Homework1_24206239

Jing Sun

1. Introduction

I'm using the *Iris* dataset for this homework1. It's a famous dataset which is built-in dataset in R with **150 observations** of flowers from three species: *setosa*, *versicolor*, and *virginica*. Each flower is described by **four features**: sepal length, sepal width, petal length, and petal width (in cm).

In the rest of this report, I look at the relationship between **petal length** and **petal width**, which helps to show how the species differ.

2. Method

I first explored the data briefly using `head()`, `summary()`, and `table()` to check the structure and basic statistics before making the plot:

Exploring the Data

```
  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
```

```
  Petal.Length    Petal.Width
 Min.   :1.000   Min.   :0.100
 1st Qu.:1.600   1st Qu.:0.300
 Median :4.350   Median :1.300
```

```
Mean   :3.758   Mean   :1.199
3rd Qu.:5.100   3rd Qu.:1.800
Max.   :6.900   Max.   :2.500
```

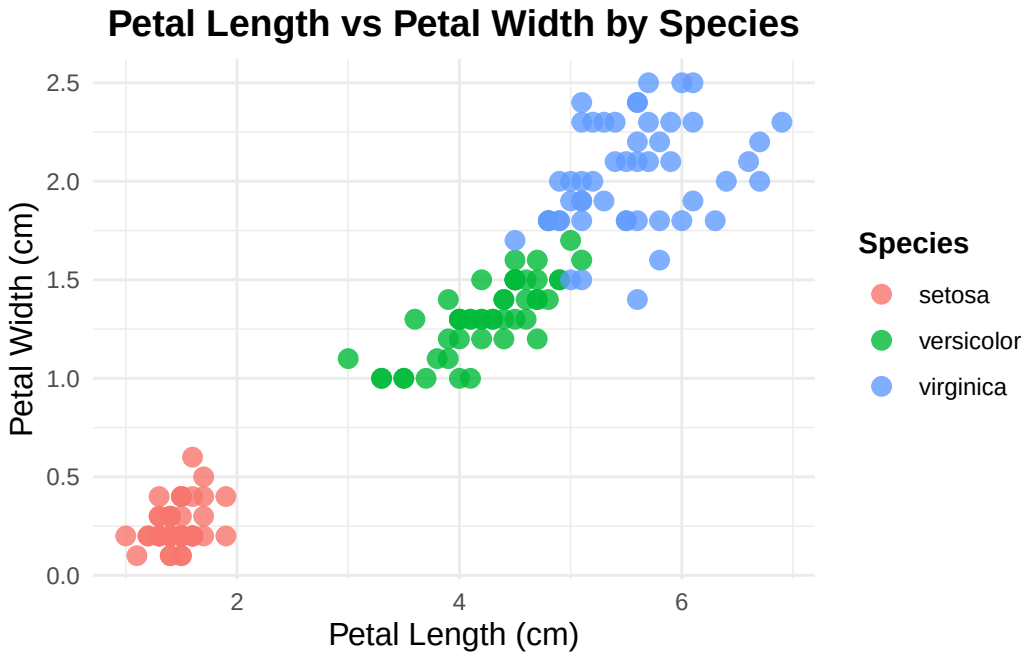
```
    setosa versicolor  virginica
        50         50         50
```

From the output above, we can obtain these information:

- The dataset has **150 rows** and **5 columns**.
- Each species appears exactly **50 times**.
- The average petal length is around **3.76 cm** and average petal width is about **1.20 cm**.
- *Setosa* tends to have much smaller petals than *virginica*, as seen clearly in the summary and scatter plot.

3. Scatter Plot: Petal Length vs Petal Width

I drew this scatter plot to show how petal length and width vary across species, since that's the main focus of this report:



Plot Description

The scatter plot shows a clear separation between the three iris species:

- *Setosa* flowers have short and narrow petals.
- *Versicolor* falls in the mid-range.
- *Virginica* flowers have the longest and widest petals.

This suggests that petal size is a strong indicator for classifying iris species.