Faculty of Computing

Year 2 Semester 1 (2025)

IT2120 - Probability and Statistics

Lab Sheet 05

IT24103878

Amarakoon M.B.V.B.A

```r
#1
getwd()
setwd("C:\\Users\\It24103878\\Desktop\\IT24103878")
getwd()
Delivery_Times <- read.table('Exercise - Lab 05.txt', header = TRUE)
```

```
> #1
> getwd()
[1] "C:/Users/It24103878/Desktop/IT24103878"
> setwd("C:\\Users\\It24103878\\Desktop\\IT24103878")
> getwd()
[1] "C:/Users/It24103878/Desktop/IT24103878"
> Delivery_Times <- read.table('Exercise - Lab 05.txt', header = TRUE)
>
```

```r
#2
cat("Dataset structure:\n")
str(Delivery_Times)
cat("\nFirst few rows:\n")
head(Delivery_Times)
breaks <- seq(20, 70, length.out = 10)

```

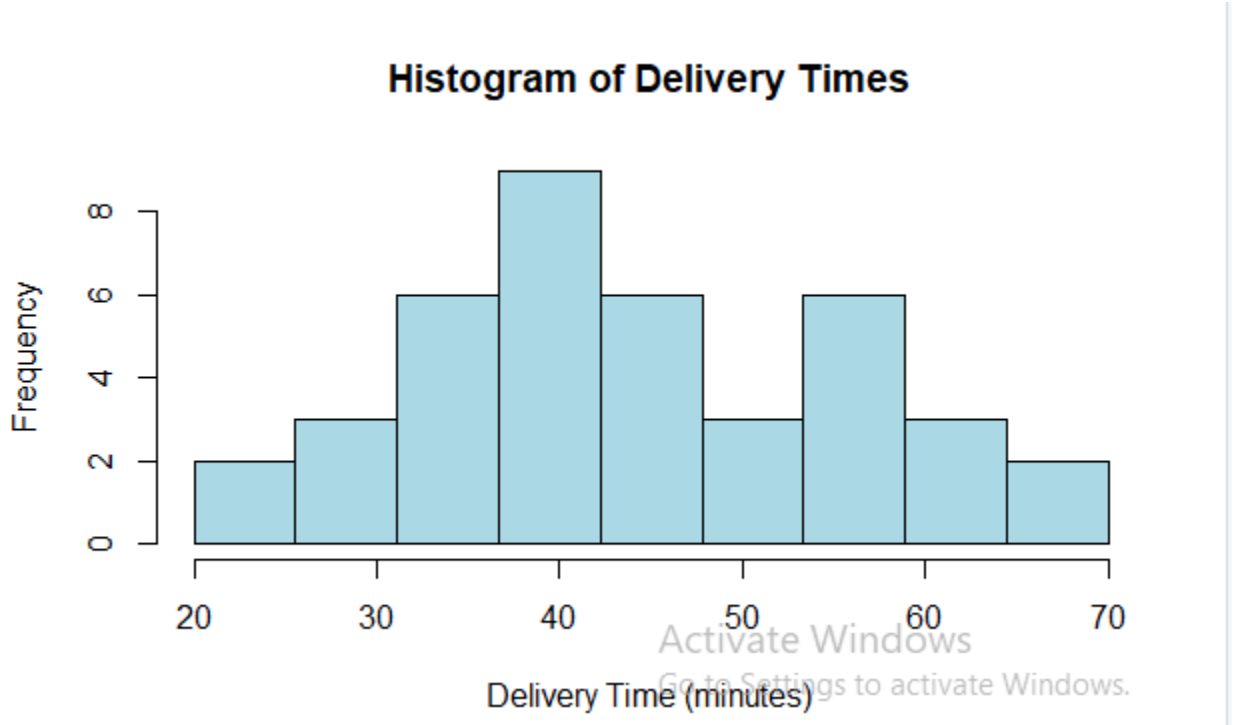
```
> #2
> cat("Dataset structure:\n")
Dataset structure:
> str(Delivery_Times)
'data.frame':   40 obs. of  1 variable:
 $ Delivery_Time_.minutes.: int  34 54 47 29 39 61 20 40 57 36 ...
> cat("\nFirst few rows:\n")

First few rows:
> head(Delivery_Times)
  Delivery_Time_.minutes.
1                      34
2                      54
3                      47
4                      29
5                      39
6                      61
> breaks <- seq(20, 70, length.out = 10)
> 
```

```
14  cat("\nClass intervals (right open):\n")
15  intervals <- paste0("(", head(breaks, -1), ", ", tail(breaks, -1), "]")
16  print(intervals)
```

```
> breaks <- seq(20, 70, length.out = 10)
> cat("\nClass intervals (right open):\n")

Class intervals (right open):
> intervals <- paste0("(", head(breaks, -1), ", ", tail(breaks, -1), "]")
> print(intervals)
[1] "(20, 25.5555555555556]"               "(25.5555555555556, 31.1111111111111]"
[3] "(31.1111111111111, 36.6666666666667]" "(36.6666666666667, 42.2222222222222]"
[5] "(42.2222222222222, 47.7777777777778]" "(47.7777777777778, 53.3333333333333]"
[7] "(53.3333333333333, 58.8888888888889]" "(58.8888888888889, 64.4444444444444]"
[9] "(64.4444444444444, 70]"
> 
```

```
hist(Delivery_Times$Delivery_Time_.minutes.,
     breaks = breaks,
