Sheet 0

Pattern Recognition

Saeed Elsayed Ahmed Abokhatwa

6829

About:

This is perhaps the best known database to be found in the pattern recognition literature. Fisher's paper is a classic in the field and is referenced frequently to this day. The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other 2; the latter are NOT linearly separable from each other. Each class has 4 attributes which are Sepal length in cm, Sepal width in cm, Petal length in cm Petal width in cm. Classes are Iris Setosa, Iris Versicolour, Iris Virginica.

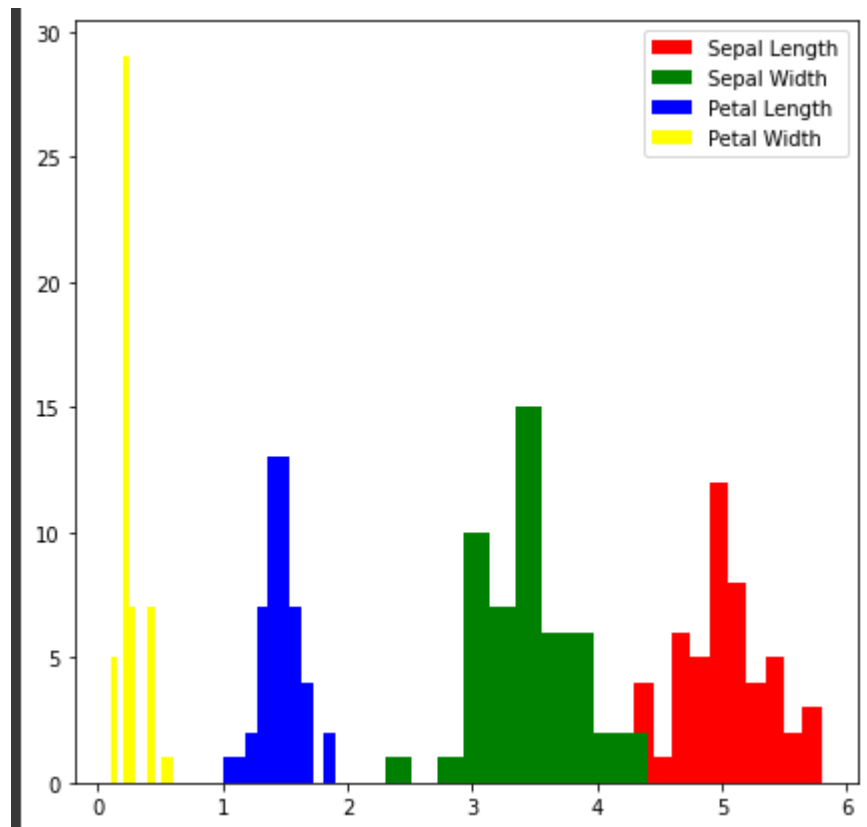| Number of classes	3
Number of Attributes	4
Number of Instances	150 (50 for each class)

