

Sri Lanka Institute of Information Technology



Lab Submission
Lab sheet No: 05

IT24101603

Palihawadana T. S.

Probability and Statistics | IT2120

B.Sc. (Hons) in Information Technology

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

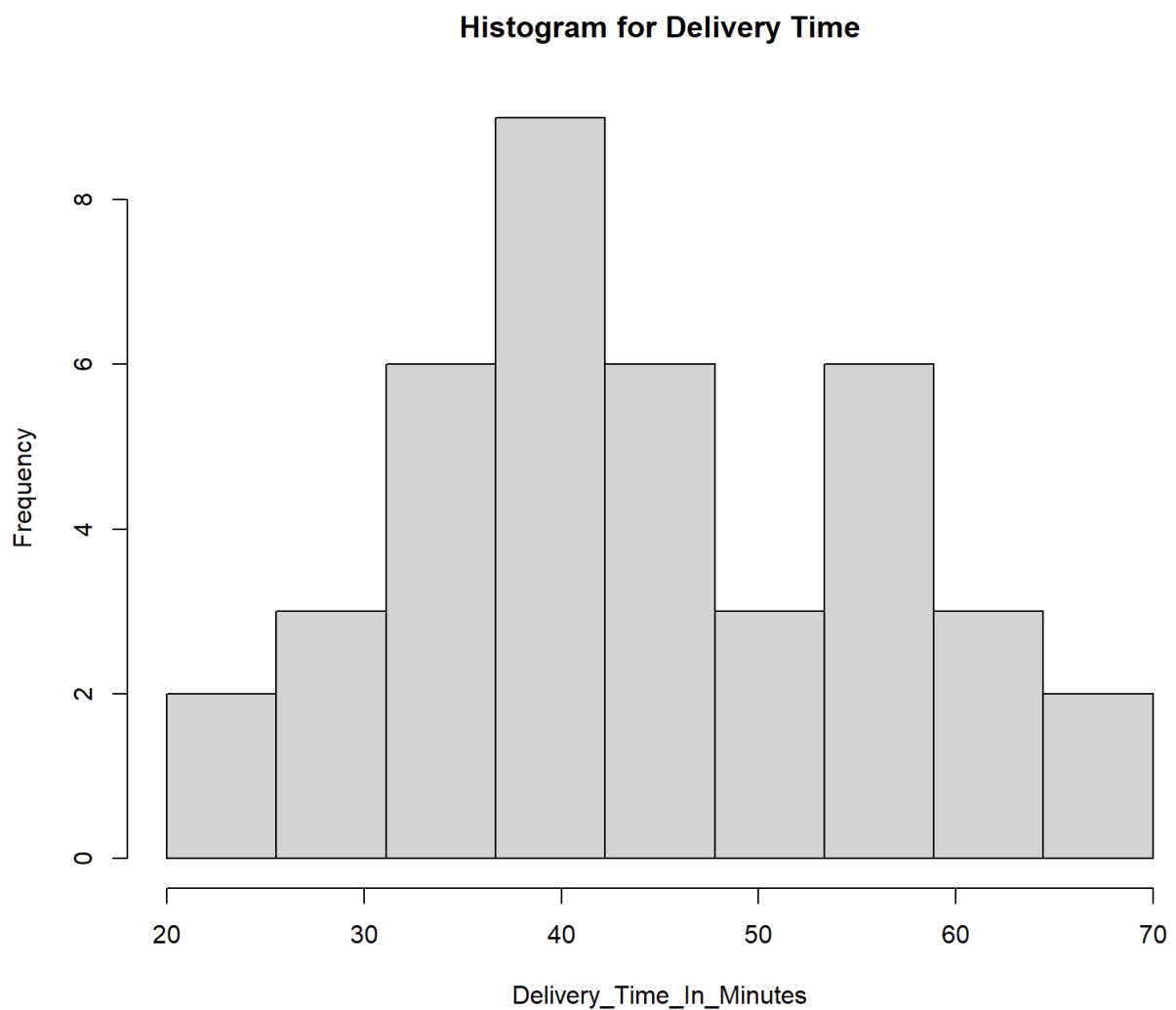
```
getwd()
setwd("C:\\Users\\teran\\OneDrive\\Desktop\\IT24101603")

Delivery_Times <- read.table("Exercise - Lab 05.txt", header=TRUE, sep=",")
fix(Delivery_Times)
attach(Delivery_Times)

names(Delivery_Times) <- c("Delivery_Time_In_Minutes")
attach(Delivery_Times)
```

