

Homework 1

Sarosh Farhan (24210969)

Introduction

The [Iris dataset](#) is a classic dataset in pattern recognition and multivariate statistics. It contains measurements of *sepal length*, *sepal width*, *petal length*, and *petal width* for 150 flowers from three species of iris: *setosa*, *versicolor*, and *virginica*.

Approach

I will explore the distribution of petal length versus petal width for each species using a scatter plot. This will help me visualize how distinct the species are based on these two features.

Statistics

The following table provides the count, mean, and standard deviation of **petal length** and **petal width** for each iris species.

Table 1: Summary Statistics by Species

Species	Count	MeanPetalLength	SDPetalLength	MeanPetalWidth	SDPetalWidth
setosa	50	1.46	0.17	0.25	0.11
versicolor	50	4.26	0.47	1.33	0.20
virginica	50	5.55	0.55	2.03	0.27

Summary table inferences

Setosa has the *smallest petals* overall, with an average petal length of **1.46 cm** and width of **0.25 cm**. The low standard deviations indicate that **Setosa** flowers are consistently small in size.

Versicolor shows *moderate petal* size, averaging **4.26 cm** in length and **1.33 cm** in width. It also exhibits slightly higher variation compared to **Setosa**.

Virginica has the *largest petals*, with mean petal dimensions of **5.55 cm** (length) and **2.03 cm** (width). It also displays the greatest variability, especially in petal width.

These differences are **substantial and consistent**, making **petal length and width excellent predictors** for classifying iris species.

Plot

The scatter plot below shows the relationship between petal length and petal width, colored by species:

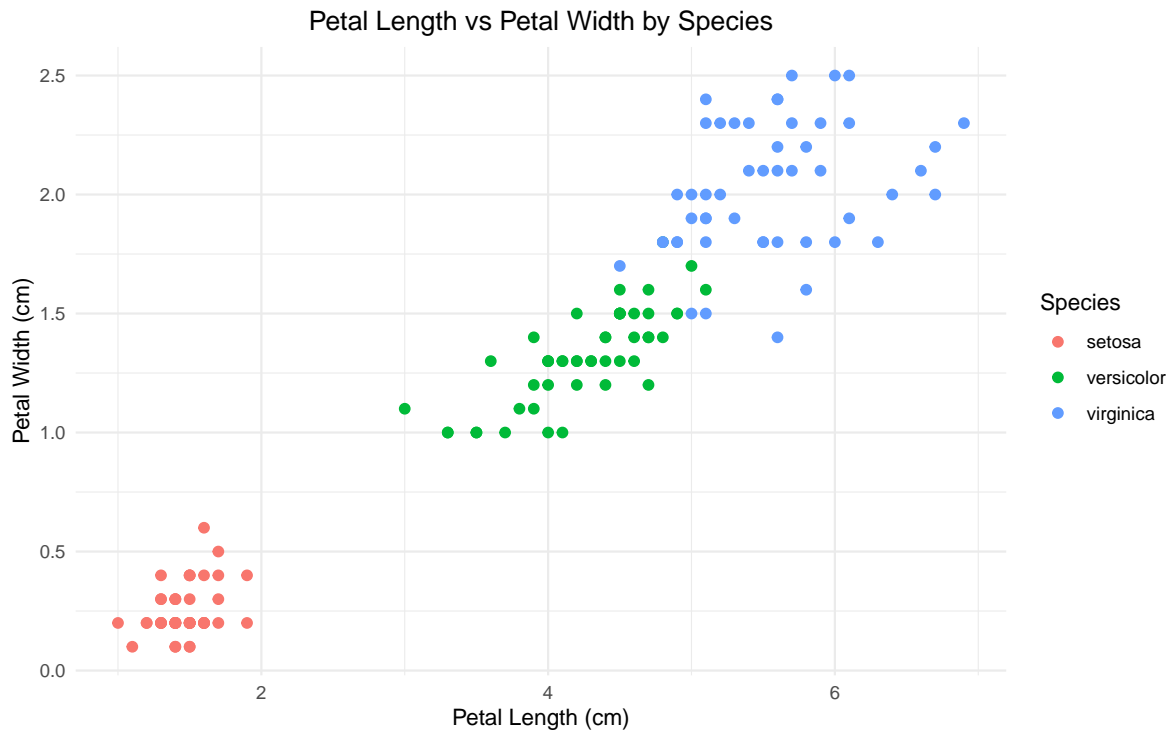


Figure 1: Scatter plot of Petal Length vs Petal Width by Species

Interpretation of the plot

From Figure 1 we can infer the following:

Setosa is clearly separable, its petals are significantly smaller in both length and width compared to other two species.

Versicolor lies in the middle range for both dimensions.

Virginica has the largest petals, both in length and width.

There is a visible positive correlation between petal length and petal width within each species. The species form distinct clusters, particularly **Setosa**, which is completely separated from the others. This makes petal measurements powerful features for classification tasks like species prediction.