Homework 1- R

GurtejSingh 24216077

Introduction

The iris dataset is a classic example used in both statistics and machine learning. It contains measurements of sepal and petal sizes for 150 iris flowers, divided evenly across three species: setosa, versicolor, and virginica.

Dataset Overview

```
'data.frame':   150 obs. of  5 variables:
 $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

This dataset comes built into R and includes five variables:

- Sepal.Length
- Sepal.Width
- Petal.Length
- Petal.Width
- Species

There are 150 observations in total, with 50 flowers from each of the three species.

Update

```
 Sepal.Length     Sepal.Width      Petal.Length     Petal.Width
 Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
 1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
 Median :5.800   Median :3.000   Median :4.350   Median :1.300
 Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
 3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
 Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
       Species
 setosa    :50
 versicolor:50
 virginica :50
```
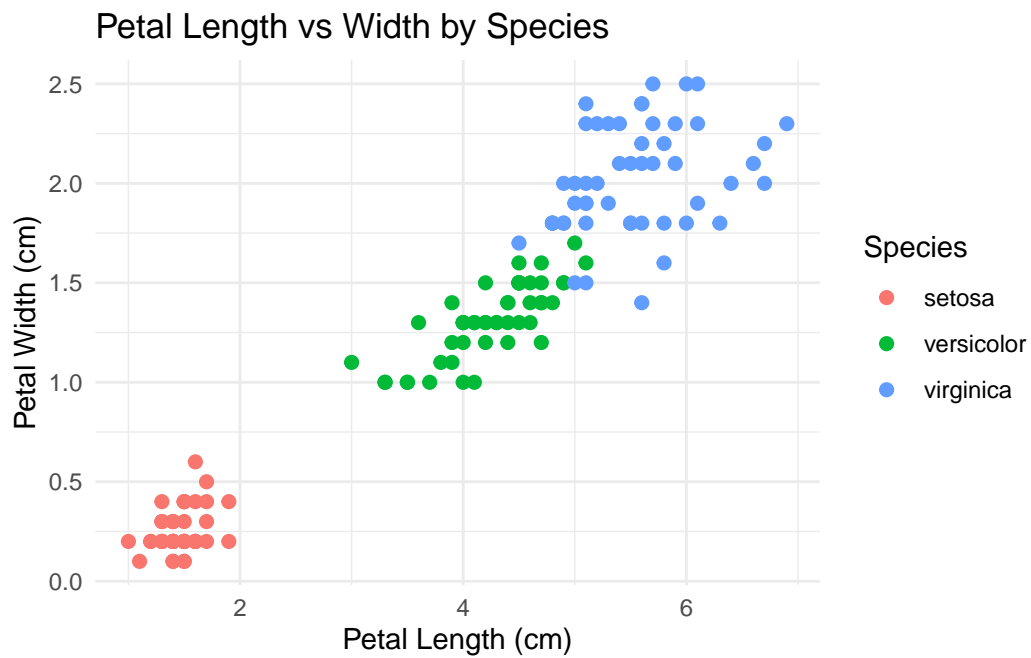
Observations

- Sepal.Length ranges from 4.3 to 7.9, with a mean of 5.84.

- Sepal.Width ranges from 2.0 to 4.4, with a mean of 3.06.

- Petal.Length ranges from 1.0 to 6.9, with a mean of 3.76.

- Petal.Width ranges from 0.1 to 2.5, with a mean of 1.20.

- Species has 3 categories: setosa, versicolor, and virginica, each with 50 observations.

Plot: Petal Dimensions by Species

Petal Length vs Width by Species

Conclusion

The scatter plot shows clear differences between the three iris species based on their petal size. Setosa flowers have the smallest petals, Virginica the largest, and Versicolor sits in the middle, with some overlap. This pattern suggests that petal length and width are useful traits for telling the species apart, both visually and in predictive models.