a) Relationship between two sepal variables

The below scatter plot shows the relationship between sepal width and sepal length. The points are plotted with different colours based on the classes.
If the best fit line is plotted based on different classes, there is a positive strong correlation between sepal width and sepal length.
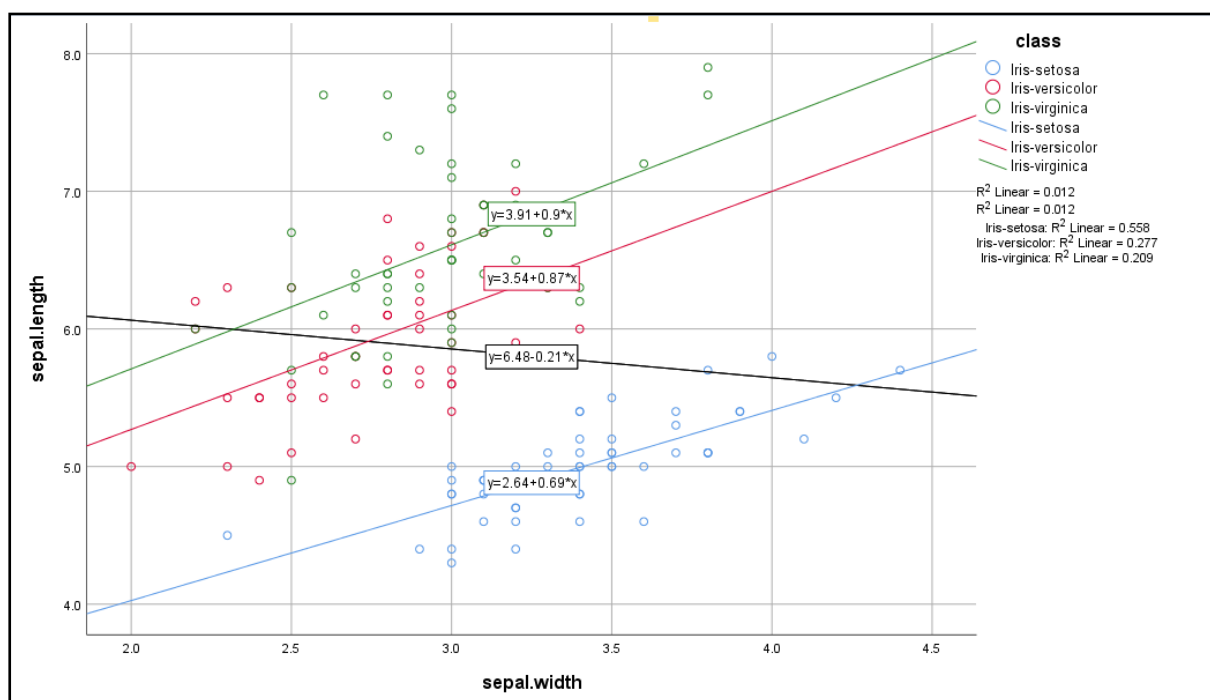If the fit line is plotted on overall data, there is a negative correlation between sepal width and sepal length.

Classification algorithm will be successful in classifying the data based on sepal length and sepal width.
Considering the values of each variable are defined or bounded by certain boundaries, it will be easier to classify a class on a record for the given set of variables.

Example: For 'Setosa', the sepal width is between 2.9 cm and 4.4 cm while the sepal length is between 4.2 cm and 5.9 cm.
Any record whose sepal length and width is bounded by these values, they can be classified as 'Setosa'.

b) Relationship between two petal variables

The below scatter plot shows the relationship between petal width and petal length. The points are plotted with different colours based on the classes.
If the best fit line is plotted based on different classes, there is a positive strong correlation between petal width and petal length.
If the fit line is plotted on overall data, there is a positive correlation between petal width and petal length.
As the petal width increases, petal length increases.

Classification algorithm will be successful in classifying the data based on petal length and petal width.
Considering the values of each variable are defined or bounded by certain boundaries, it will be easier to classify a class on a record for the given set of variables.