     right = TRUE,
     main = "Histogram of Delivery Times",
     xlab = "Delivery Time (minutes)",
     ylab = "Frequency",
     col = "lightblue",
     border = "black",
     xlim = c(20, 70))
```

```
Class intervals (right open):
> intervals <- paste0("(", head(breaks, -1), ", ", tail(breaks, -1), "]")
> print(intervals)
[1] "(20, 25.5555555555556]"                "(25.5555555555556, 31.1111111111111]"
[3] "(31.1111111111111, 36.6666666666667]" "(36.6666666666667, 42.2222222222222]"
[5] "(42.2222222222222, 47.7777777777778]" "(47.7777777777778, 53.3333333333333]"
[7] "(53.3333333333333, 58.8888888888889]" "(58.8888888888889, 64.4444444444444]"
[9] "(64.4444444444444, 70]"
> hist(Delivery_Times$Delivery_Time_.minutes.,
+      breaks = breaks,
+      right = TRUE,
+      main = "Histogram of Delivery Times",
+      xlab = "Delivery Time (minutes)",
+      ylab = "Frequency",
+      col = "lightblue",
+      border = "black",
+      xlim = c(20, 70))
>
```



```
cat("\n3. Shape of the distribution:\n")
dist_shape <- "The distribution appears to be approximately symmetric with a sligh
dist_shape <- paste0(dist_shape, "Most delivery times are concentrated between 35-
dist_shape <- paste0(dist_shape, "There are fewer deliveries at the extremes (very
cat(dist_shape, "\n")
```

```
3. Shape of the distribution:
> dist_shape <- "The distribution appears to be approximately symmetric with a slight r
ight skew. "
> dist_shape <- paste0(dist_shape, "Most delivery times are concentrated between 35-55
minutes. ")
> dist_shape <- paste0(dist_shape, "There are fewer deliveries at the extremes (very fa
st or very slow delivery times).")
> cat(dist_shape, "\n")
The distribution appears to be approximately symmetric with a slight right skew. Most d
elivery times are concentrated between 35-55 minutes. There are fewer deliveries at the
extremes (very fast or very slow delivery times).
>
```

```
freq <- hist(Delivery_Times$Delivery_Time_.minutes., breaks = breaks, plot = FALSE
cum_freq <- cumsum(freq)
```