Plots

Histogram

Setosa

Figure 1.1

As shown in Figure 1.1 so samples has wide width with almost 30 cm which might be outlier
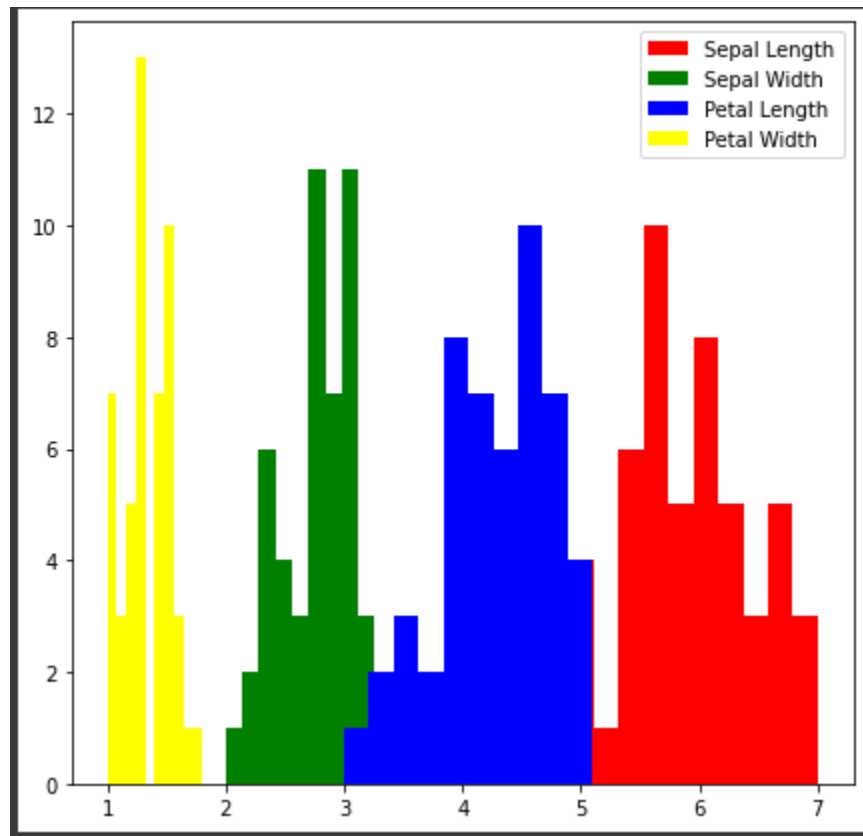
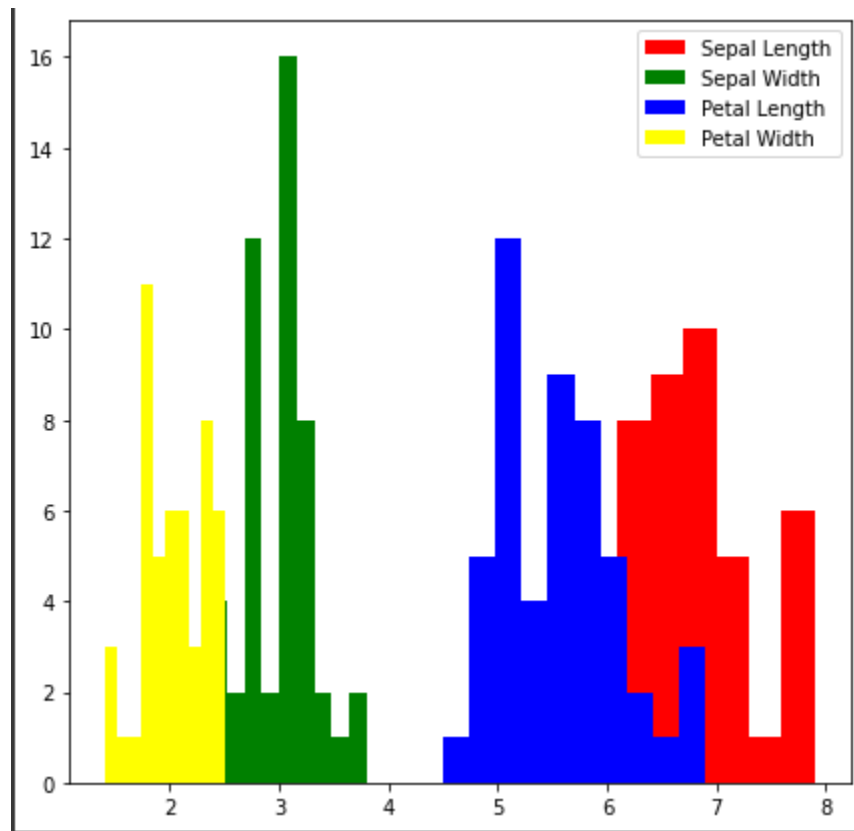# Versicolour

Figure 1.2

Virginica

Figure 1.3

2D Scatter

A 2D scatter is used to visualize the distribution of our samples and see if there is a relationship between two variables.
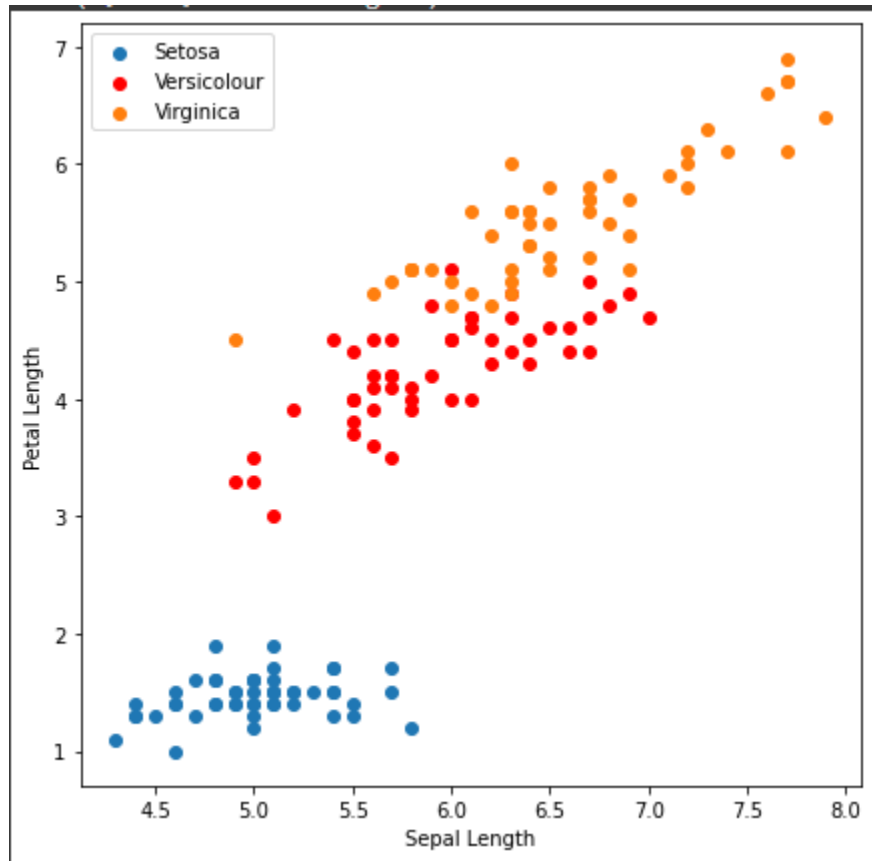
1. Sepal Length and Petal Length