Example: For 'Setosa', the petal width is between 0.1cm and 0.7 cm while the petal length is between 1.0 cm and 2.0 cm.
Any record whose petal length and width is bounded by these values, they can be classified as 'Setosa'.
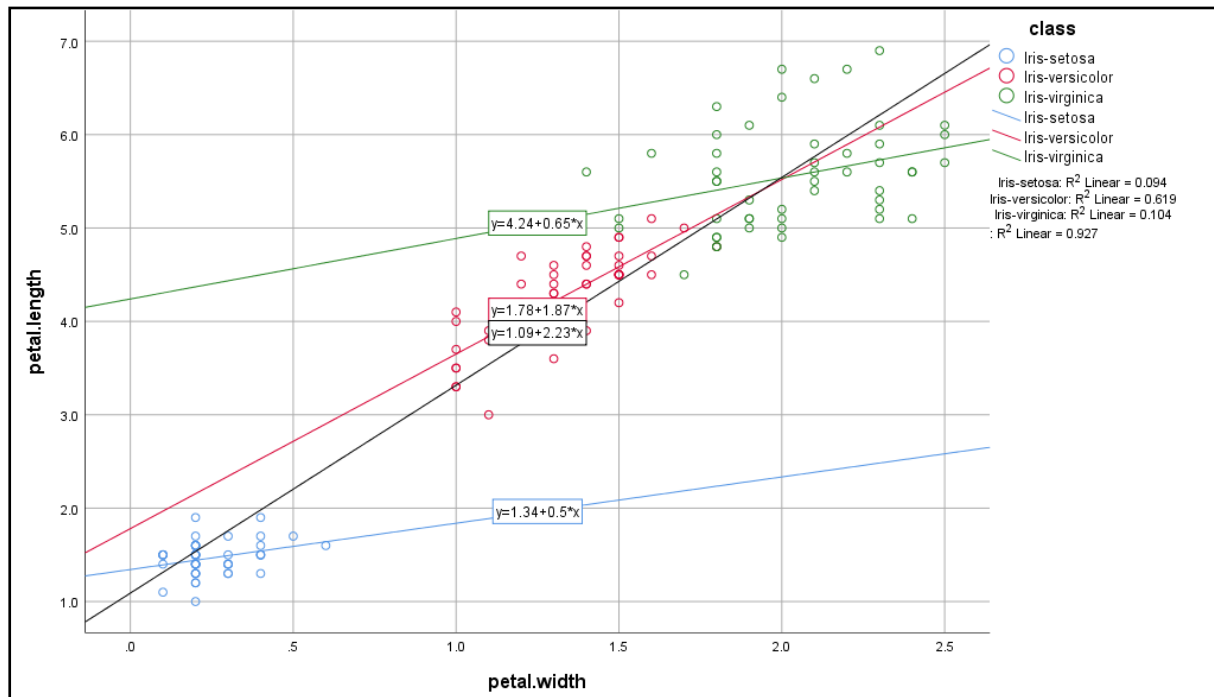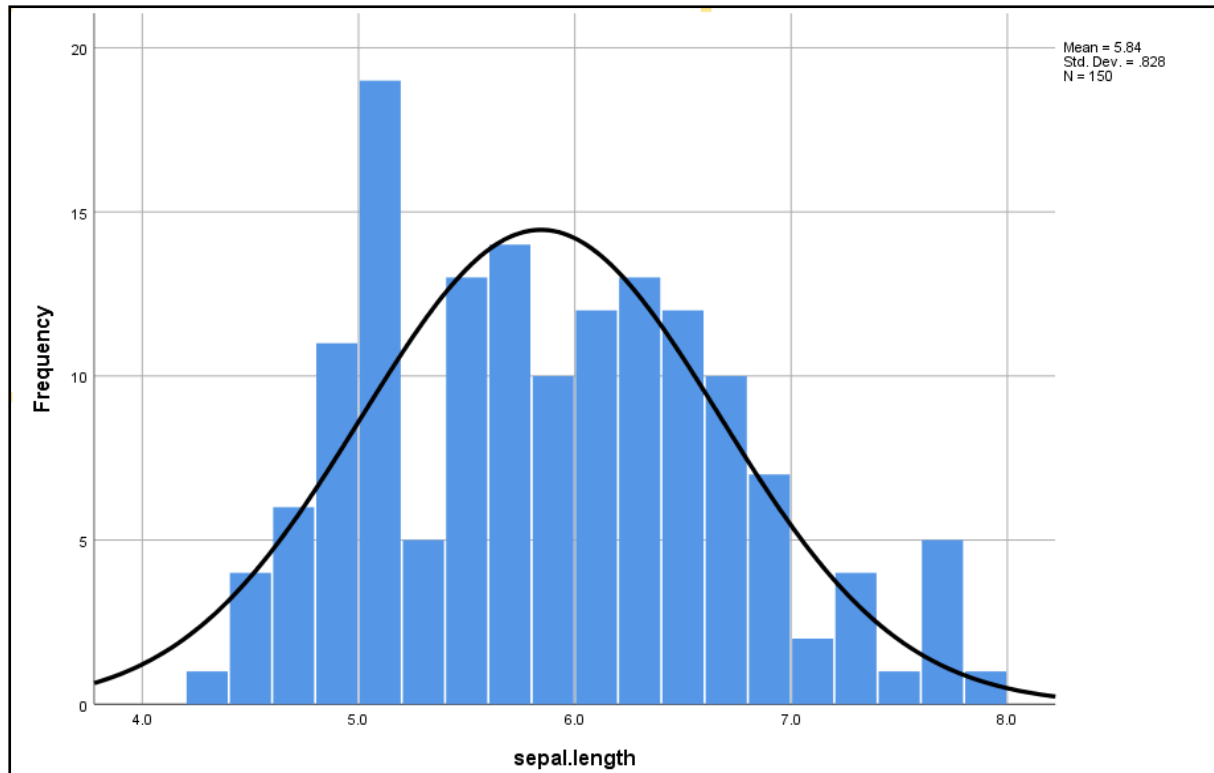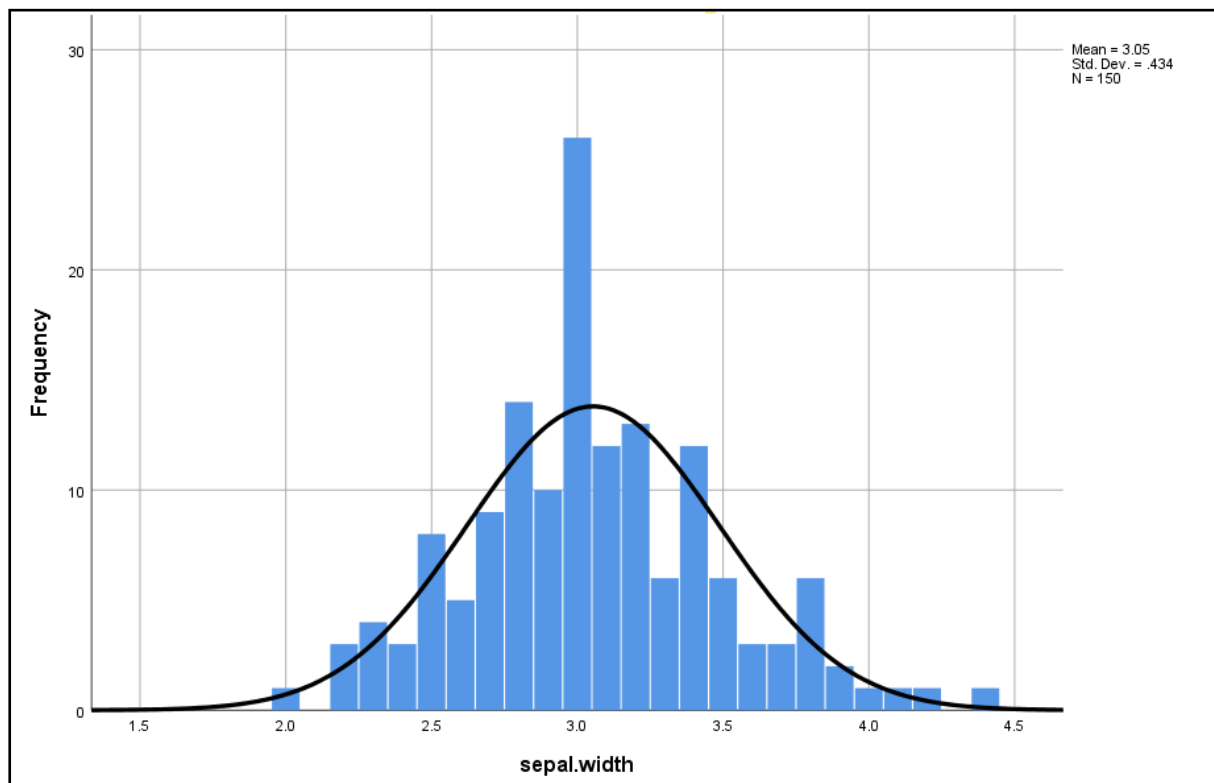

