

# Assignment 1

2025/7/16

This is the report for the assignment 1:

Code used in the summary table:

```
mean_val <- mean(dataset$V1)
```

```
median_val <- median(dataset$V1)
```

```
sd_val <- sd(dataset$V1)
```

```
quantiles <- quantile(dataset$V1)
```

```
min_val <- min(dataset$V1)
```

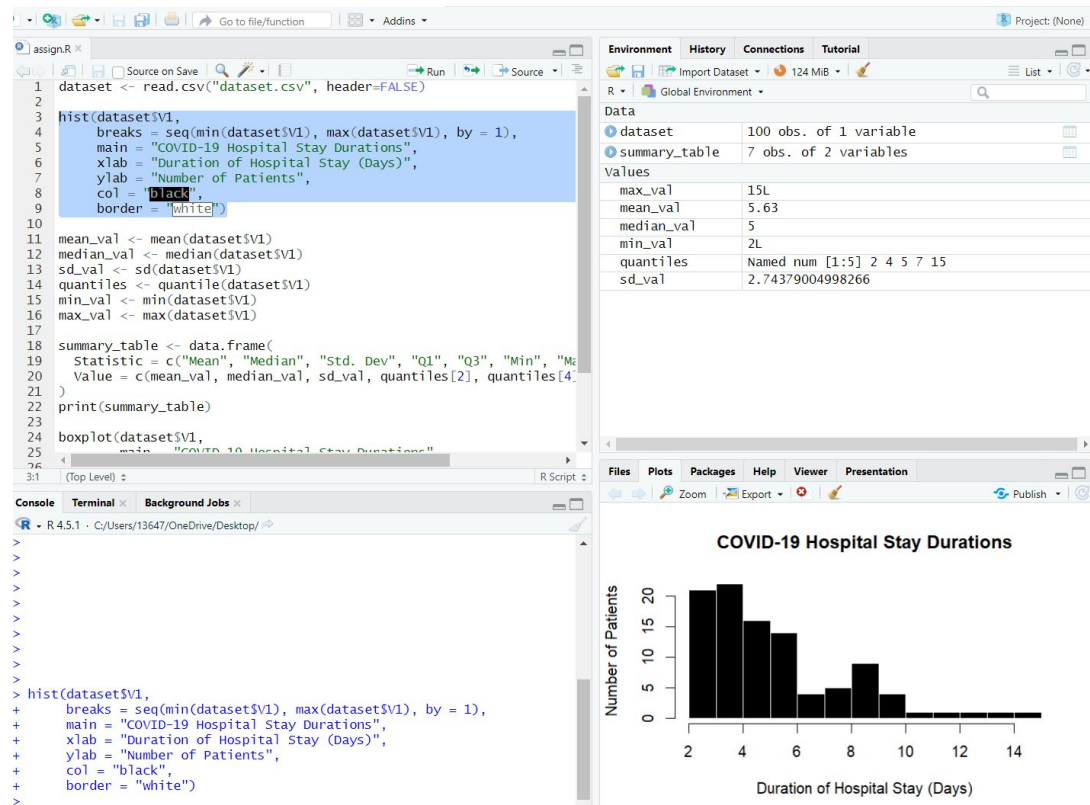
```
max_val <- max(dataset$V1)
```

## ***The Summary Table:***

Statistic	Value
Mean	5.32
Median	5.00
Standard Deviation	2.32
1st Quartile (Q1)	4.00
3rd Quartile (Q3)	6.00
Minimum	2.00
Maximum	15.00

I personally think that the Mean is the best way to describe the entire dataset, since it is reflecting the average the duration of hospital stays of patients admitted to the hospital with COVID-19.

1. Histogram:



This distribution implies that the majority of COVID-19 patients had relatively short hospital stays, but a small subset required extended care.

## 2. Boxplot

File Edit Code view Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

assign.R

```
7 ylab = "Number of Patients",
8 col = "black",
9 border = "white")
10
11 mean_val <- mean(dataset$V1)
12 median_val <- median(dataset$V1)
13 sd_val <- sd(dataset$V1)
14 quantiles <- quantile(dataset$V1)
15 min_val <- min(dataset$V1)
16 max_val <- max(dataset$V1)
17
18 summary_table <- data.frame(
19   Statistic = c("Mean", "Median", "Std. Dev", "Q1", "Q3", "Min", "Max"),
20   Value = c(mean_val, median_val, sd_val, quantiles[2], quantiles[4], min_val, max_val)
21 )
22 print(summary_table)
23
24 boxplot(dataset$V1,
25   main = "COVID-19 Hospital Stay Durations",
26   ylab = "Duration (Days)",
27   col = "darkgreen",
28   horizontal = TRUE)
29
30
```

Environment History Connections Tutorial

R Global Environment

Data

- dataset 100 obs. of 1 variable
- summary\_table 7 obs. of 2 variables

Values

max_val	15L
mean_val	5.63
median_val	5
min_val	2L
quantiles	Named num [1:5] 2 4 5 7 15
sd_val	2.74379004998266

Files Plots Packages Help Viewer Presentation

Zoom Export Publish

Console Terminal Background Jobs

R 4.5.1 - C:/Users/13647/OneDrive/Desktop/

```
>
>
>
> hist(dataset$V1,
+   breaks = seq(min(dataset$V1), max(dataset$V1), by = 1),
+   main = "COVID-19 Hospital Stay Durations",
+   xlab = "Duration of Hospital Stay (Days)",
+   ylab = "Number of Patients",
+   col = "black",
+   border = "white")
>
> boxplot(dataset$V1,
+   main = "COVID-19 Hospital Stay Durations",
+   ylab = "Duration (Days)",
+   col = "darkgreen",
+   horizontal = TRUE)
> |
```

COVID-19 Hospital Stay Durations

Duration (Days)

Statistic	Value
Mean	5.63
Median	5
Std. Dev	2.74379004998266
Q1	4
Q3	5
Min	2
Max	15