IT2120

Probability and Statistics

Year 2 – Semester 1

Lab Report - 05

Submitted to

Sri Lanka Institute of Information Technology

In partial fulfillment of the requirements for the

Bachelor of Science Special Honors Degree in Information Technology

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins

PS lab 3.R | PS lab 3.R | Untitled1* | Exercise...Lab.05 | `Delivery Times`

Source on Save | Run | Source

```r
setwd("C:\\Users\\Piyumi Samaraweera\\Downloads\\IT24103858")
data <- read.table("Exercise - Lab 05.txt", header = TRUE)
names(data) <- c("Delivery_Time")
attach(data)

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

breaks <- round(histogram$breaks)
freq <- histogram$counts
mids <- histogram$mids

classes <- c()
for (i in 1:(length(breaks)-1)) {
  classes[i] <- paste0("[", breaks[i], ",", breaks[i+1], ")")
}
freq_dist <- cbind(Classes = classes, Frequency = freq)
print(freq_dist)

plot(mids, freq, type = 'l', main = "Frequency Polygon for Delivery Times", xlab = "Delivery Time (minutes)", ylab = "Frequency", ylim = c(0, max(freq)))

cum_freq <- cumsum(freq)
new <- c()
for (i in 1:length(breaks)) {
  if (i == 1) {
    new[i] = 0
  } else {
    new[i] = cum_freq[i-1]
  }
}
plot(breaks, new, type = 'l', main = "Cumulative Frequency Polygon (Ogive)", xlab = "Delivery Time (minutes)", ylab = "Cumulative Frequency", ylim = c(0, max(cum_freq)))
cbind(Upper = breaks, CumFreq = new)
```

Filter

	Delivery_Time_.minutes.
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
20	48
21	52
22	59
23	35
24	56
25	32
26	38
27	54
28	30
29	43
30	36

Showing 1 to 30 of 40 entries, 1 total columns

Console

Source

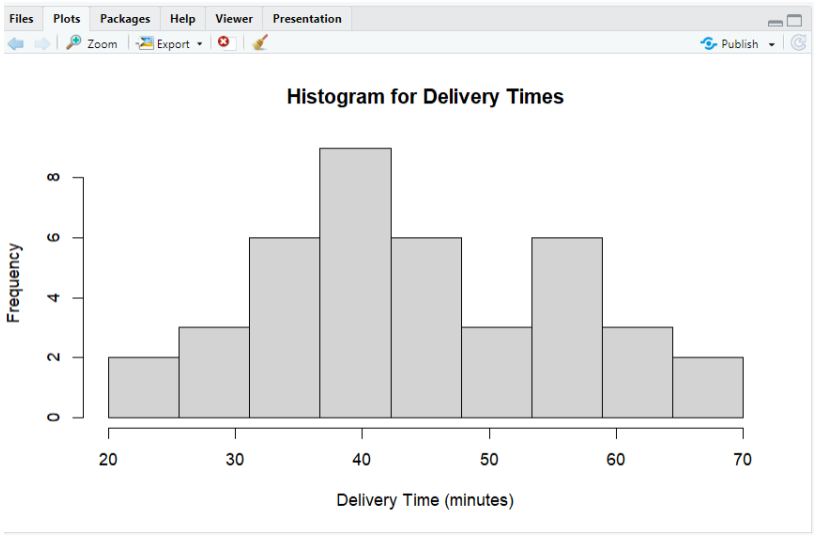
Console | Background Jobs

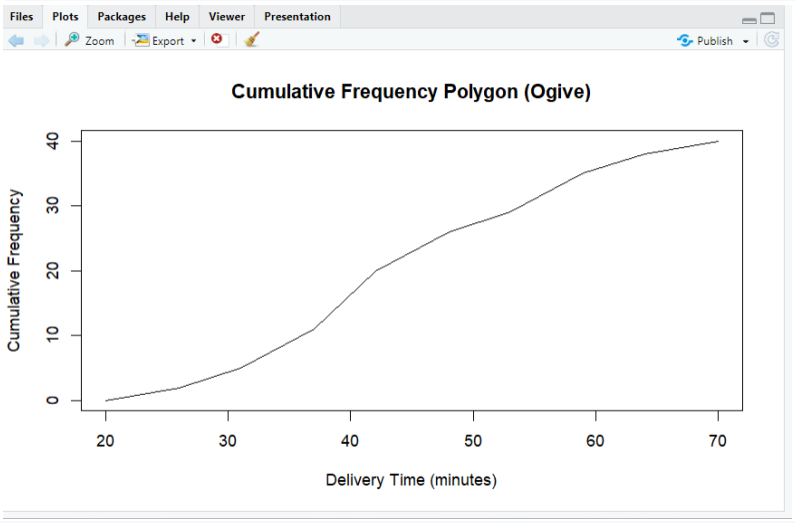
R 4.5.0 · C:/Users/Piyumi Samaraweera/Downloads/IT24103858/