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_In_Minutes,
  main = "Histogram for Delivery Time",
  breaks = seq(20, 70, length = 10),
  right = FALSE)
```



3. Comment on the shape of the distribution.

- The distribution is roughly bell-shaped (unimodal).
- Most of the delivery times cluster around 40 minutes, which is the peak.
- The frequencies decrease fairly evenly as you move away from the center in both directions, so it is approximately symmetric.
- There's only a slight right tail (because a few deliveries stretched closer to 65–70 minutes), but it's not strongly skewed.

The distribution of delivery times is roughly symmetric and unimodal, centered around 40 minutes. It resembles a bell-shaped curve with a slight tendency toward right skewness.

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

```
##Assign class limits of the frequency distribution into a variable  
breaks <- round(histogram$breaks)
```

```
##Assign class frequencies of the histogram into a variable  
freq <- histogram$counts
```

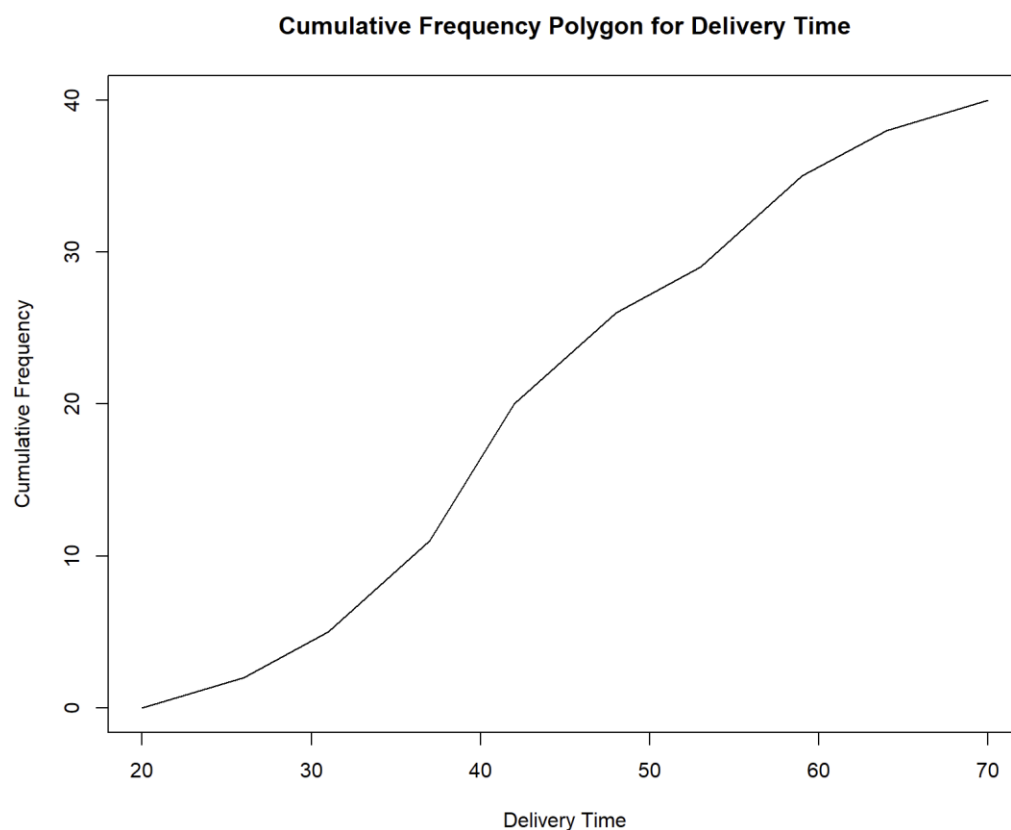
```
##Getting cumulative frequencies  
cum.freq <- cumsum(freq)
```

```
##Creating a null variable  
new <- c()
```

```
##Storing cumulative frequencies
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 for Delivery Time",
     xlab= "Delivery Time",
     ylab="Cumulative Frequency",
     ylim = c(0, max(cum.freq)))
```

```
cbind(Upper = breaks, CumFreq = new)
```



```
> 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