Sri Lanka Institute of Information Technology

Probability and Statistics

IT2120

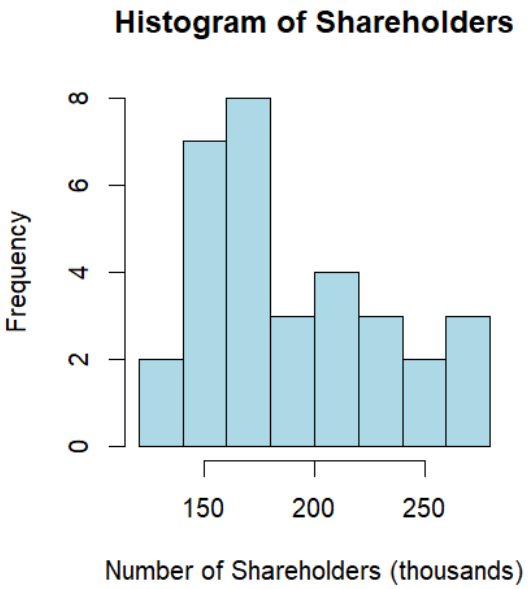
Lab Submission
Labsheet 05

IT23453074

Rusiruka R.A.C

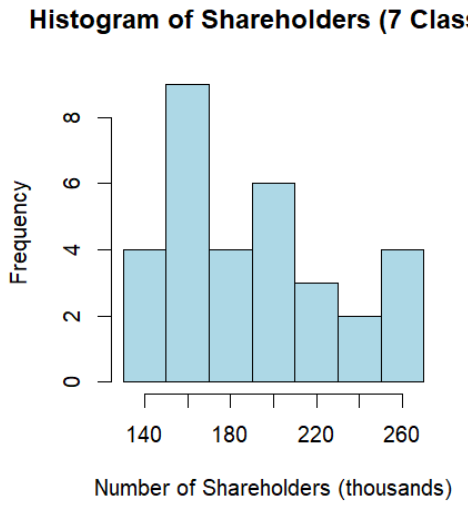
```r
setwd("C:/Users/rusir/OneDrive/Desktop/IT23453074")

# Read the data (it's comma-separated with a header)
shareholders_data <- read.csv("Data.txt", header = TRUE)

# Extract the shareholder numbers into a vector
shareholders <- shareholders_data$Number_of_Shareholders.thousands.

# View the data to confirm
print(shareholders)

# Default histogram
hist(shareholders, main = "Histogram of Shareholders", xlab = "Number of Shareholders (thousands)", col = "lightblue")
```

```
R 4.5.1 · C:/Users/rusir/OneDrive/Desktop/IT23453074/
> setwd("C:/Users/rusir/OneDrive/Desktop/IT23453074")
> shareholders_data <- read.csv("Data.txt", header = TRUE)
> shareholders <- shareholders_data$Number_of_Shareholders.thousands.
> print(shareholders)
 [1] 144 266 177 133 209 264 160 143 246 151 239 204 204 195 176
[16] 175 200 173 195 220 251 137 150 262 158 162 165 223 158 162
[31] 225 161
> hist(shareholders, main = "Histogram of Shareholders", xlab = "Number of Shareholders (thousands)", col = "lightblue")
```



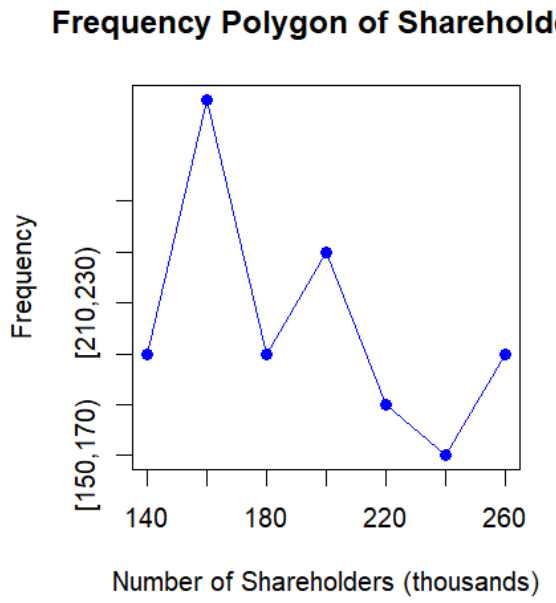

```r
# Define breaks for seven classes: 130 to 150, 150 to 170, ..., 250 to 270
breaks <- seq(130, 270, by = 20)

# Histogram with specified breaks
hist(shareholders, breaks = breaks, main = "Histogram of Shareholders (7 Classes)", xlab = "Number of Shareholders (thousands)", col =

# Create frequency distribution
classes <- cut(shareholders, breaks = breaks, include.lowest = TRUE, right = FALSE)
freq_table <- table(classes)

# Display the table
print(freq_table)

# Get midpoints of the classes
midpoints <- seq(140, 260, by = 20)  # Midpoints: 140, 160, ..., 260
```