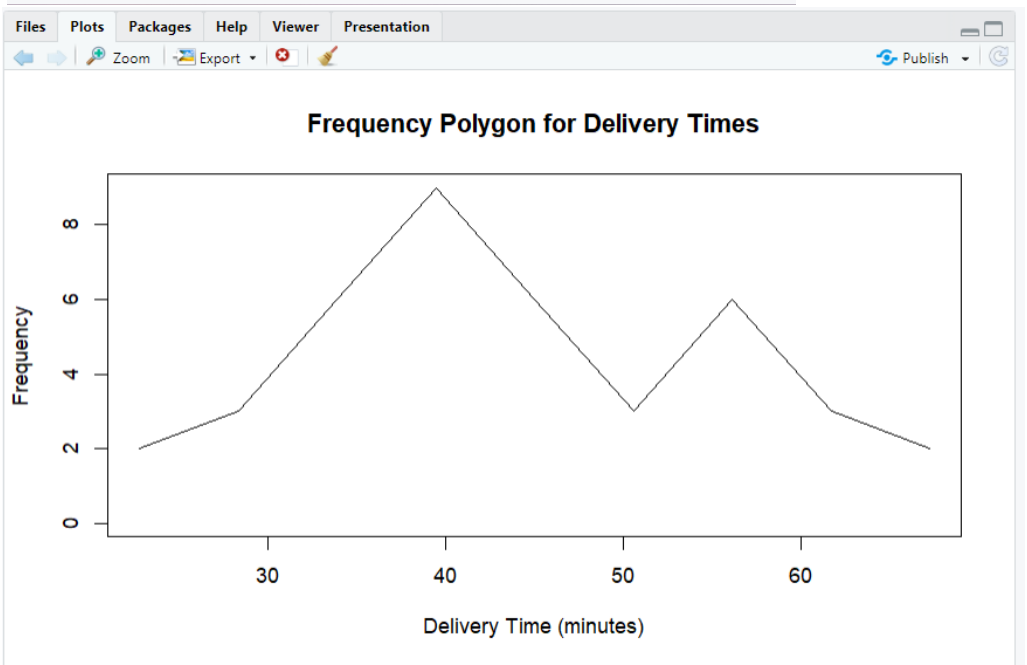
```
> setwd("C:\\Users\\Piyumi Samaraweera\\Downloads\\IT24103858")
> data <- read.table("Exercise - Lab 05.txt", header = TRUE)
> names(data) <- c("Delivery_Time")
> attach(data)
The following object is masked from data (pos = 3):

    Delivery_Time

The following object is masked from data (pos = 4):

    Delivery_Time

> histogram <- hist(Delivery_Time, main = "Histogram for Delivery Times", breaks = seq(20, 70, length = 10), right = FALSE, xlab = "Delivery Time (minutes)")
> breaks <- round(histogram$breaks)
> freq <- histogram$counts
> mids <- histogram$mids
> classes <- c()
> for (i in 1:(length(breaks)-1)) {
+   classes[i] <- paste0("[", breaks[i], ",", breaks[i+1], ")")
+ }
> freq_dist <- cbind(Classes = classes, Frequency = freq)
> print(freq_dist)
      Classes   Frequency
 [1,] "[20,26)" "2"
 [2,] "[26,31)" "3"
 [3,] "[31,37)" "6"
 [4,] "[37,42)" "9"
 [5,] "[42,48)" "6"
 [6,] "[48,53)" "3"
 [7,] "[53,59)" "6"
 [8,] "[59,64)" "3"
 [9,] "[64,70)" "2"


> plot(mids, freq, type = 'l', main = "Frequency Polygon for Delivery Times", xlab = "Delivery Time (minutes)", ylab = "Frequency", ylim = c(0, max(freq)))
> cum_freq <- cumsum(freq)
> new <- c()
> for (i in 1:length(breaks)) {
+   if (i == 1) {
+     new[i] = 0
+   } else {
+     new[i] = cum_freq[i-1]
+   }
+ }
> plot(breaks, new, type = 'l', main = "Cumulative Frequency Polygon (Ogive)", xlab = "Delivery Time (minutes)", ylab = "Cumulative Frequency", ylim = c(0, max(cum_freq)))
> plot(mids, freq, type = 'l', main = "Frequency Polygon for Delivery Times", xlab = "Delivery Time (minutes)", ylab = "Frequency", ylim = c(0, max(freq)))
> plot(breaks, new, type = 'l', main = "Cumulative Frequency Polygon (Ogive)", xlab = "Delivery Time (minutes)", ylab = "Cumulative Frequency", ylim = c(0, max(cum_freq)))
> cbind(Upper = breaks, CumFreq = new)
      Upper CumFreq
 [1,]    20       0
 [2,]    26       2
 [3,]    31       5
 [4,]    37      11
 [5,]    42      20
 [6,]    48      26
 [7,]    53      29
 [8,]    59      35
 [9,]    64      38
[10,]    70      40
>
```