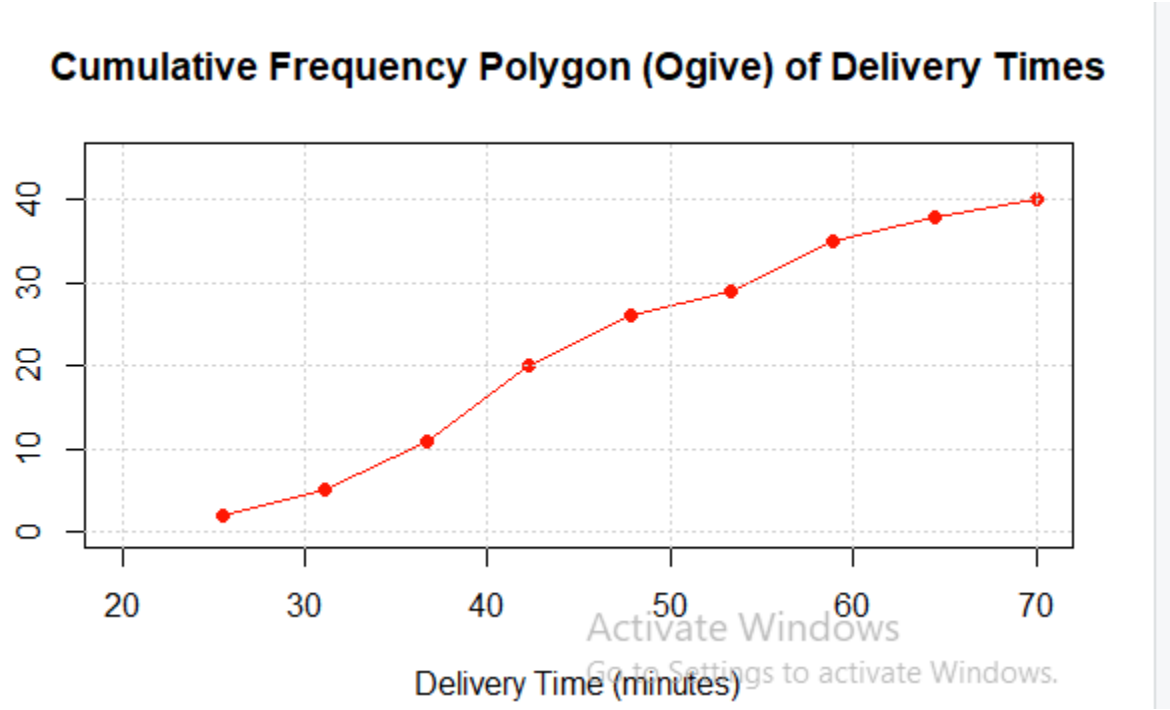
```
> freq <- hist(Delivery_Times$Delivery_Time_.minutes., breaks = breaks, plot = FALSE)$c
ounts
> cum_freq <- cumsum(freq)
>
```

```
cat("\nFrequency distribution:\n")
freq_table <- data.frame(Interval = intervals, Frequency = freq, Cumulative = cum_
print(freq_table)
```