Figure 2.1
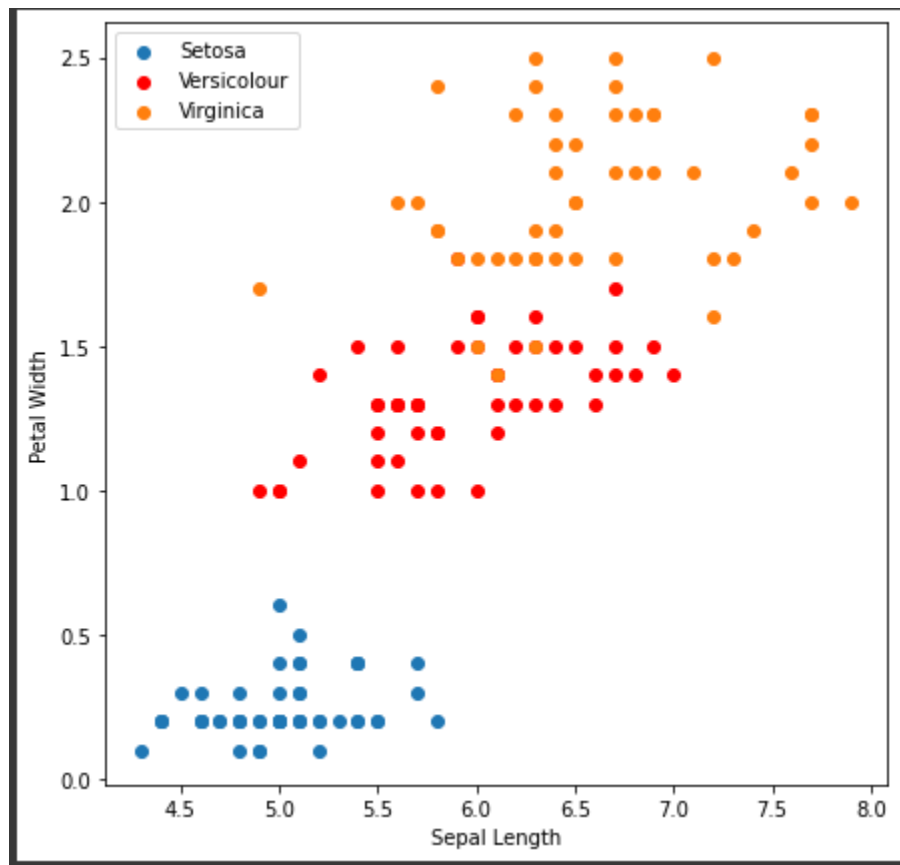
2. Sepal Length and Petal Width

Figure 2.2
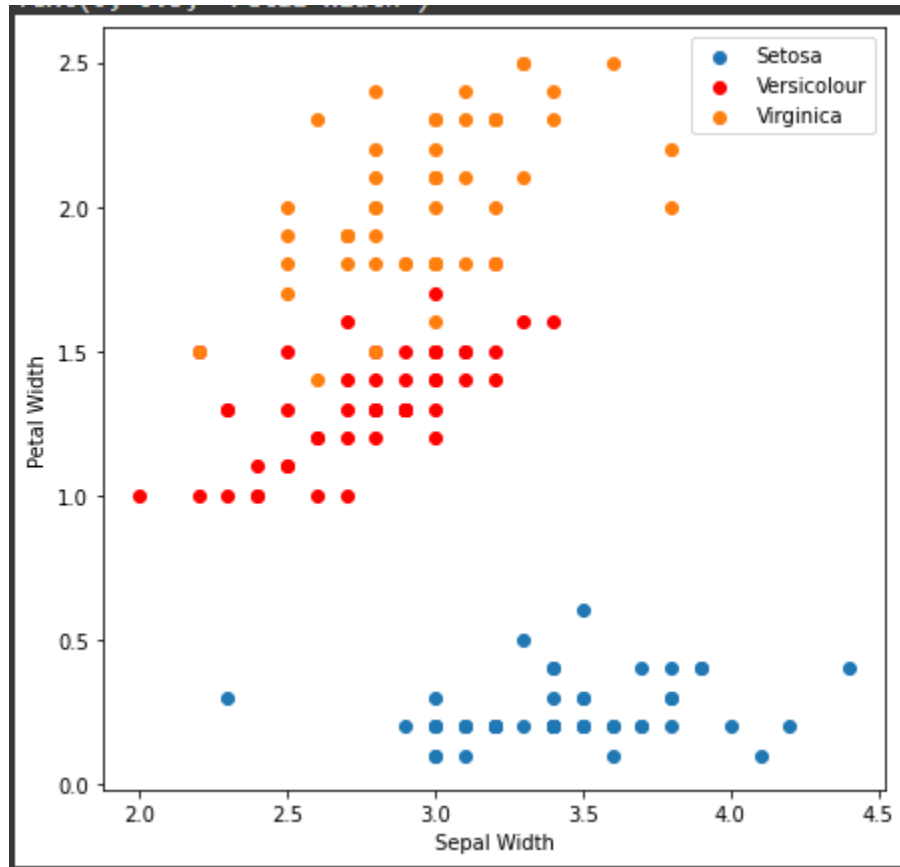
3. Sepal Width and Petal Width

