

# YPDC Session

Haseeb Raza

2024-07-08

## Importing iris data set

The Iris dataset is a multivariate dataset

The Iris dataset is a multivariate dataset introduced by the British biologist and statistician Ronald A. Fisher in his 1936 paper “The use of multiple measurements in taxonomic problems.” It is often used as a beginner’s dataset for machine learning and statistical classification techniques.

```
head(iris)
```

```
##      Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1           5.1         3.5         1.4         0.2  setosa
## 2           4.9         3.0         1.4         0.2  setosa
## 3           4.7         3.2         1.3         0.2  setosa
## 4           4.6         3.1         1.5         0.2  setosa
## 5           5.0         3.6         1.4         0.2  setosa
## 6           5.4         3.9         1.7         0.4  setosa
```

```
tail(iris)
```

```
##      Sepal.Length Sepal.Width Petal.Length Petal.Width  Species
## 145           6.7         3.3         5.7         2.5 virginica
## 146           6.7         3.0         5.2         2.3 virginica
## 147           6.3         2.5         5.0         1.9 virginica
## 148           6.5         3.0         5.2         2.0 virginica
## 149           6.2         3.4         5.4         2.3 virginica
## 150           5.9         3.0         5.1         1.8 virginica
```

```
# Load the ggplot2 package
```

```
library(ggplot2)
```

```
# Create a scatter plot of Petal.Length vs Petal.Width, colored by Species
```

```
ggplot(data = iris, aes(x = Petal.Length, y = Petal.Width, color = Species))+
```

```
  geom_point(size = 3, alpha = 0.7) + # Add points with custom size and transparency
```

```
  geom_smooth(method = "lm", se = FALSE, linetype = "dashed") + # Add linear model lines
```

```
  labs(
```

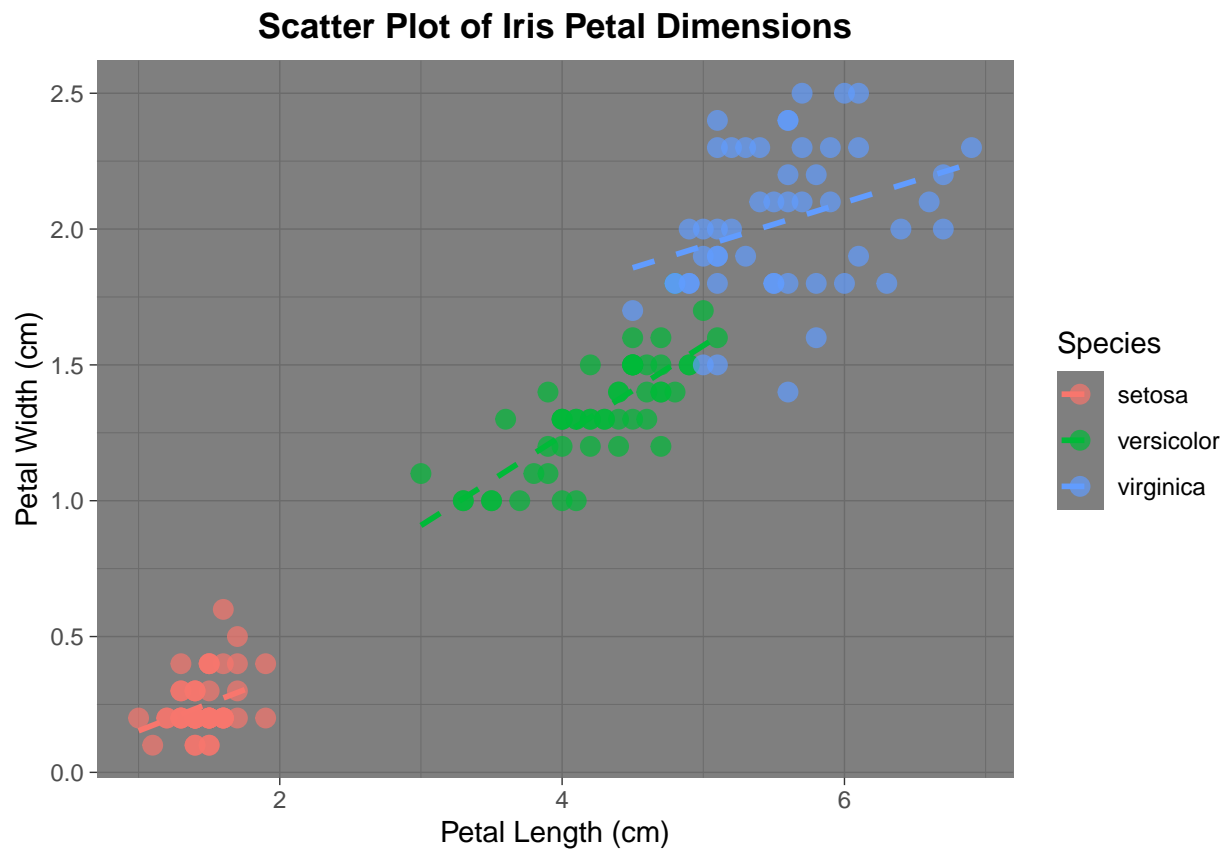
```
    title = "Scatter Plot of Iris Petal Dimensions",
```

```

x = "Petal Length (cm)",
y = "Petal Width (cm)"
) +
theme_dark() + # Use a minimal theme for a clean look
theme(plot.title = element_text(hjust = 0.5, face = "bold"))

```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
# Center and bold the title
```