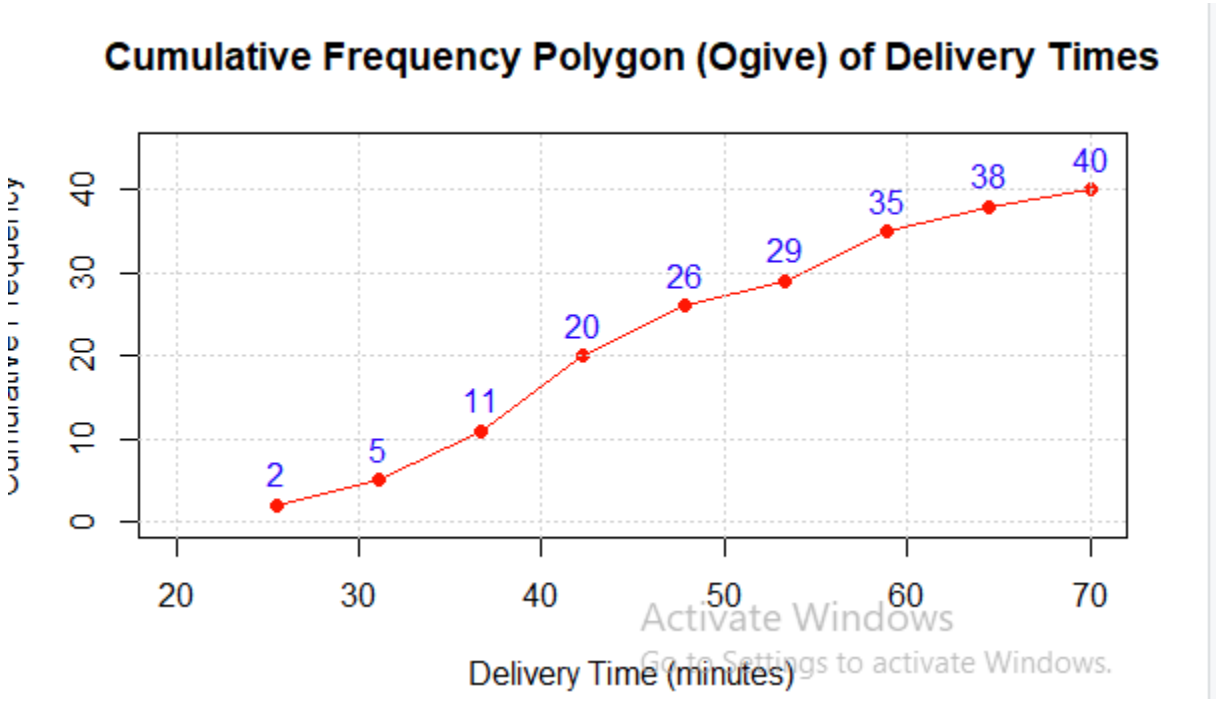
```
> print(freq_table)
                          Interval Frequency Cumulative
1               (20, 25.5555555555556]         2          2
2 (25.5555555555556, 31.1111111111111]         3          5
3 (31.1111111111111, 36.6666666666667]         6         11
4 (36.6666666666667, 42.2222222222222]         9         20
5 (42.2222222222222, 47.7777777777778]         6         26
6 (47.7777777777778, 53.3333333333333]         3         29
7 (53.3333333333333, 58.8888888888889]         6         35
8 (58.8888888888889, 64.4444444444444]         3         38
9               (64.4444444444444, 70]         2         40
>
```

```
plot(breaks[-1], cum_freq,
     type = "o",
     pch = 16,
     col = "red",
     main = "Cumulative Frequency Polygon (Ogive) of Delivery Times",
     xlab = "Delivery Time (minutes)",
     ylab = "Cumulative Frequency",
     xlim = c(20, 70),
     ylim = c(0, max(cum_freq) + 5))

grid()
```

```
9              (64.444444444444, 70]          2          40
> plot(breaks[-1], cum_freq,
+      type = "o",
+      pch = 16,
+      col = "red",
+      main = "Cumulative Frequency Polygon (Ogive) of Delivery Times",
+      xlab = "Delivery Time (minutes)",
+      ylab = "Cumulative Frequency",
+      xlim = c(20, 70),
+      ylim = c(0, max(cum_freq) + 5))
>
> grid()
>
```



```
text(breaks[-1], cum_freq, labels = cum_freq, pos = 3, col = "blue")

png("delivery_times_histogram.png", width = 800, height = 600)
```



```r
png("delivery_times_histogram.png", width = 800, height = 600)
hist(Delivery_Times$Delivery_Time_.minutes.,
     breaks = breaks,
     right = TRUE,
     main = "Histogram of Delivery Times",
     xlab = "Delivery Time (minutes)",
     ylab = "Frequency",
     col = "lightblue",
     border = "black",
     xlim = c(20, 70))
dev.off()

png("delivery_times_ogive.png", width = 800, height = 600)
```