Figure 2.3

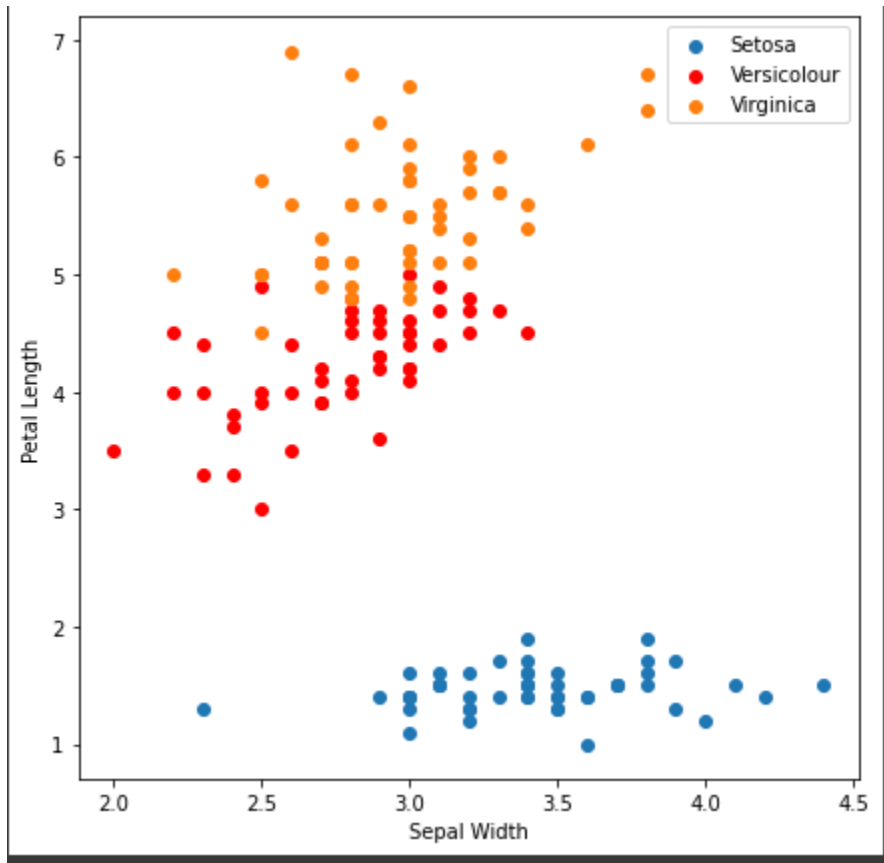## 4. Sepal Width and Petal Length

Figure 2.4

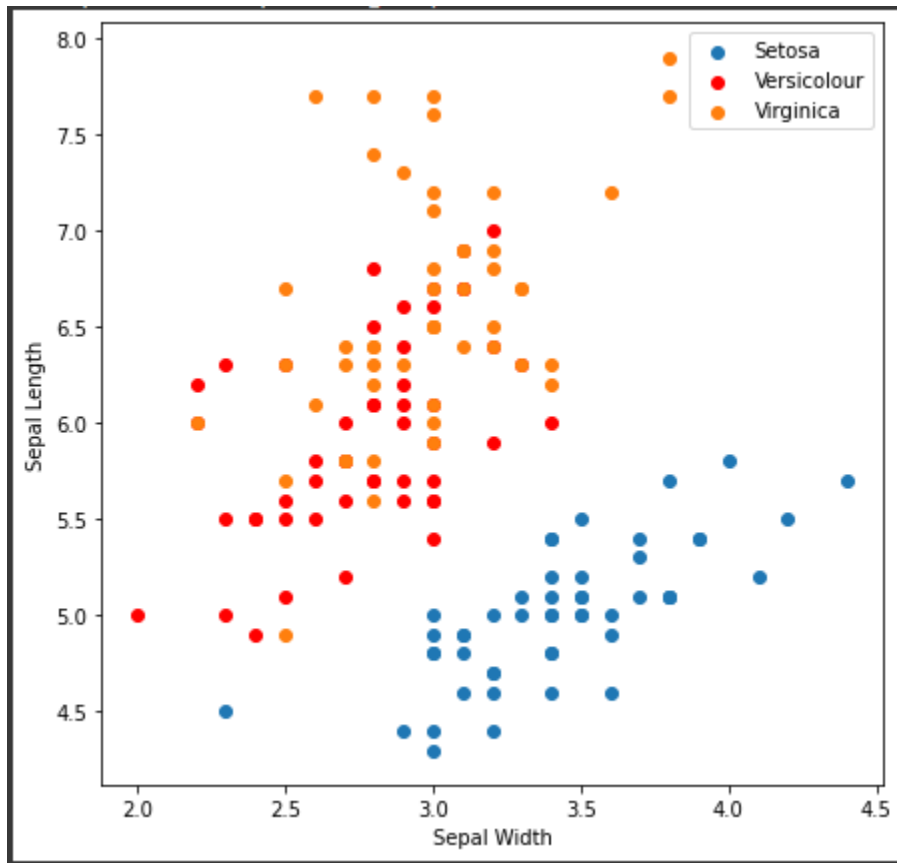## 5. Sepal Width and Sepal Length

Figure 2.5

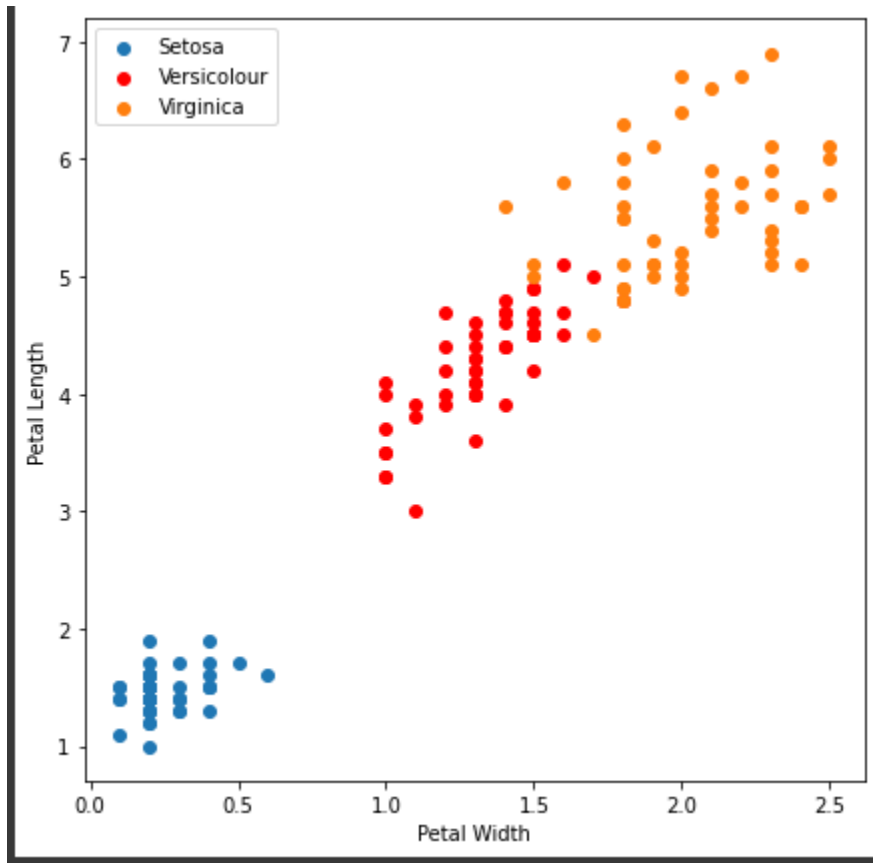6. Petal Width and Pental Length

