiB
List
R
Global Environment
Data
data
32 obs. of 2 variables
Delivery_Times
40 obs. of 1 variable
hist_data
List of 6
histogram
List of 6
values
breaks
num [1:10] 20 25.6 31.1 36.7 42.2 ...
classes
chr [1:7] "(130,150)" "(150,170)" "(170,190)" "(190,210)" "(210,230)" "(230,250)" "(250,270)"
cum_freq
int [1:9] 2 5 11 20 26 29 35 38 40
cum.freq
int [1:7] 4 13 17 23 26 28 32
freq
int [1:7] 4 9 4 6 3 2 4
frequencies
int [1:9] 2 3 6 9 6 3 6 3 2
i
8L
midpoints
num [1:9] 22.8 28.3 33.9 39.4 45 ...
mids
num [1:7] 140 160 180 200 220 240 260
new
num [1:8] 0 4 13 17 23 26 28 32