

**Muhammad Nadeem**

**201980050**

## **Data Visualization**

CP-1

Here is Code that meet our Class Participation Requirement.

- **Load the iris dataset using readr library and View its content.**

```
library(readr)
```

```
dataset <- read_csv("Dataset-CP/iris.csv")
```

```
View(dataset)
```

- **Explore ggplot2 and plot the following scatter plots:**

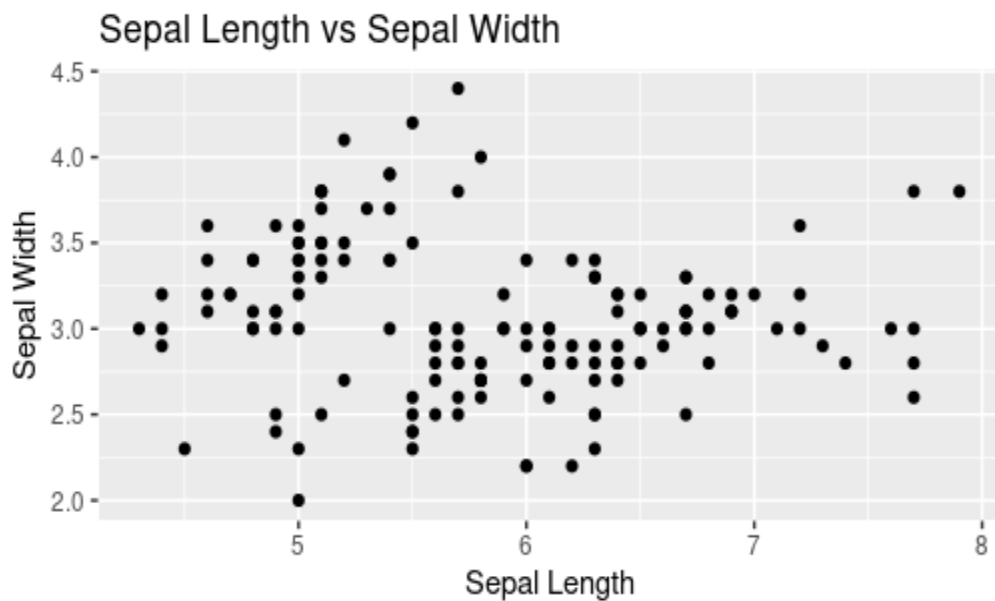
```
library(ggplot2)
```

```
ggplot(data = dataset, aes(x = sepal.length, y = sepal.width)) +
```

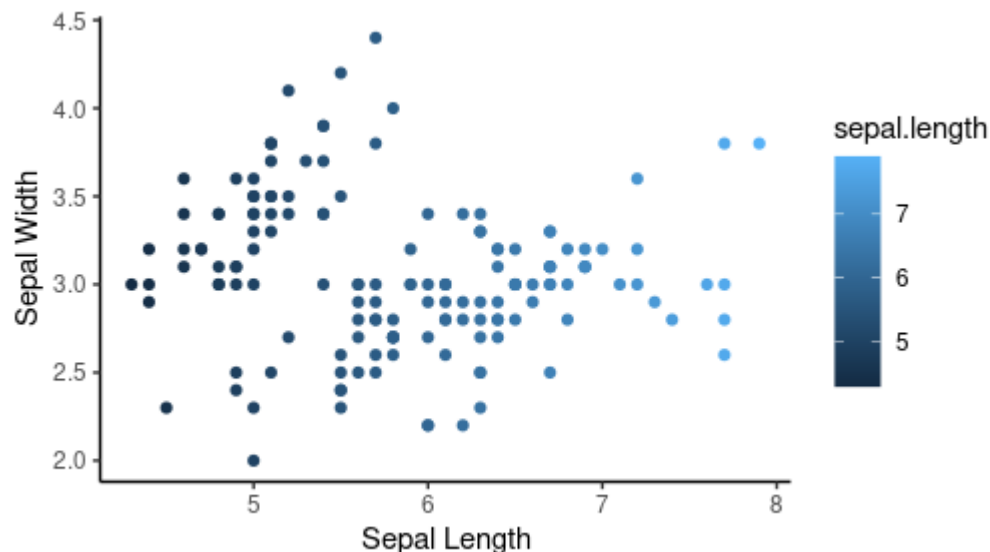
```
  geom_point() +
```

```
  labs(x = "Sepal Length", y = "Sepal Width") +
```

```
  ggtitle("Sepal Length vs Sepal Width")
```



```
ggplot(dataset, aes(x = sepal.length, y = sepal.width, color = sepal.length)) +
  geom_point() +
  labs(x = "Sepal Length", y = "Sepal Width") +
  theme_classic()
```



## ggplot2 in my own words.

This package is used for creating visualizations in R language to make more efficient visualization rather than Python and some other tools like Orange, Tableau, PowerBI, Excel etc. In Programming Language we have flexibility to make customized visualization in it.

ggplot2 is installed as `install.packages("ggplot2")` and imported as `library(ggplot2)`.

I suggest you, mostly before visualizing it. We need to do some Pre-Processing and sometimes we need to read files or perform many more operations before it.

We have a package that is named as tidyverse. First install it and then import it as `library(tidyverse)`. It gives multiple packages to do work efficiently and make code beautiful.

It gives below libraries to use...

```
> library(tidyverse)
— Attaching core tidyverse packages — tidyverse 2.0.0 —
✓ dplyr      1.1.4    ✓ readr      2.1.5
✓ forcats    1.0.0    ✓ stringr    1.5.1
✓ ggplot2    3.5.1    ✓ tibble     3.2.1
✓ lubridate  1.9.3    ✓ tidyr      1.3.1
✓ purrr      1.0.2
— Conflicts — tidyverse_conflicts() —
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()     masks stats::lag()
ℹ Use the conflicted package to force all conflicts to become errors
```