

IT2120 - Probability and Statistics

Lab Sheet 05

IT24102293

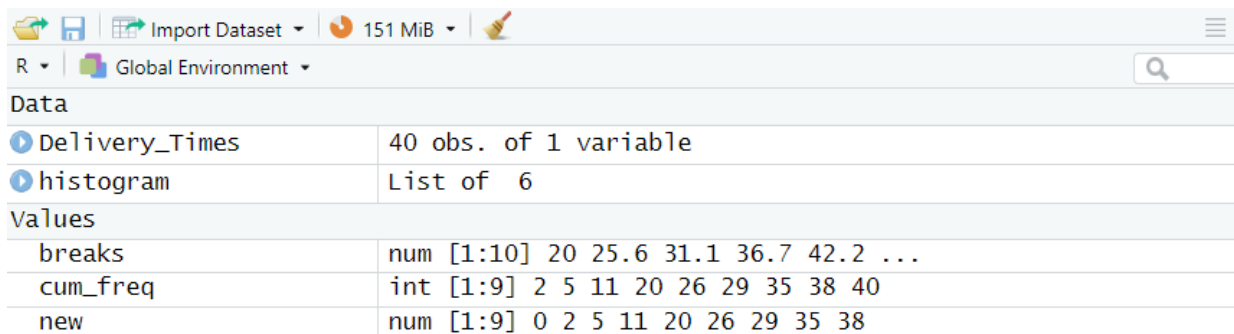
Chandrasiri M.L.M.M

1. Import the dataset ('Exercise - Lab 05.txt') into R and store it in a data frame called "Delivery Times".

```
2  
3 Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)  
4
```

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.

```
# Create histogram  
histogram <- hist(Delivery_Time,  
                  main = "Histogram for Delivery Times",  
                  breaks = seq(20, 70, length = 10),  
                  right = FALSE, |  
                  xlab = "Delivery Time (minutes)")
```



The screenshot shows the R Studio interface. The top toolbar includes icons for file operations and a memory usage indicator of 151 MiB. The 'Global Environment' pane is empty. The 'Data' pane shows two objects: 'Delivery_Times' (40 observations of 1 variable) and 'histogram' (a list of 6). The 'Values' pane displays the internal structure of the histogram object, including breaks, cumulative frequencies, and new frequencies.

Data	
Delivery_Times	40 obs. of 1 variable
histogram	List of 6
Values	
breaks	num [1:10] 20 25.6 31.1 36.7 42.2 ...
cum_freq	int [1:9] 2 5 11 20 26 29 35 38 40
new	num [1:9] 0 2 5 11 20 26 29 35 38



3. Comment on the shape of the distribution.

- The distribution is roughly symmetric with a peak around 40 minutes, but the right side is a little higher than the left.

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

```
# Create cumulative frequency polygon
new <- c(0, cum_freq[-length(cum_freq)]) # Prepend 0 and shift values
plot(breaks, c(new, cum_freq[length(cum_freq)]),
     type = 'l',
     main = "Cumulative Frequency Polygon for Delivery Times",
     xlab = "Delivery Time (minutes)",
     ylab = "Cumulative Frequency",
     ylim = c(0, max(cum_freq)))
points(breaks, c(new, cum_freq[length(cum_freq)]), pch = 16)

# Combine breaks and cumulative frequencies
cbind(Upper = breaks, CumFreq = c(new, cum_freq[length(cum_freq)]))
```

(Top Level) ↕

```

>
> # Combine breaks and cumulative frequencies
> cbind(Upper = breaks, CumFreq = c(new, cum_freq[length(cum_freq)]))
      Upper CumFreq
[1,] 20.00000      0
[2,] 25.55556      2
[3,] 31.11111      5
[4,] 36.66667     11
[5,] 42.22222     20
[6,] 47.77778     26
[7,] 53.33333     29
[8,] 58.88889     35
[9,] 64.44444     38
[10,] 70.00000     40
> |

```

R	Global Environment
Data	
Delivery_Times	40 obs. of 1 variable
histogram	List of 6
Values	
breaks	num [1:10] 20 25.6 31.1 36.7 42.2 ...
cum_freq	int [1:9] 2 5 11 20 26 29 35 38 40
new	num [1:9] 0 2 5 11 20 26 29 35 38