c) **Histograms**

The below histogram for sepal length is divided into bins of 0.2 cm.
It shows a normal distribution with 95% of the data is between two standard deviation.
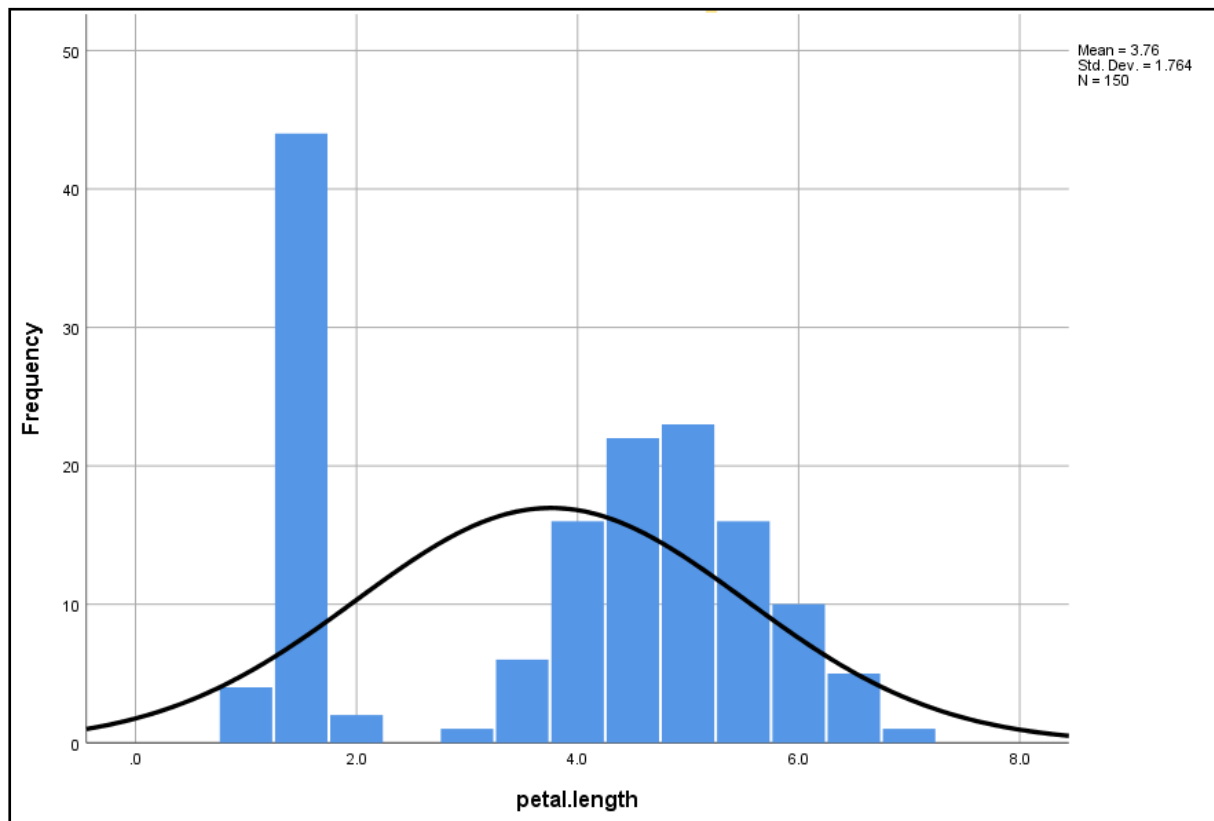The mean of sepal length is 5.84 and standard deviation is 0.828.