(Top Level) R

Terminal × Background Jobs ×

R 4.2.2 · C:/Users/It24103878/Desktop/IT24103878/

```r
     ylim = c(0, max(cum_freq) + 5))

grid()
text(breaks[-1], cum_freq, labels = cum_freq, pos = 3, col = "blue")
png("delivery_times_histogram.png", width = 800, height = 600)
hist(Delivery_Times$Delivery_Time_.minutes.,
     breaks = breaks,
     right = TRUE,
     main = "Histogram of Delivery Times",
     xlab = "Delivery Time (minutes)",
     ylab = "Frequency",
     col = "lightblue",
     border = "black",
     xlim = c(20, 70))
dev.off()
RStudioGD
        2
```

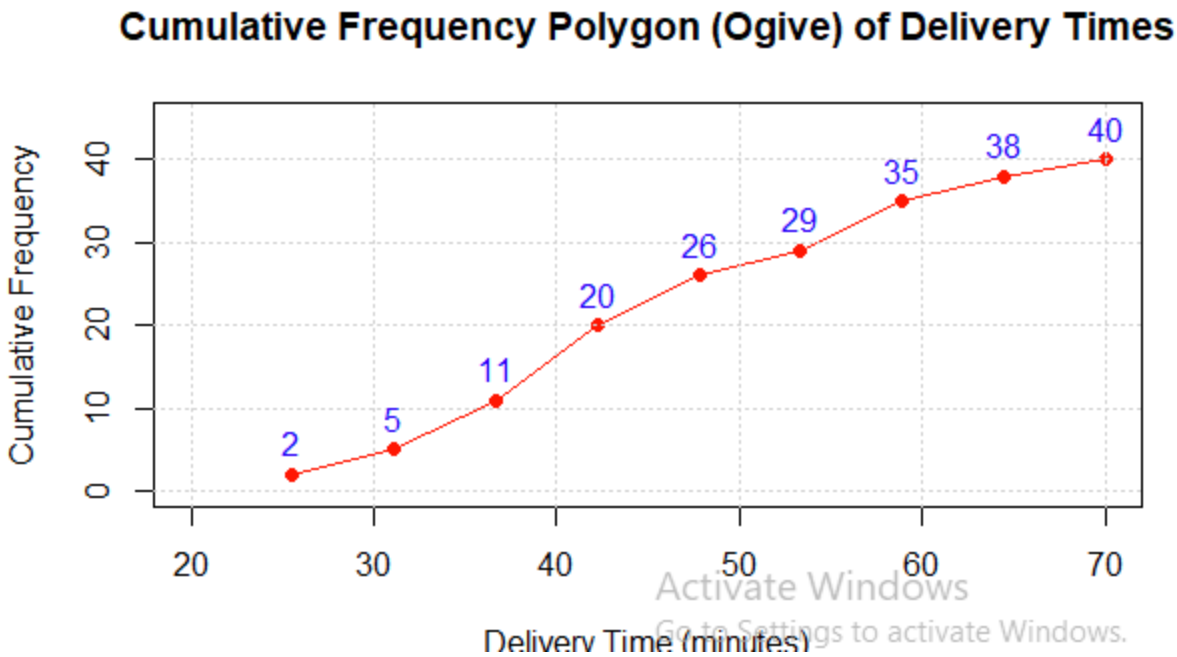
```
     xlab = "Delivery Time (minutes)",
     ylab = "Cumulative Frequency",
     xlim = c(20, 70),
     ylim = c(0, max(cum_freq) + 5))
grid()
text(breaks[-1], cum_freq, labels = cum_freq, pos = 3, col = "blue")
dev.off()

cat("\nAnalysis completed successfully!\n")
cat("Plots have been saved as 'delivery_times_histogram.png' and 'delivery_times
```