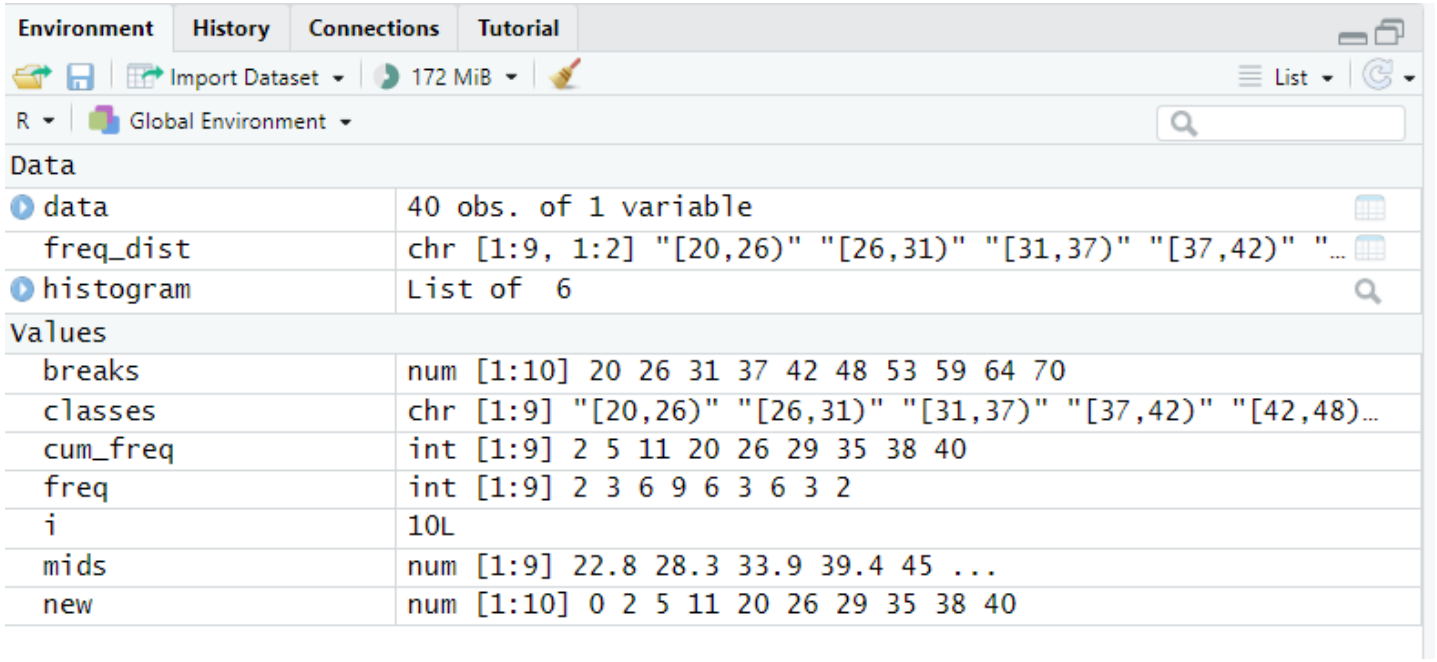
3) Comment on the shape of the distribution

The distribution has 8 classes and the class with the highest frequency is the 4th one.








Environment
History
Connections
Tutorial
Import Dataset
172 MiB
List
R
Global Environment
Data
data
40 obs. of 1 variable
freq_dist
chr [1:9, 1:2] "[20,26)" "[26,31)" "[31,37)" "[37,42)" "...
histogram
List of 6
Values
breaks
num [1:10] 20 26 31 37 42 48 53 59 64 70
classes
chr [1:9] "[20,26)" "[26,31)" "[31,37)" "[37,42)" "[42,48)...
cum_freq
int [1:9] 2 5 11 20 26 29 35 38 40
freq
int [1:9] 2 3 6 9 6 3 6 3 2
i
10L
mids
num [1:9] 22.8 28.3 33.9 39.4 45 ...
new
num [1:10] 0 2 5 11 20 26 29 35 38 40