Figure 2.6

We can clearly see that Setosa is linearly Separable from the others but Versicolour and Virginicia are not linearly Separable.

3D scatter plot
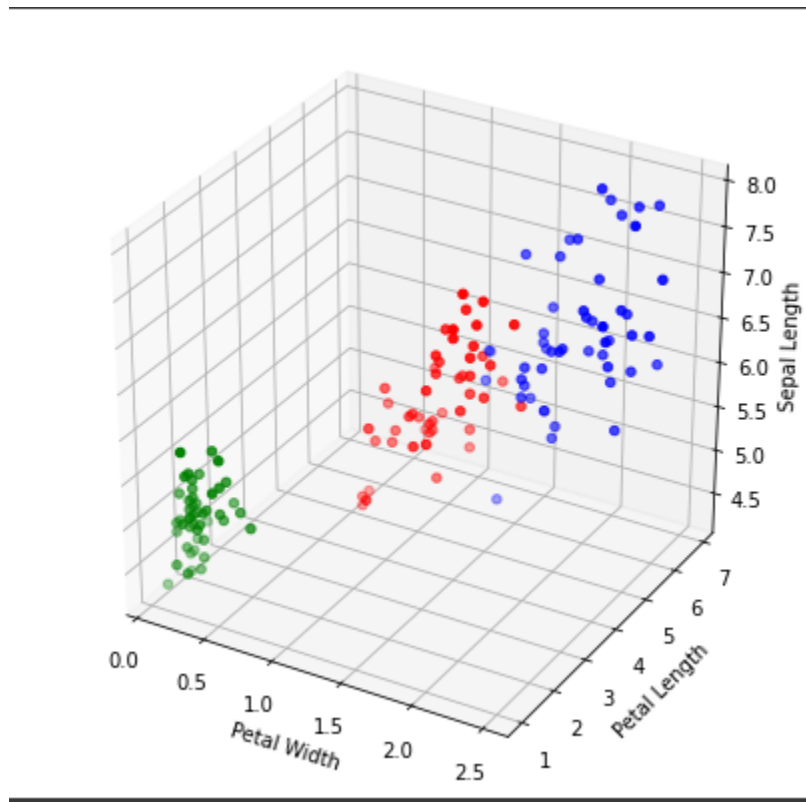# 1. Petal Length, Petal Width and Sepal Length