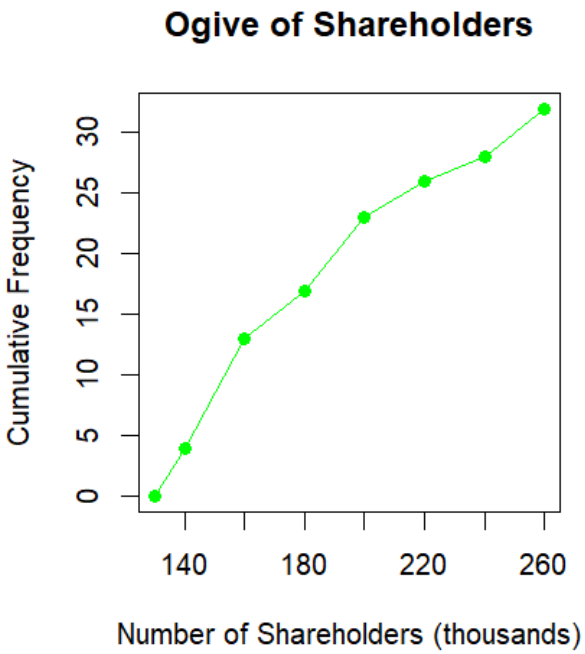
```
> breaks <- seq(130, 270, by = 20)
> hist(shareholders, breaks = breaks, main = "Histogram of Shareholders (7 Classes)", xlab = "Number of Shareholders (thousands)", col = "lig
htblue", right = FALSE)  # right=FALSE for left-closed intervals
> classes <- cut(shareholders, breaks = breaks, include.lowest = TRUE, right = FALSE)
> freq_table <- table(classes)
> print(freq_table)
classes
[130,150) [150,170) [170,190) [190,210) [210,230) [230,250)
        4         9         4         6         3         2
[250,270]
        4
> midpoints <- seq(140, 260, by = 20)  # Midpoints: 140, 160, ..., 260
```




```r
# Plot frequency polygon
plot(midpoints, freq_table, type = "o", main = "Frequency Polygon of Shareholders", xlab = "Number of Shareholders (thousands)", ylab =

# Cumulative frequencies
cum_freq <- cumsum(freq_table)

# Plot ogive (add 0 at the start for the lower bound)
plot(c(130, midpoints), c(0, cum_freq), type = "o", main = "Ogive of Shareholders", xlab = "Number of Shareholders (thousands)", ylab =

# Import the data (assuming it's tab-separated or space-separated; adjust sep if needed)
delivery_times <- read.table("Exercise - Lab 05.txt", header = TRUE, sep = ",")

# If it's one column, extract to a vector
delivery <- delivery_times$Delivery_Time_.minutes.

# View the data to confirm
print(delivery)
```

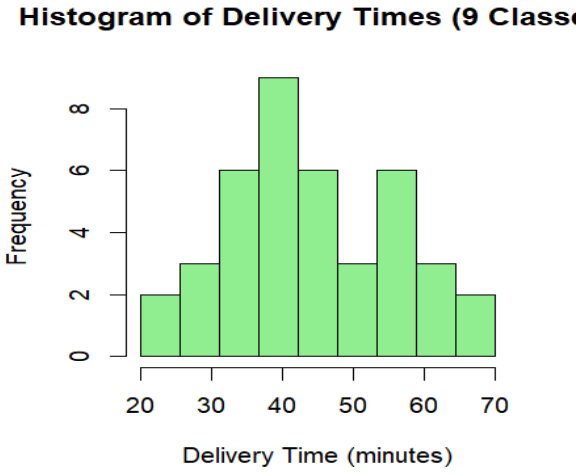
```
> plot(midpoints, freq_table, type = "o", main = "Frequency Polygon of Shareholders", xlab = "Number of Shareholders (thousands)", ylab = "F
equency", col = "blue", pch = 19)
> cum_freq <- cumsum(freq_table)
> plot(c(130, midpoints), c(0, cum_freq), type = "o", main = "Ogive of Shareholders", xlab = "Number of Shareholders (thousands)", ylab = "C
mulative Frequency", col = "green", pch = 19)
> delivery_times <- read.table("Exercise - Lab 05.txt", header = TRUE, sep = ",")
> delivery <- delivery_times$Delivery_Time_.minutes.
> print(delivery)
```






```r
# Define breaks for nine classes
breaks_delivery <- seq(20, 70, length.out = 10)  # 10 points for 9 intervals

# Histogram
hist(delivery, breaks = breaks_delivery, main = "Histogram of Delivery Times (9 Classes)", xlab = "Delivery Time (minutes)", col = "lig

# Frequency table
classes_delivery <- cut(delivery, breaks = breaks_delivery, include.lowest = TRUE, right = FALSE)
freq_table_delivery <- table(classes_delivery)

# Cumulative frequencies
cum_freq_delivery <- cumsum(freq_table_delivery)

# Midpoints (average of breaks)
midpoints_delivery <- (breaks_delivery[-1] + breaks_delivery[-length(breaks_delivery)]) / 2

# Plot ogive
plot(c(20, midpoints_delivery), c(0, cum_freq_delivery), type = "o", main = "Ogive of Delivery Times", xlab = "Delivery Time (minutes)"
```

```
> print(delivery)
 [1] 34 54 47 29 39 61 20 40 57 36 38 44 59 38 40 40 67 66 55 48
[21] 52 59 35 56 32 38 54 30 43 36 42 20 27 38 54 43 45 51 36 47
> breaks_delivery <- seq(20, 70, length.out = 10)
> hist(delivery, breaks = breaks_delivery, main = "Histogram of Delivery Times (9 Classes)", xlab = "Delivery Time (minutes)", col = "lightgr
een", right = FALSE)
> classes_delivery <- cut(delivery, breaks = breaks_delivery, include.lowest = TRUE, right = FALSE)
> freq_table_delivery <- table(classes_delivery)
> cum_freq_delivery <- cumsum(freq_table_delivery)
> midpoints_delivery <- (breaks_delivery[-1] + breaks_delivery[-length(breaks_delivery)]) / 2
> plot(c(20, midpoints_delivery), c(0, cum_freq_delivery), type = "o", main = "Ogive of Delivery Times", xlab = "Delivery Time (minutes)", yl
```



```
> plot(c(20, midpoints_delivery), c(0, cum_freq_delivery), type = "o", main = "Ogive of Delivery Times", xlab = "Delivery Time (minutes)", yl
ab = "Cumulative Frequency", col = "red", pch = 19)
> dev.off()
null device
```