# Probability and Statistics – IT2120

## LabSheet 05

### Karunarathna R.N.D.W. – IT24102730

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

```
setwd("C:\\Users\\user\\Desktop\\IT24102730")
Delivery_Times = read.table("Exercise - Lab 05.txt", header = TRUE, sep = ",")
fix(Delivery_Times)
attach(Delivery_Times)
```

```
> setwd("C:\\Users\\user\\Desktop\\IT24102730")
> Delivery_Times = read.table("Exercise - Lab 05.txt", header = TRUE, sep = ",")
> fix(Delivery_Times)
```

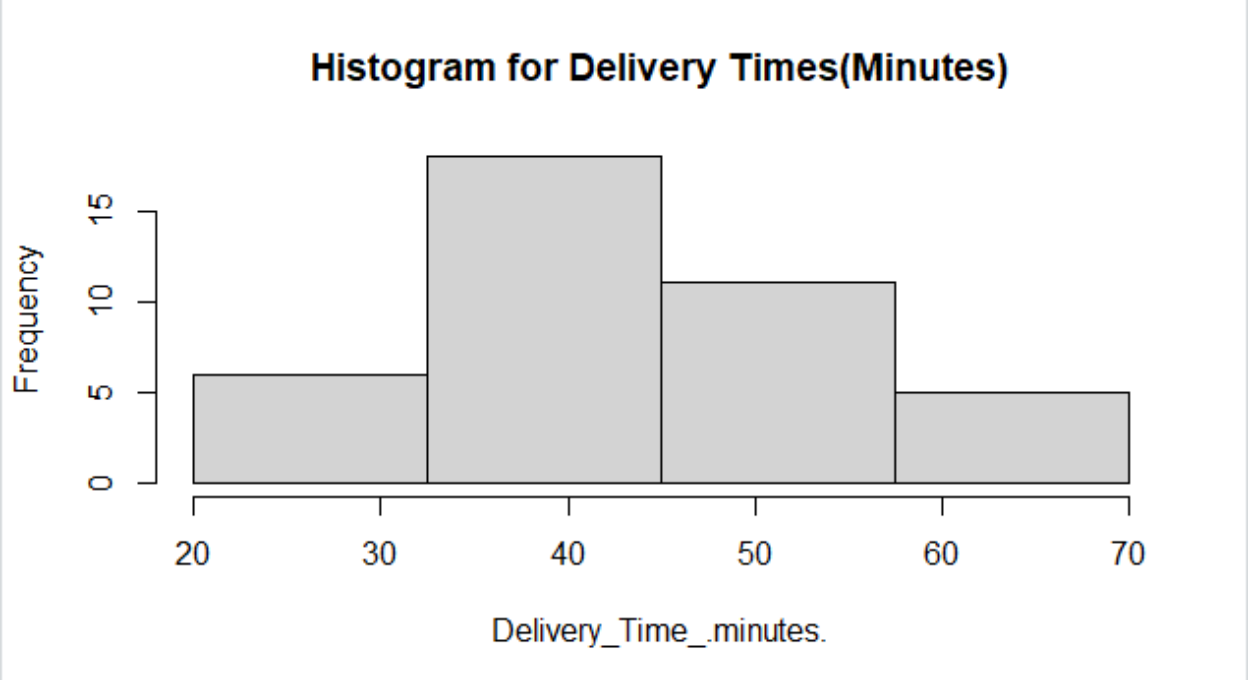
Data Editor

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.

```
histogram<-hist(Delivery_Time_.minutes.,main = "Histogram for Delivery Times(Minutes)", breaks = seq(20, , length = 5), right = TRUE)
```




3. Comment on the shape of the distribution.

   The distribution has one clear peak so most of the data values are around a single range with most delivery times between 35 to 55 minutes and overall range is 20 to 70minutes.

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

```
names(Delivery_Times)[1] <- "DeliveryTime"

str(Delivery_Times)

hist_data <- hist(Delivery_Times$DeliveryTime,
                  breaks = seq(20, 70, length.out = 10),
                  right = FALSE,
                  plot = FALSE)
cum_freq <- cumsum(hist_data$counts)
breaks <- hist_data$breaks

plot(breaks, c(0, cum_freq),
     type = "l",
     main = "Cumulative Frequency Polygon (Ogive)",
     xlab = "Delivery Time",
     ylab = "Cumulative Frequency",
     ylim = c(0, max(cum_freq)))
```

```
> names(Delivery_Times)[1] <- "DeliveryTime"
> str(Delivery_Times)
'data.frame':   40 obs. of  1 variable:
 $ DeliveryTime: num  34 54 47 29 39 61 20 40 57 36 ...
> hist_data <- hist(Delivery_Times$DeliveryTime,
+                   breaks = seq(20, 70, length.out = 10),
+                   right = FALSE,
+                   plot = FALSE)
> cum_freq <- cumsum(hist_data$counts)
> breaks <- hist_data$breaks
> plot(breaks, c(0, cum_freq),
+      type = "l",
+      main = "Cumulative Frequency Polygon (Ogive)",
+      xlab = "Delivery Time",
+      ylab = "Cumulative Frequency",
+      ylim = c(0, max(cum_freq)))
>
```