Figure 3.1

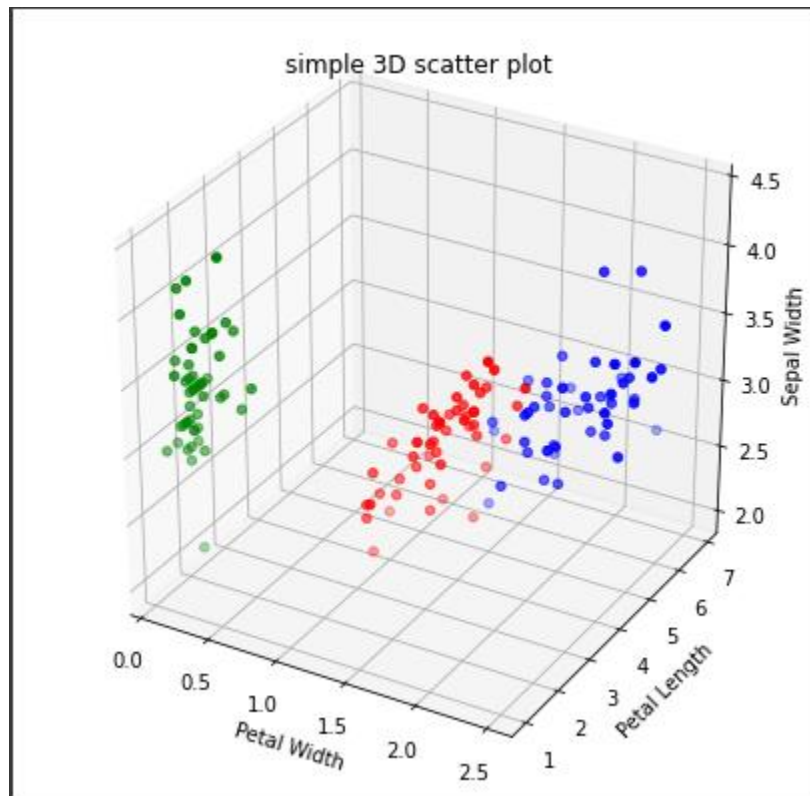## 2. Petal Length, Petal Width and Sepal Width