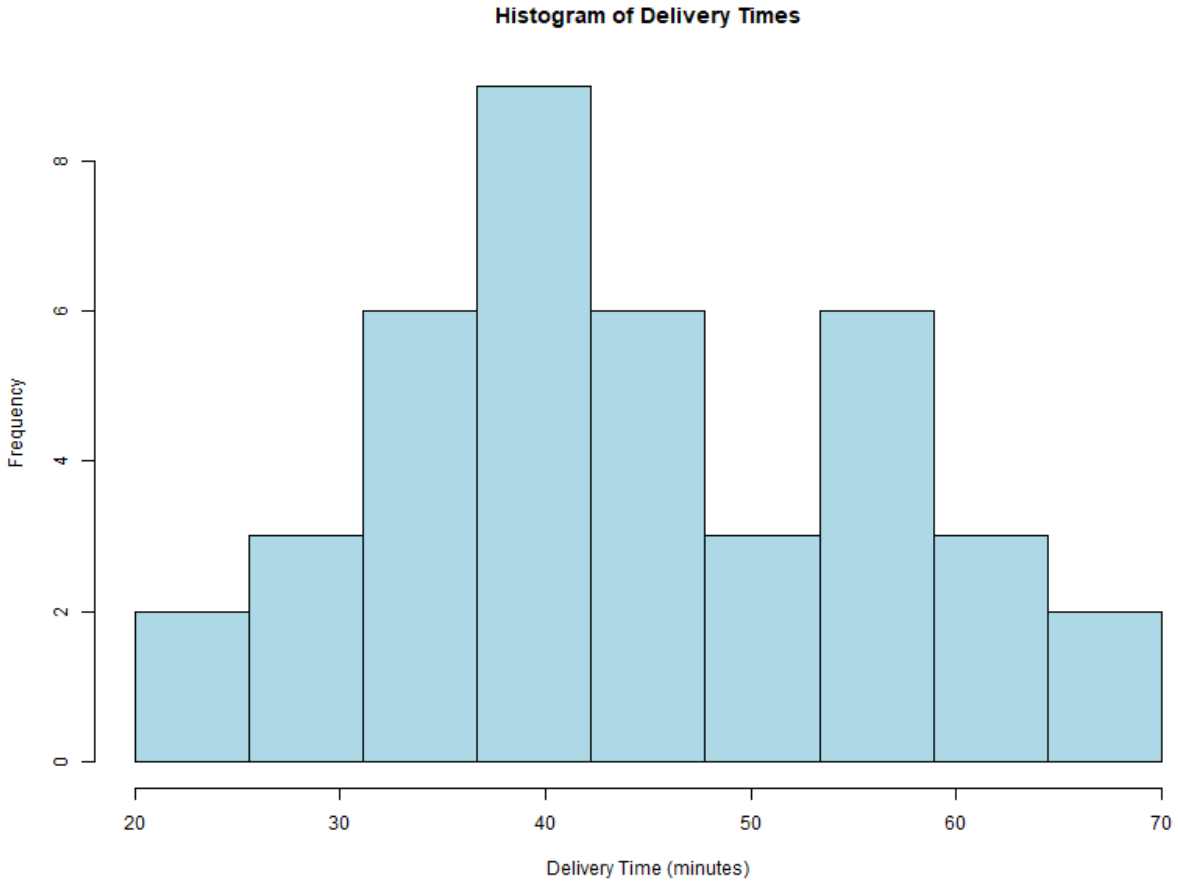
(Top Level) R Scr

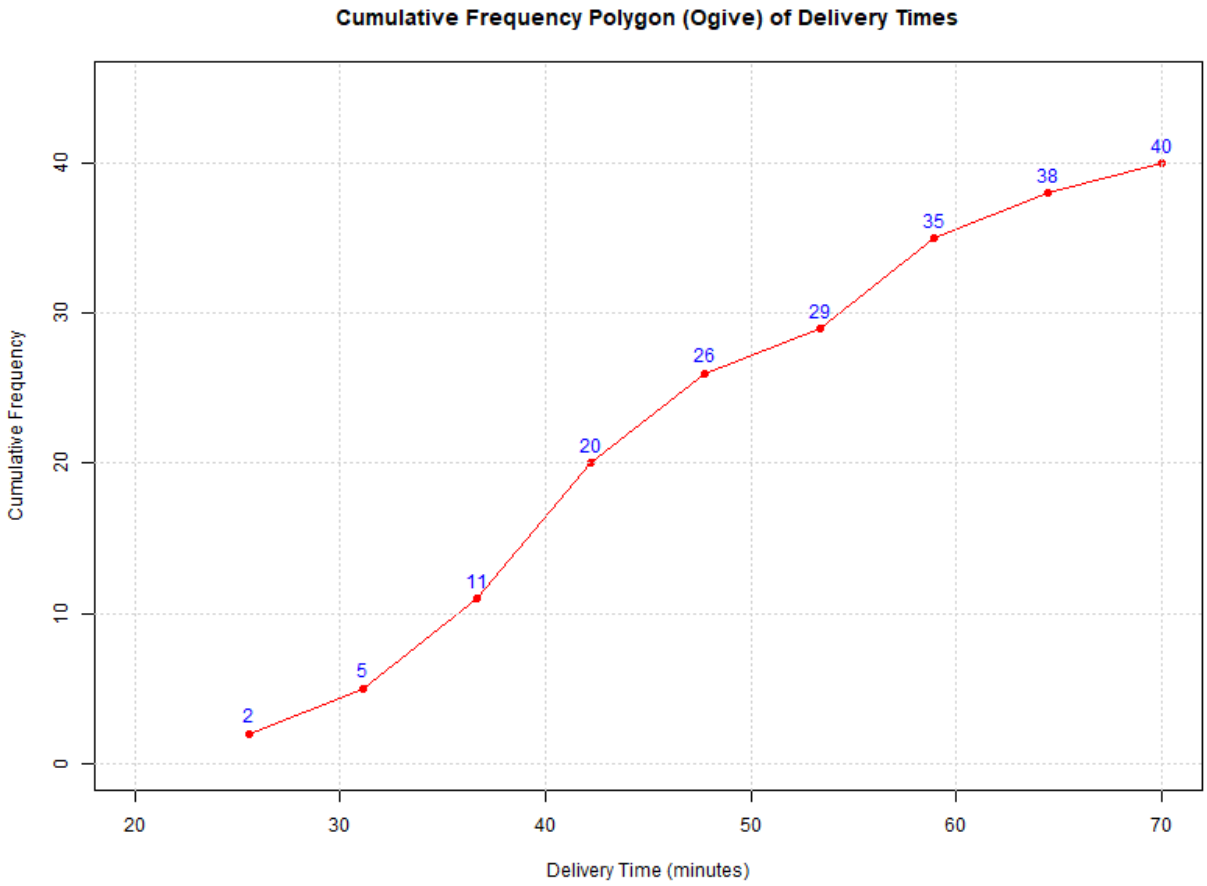
ole | Terminal × | Background Jobs ×

R 4.2.2 · C:/Users/It24103878/Desktop/IT24103878/

```
    main = "Cumulative Frequency Polygon (Ogive) of Delivery Times",
    xlab = "Delivery Time (minutes)",
    ylab = "Cumulative Frequency",
    xlim = c(20, 70),
    ylim = c(0, max(cum_freq) + 5))
id()
xt(breaks[-1], cum_freq, labels = cum_freq, pos = 3, col = "blue")
v.off()
dioGD
    2

t("\nAnalysis completed successfully!\n")

ysis completed successfully!
t("Plots have been saved as 'delivery_times_histogram.png' and 'delivery_times_ogi
g'\n")
s have been saved as 'delivery_times_histogram.png' and 'delivery_times_ogive.png'
```




Histogram of Delivery Times
Frequency
0
2
4
6
8
20
30
40
50
60
70
Delivery Time (minutes)


Cumulative Frequency Polygon (Ogive) of Delivery Times
Cumulative Frequency
0
10
20
30
40
2
5
11
20
26
29
35
38
40
20
30
40
50
60
70
Delivery Time (minutes)