The below histogram for sepal width is divided into bins of 0.1 cm.
It shows a normal distribution with 95% of the data is between two standard deviation.
The mean of sepal width is 3.05 and standard deviation is 0.434.

The below histogram for petal length is divided into bins of 0.5 cm.
It shows a normal distribution with 95% of the data is between two standard deviation.

The mean of petal length is 3.76 and standard deviation is 1.764.

The below histogram for petal width is divided into bins of 0.1 cm.
It shows a normal distribution with 95% of the data is between two standard deviation.
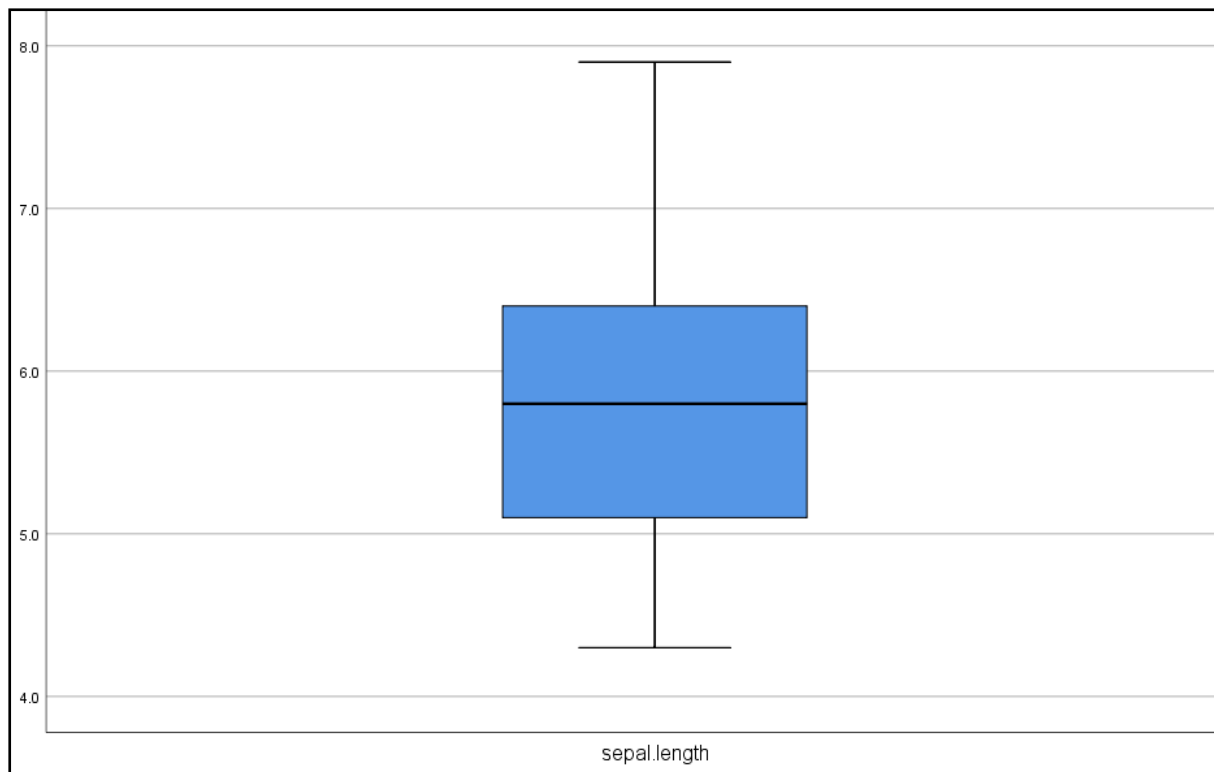The mean of petal width is 1.2 and standard deviation is 0.763.
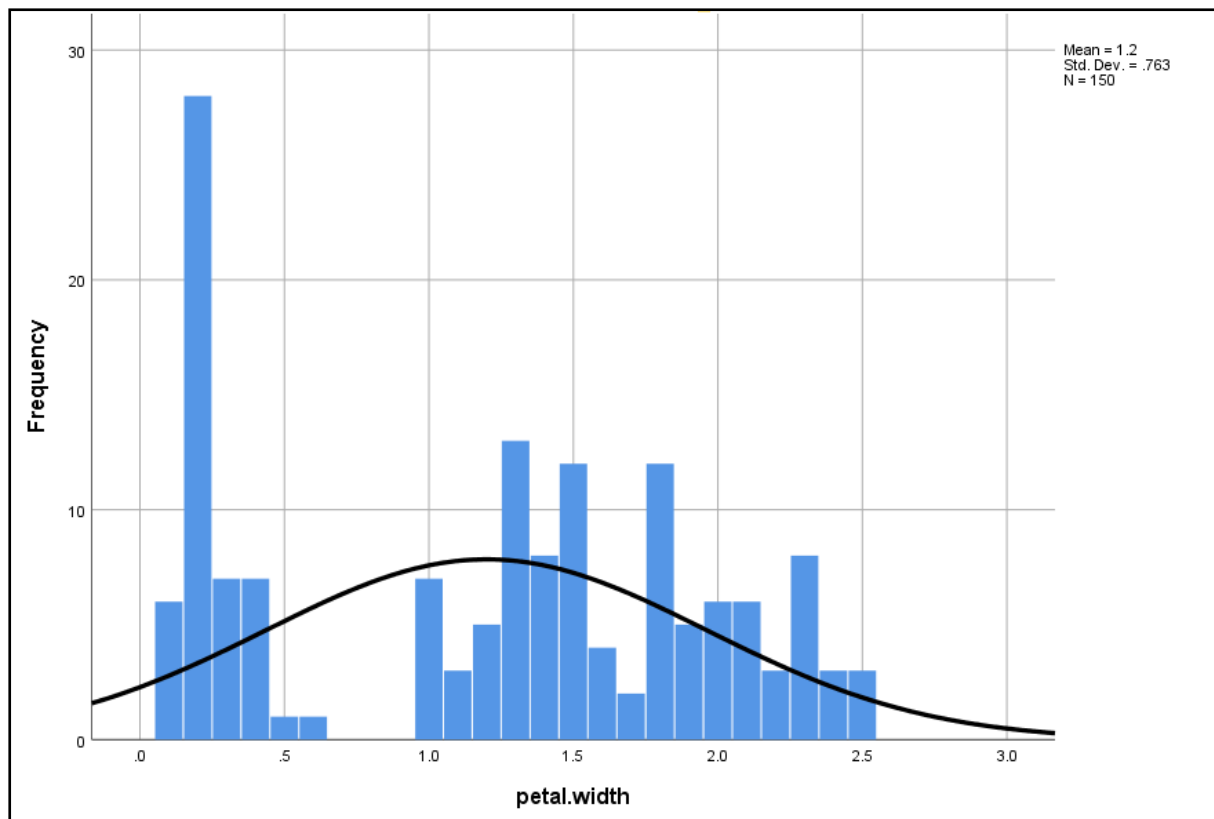
d) Sepal Length

There are no outliers in sepal length.
The below boxplot shows the maximum length at 7.9 cm and minimum length at 4.2 cm.
The 1st and 3rd quartile is between 5.1 cm and 6.3 cm. The median is at 5.8 cm.

sepal.length

e) Petal Length

There are no outliers in petal length.
The below boxplot shows the maximum length at 6.95 cm and minimum length at 1.0 cm.
The 1st and 3rd quartile is between 1.7 cm and 5.1 cm. The median is at 4.3 cm.

petal.length