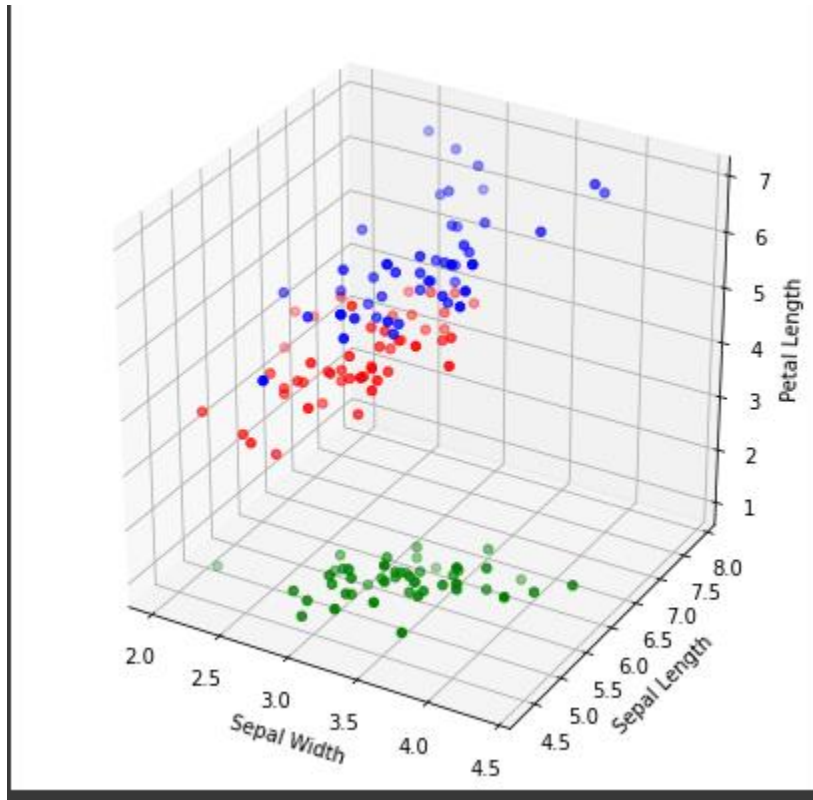Figure 3.2

3. Sepal Width, Sepal Width and Petal Length

Figure 3.3

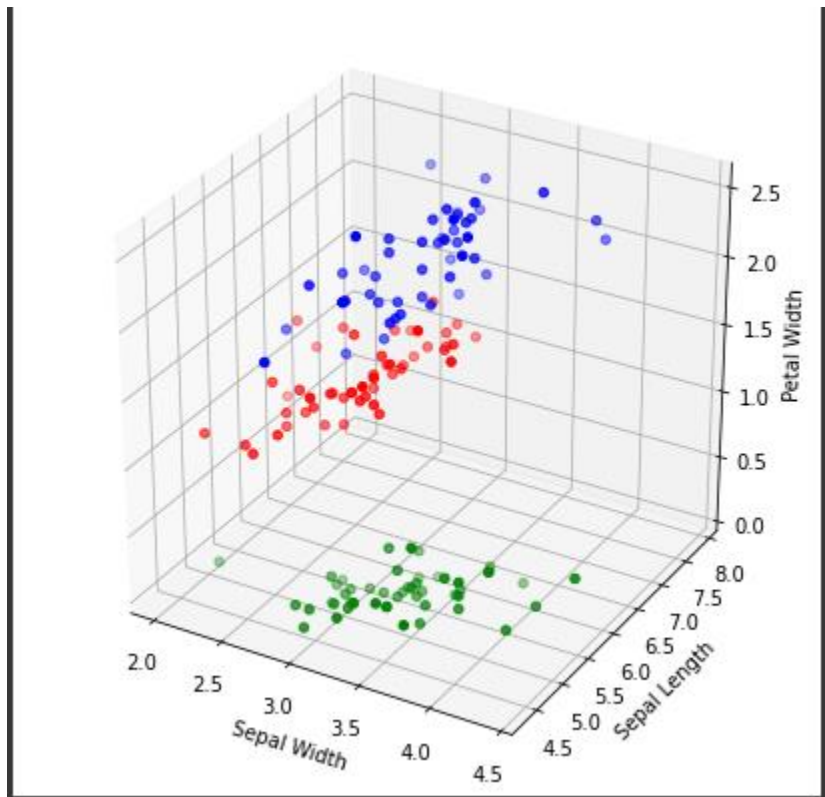## 4. Sepal Width, Sepal Width and Petal Width
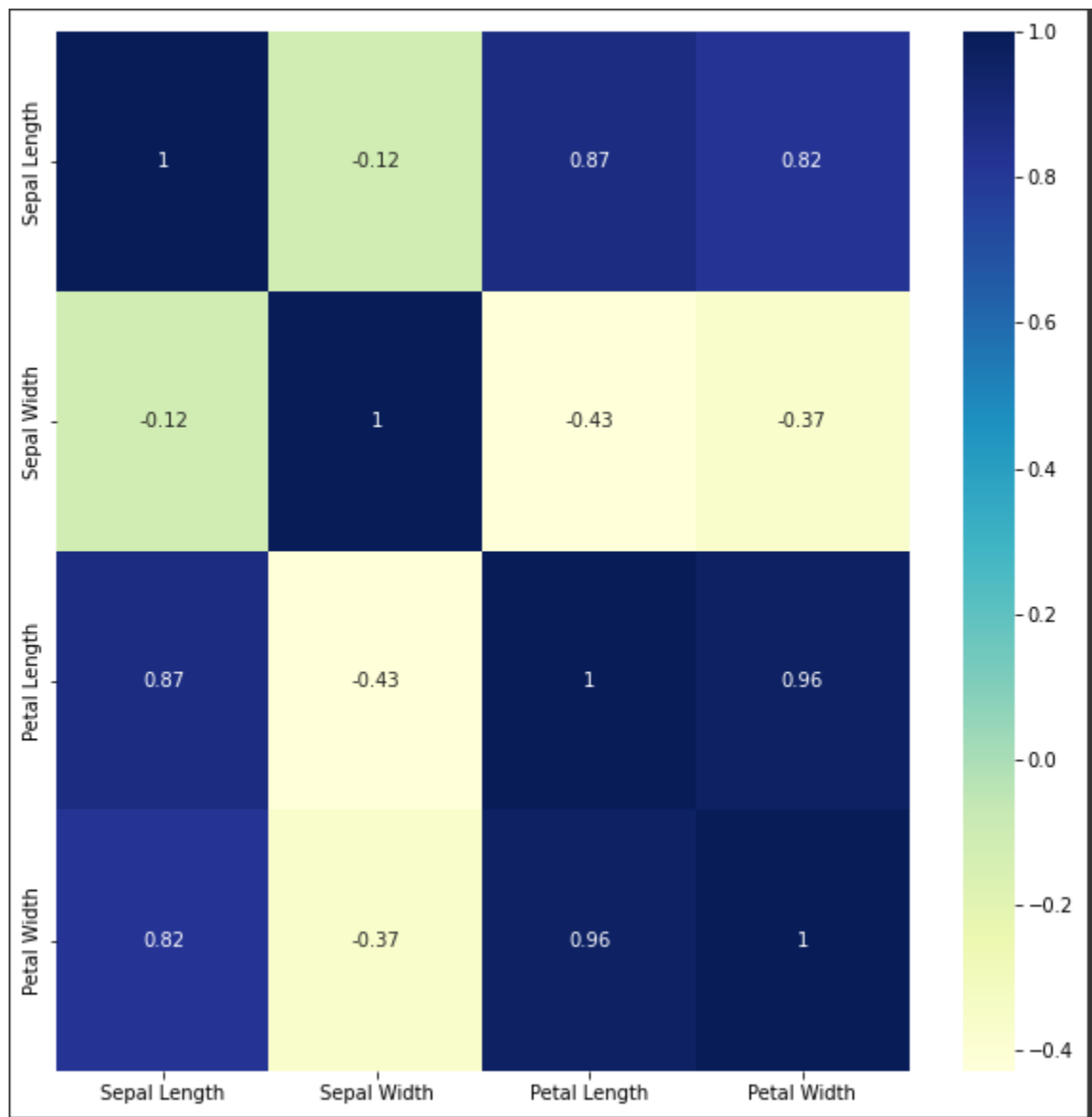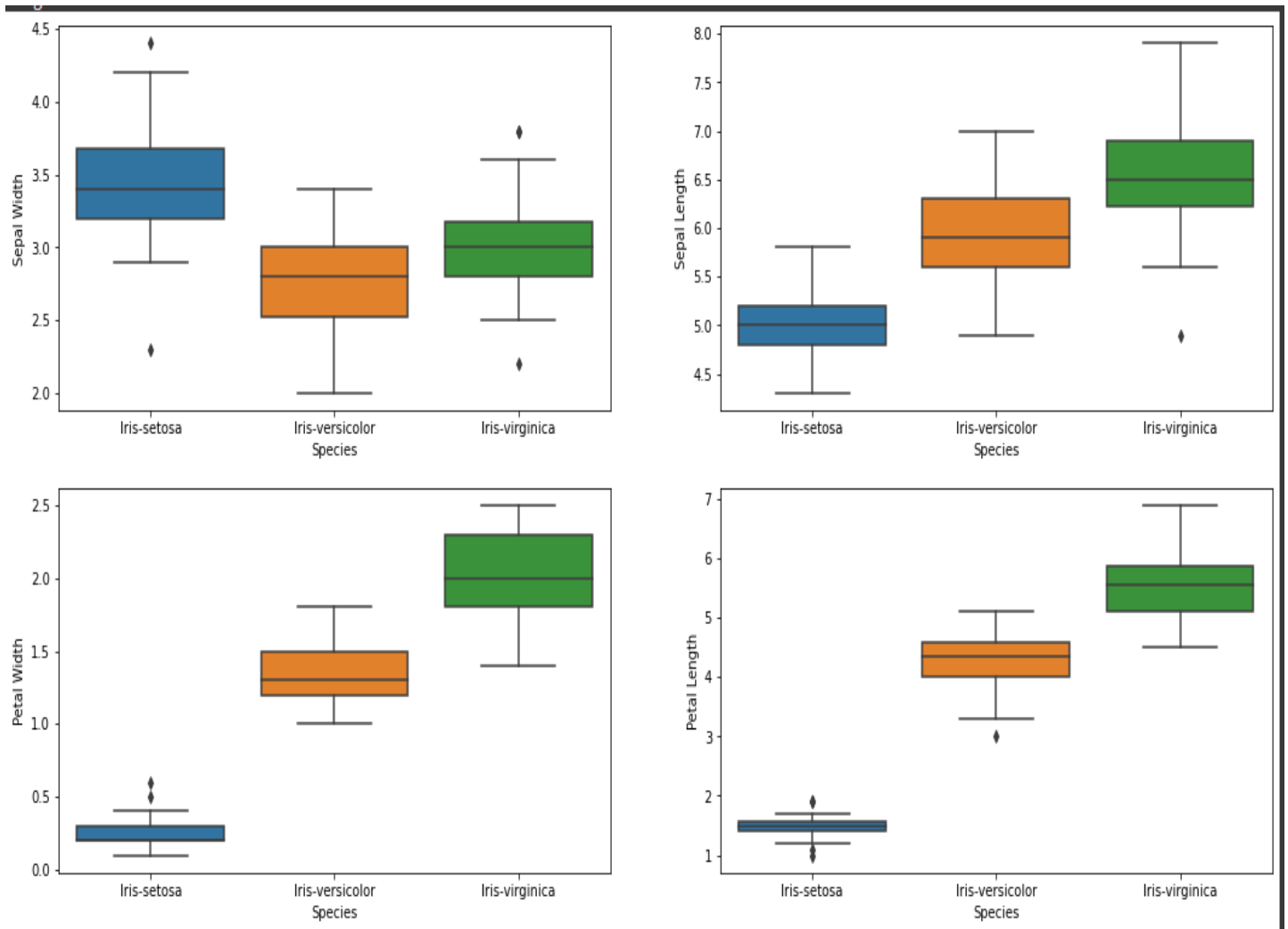

Figure 3.4

It shows the same as 2D scatter but it's so clear here that Setosa is linearly Separable.

correlation matrix

According to the correlation matrix results Petal Length and Petal Width have possitive correlation which is proved by the plot above as well as Petal Length and Sepal Width.

Box Plot

As shown the figure above, the box Plot for the Sepal Length almost doesn't have outliers except one for Virginicia. The median at Sepal width and Petal Length is almost the same as its mean. 75% Is almost close to the maximum as the 25% close to the minimum which indicates that there isn't too many outliers.