**IRIS FLOWER CLASSIFICATION**

**Executive:** The aim is to classify iris flowers among three species (setosa, versicolor or virginica) from measurements of length and width of sepals and petals.

**Background:** In 1935, a study was made of the differences in measurements (width and length of petals and sepals) of 3 closely related iris flowers by Anderson:



The iris data set consists of 50 samples from each of three species of Iris (*Iris setosa, Iris virginica* and *Iris versicolor*). Four features were measured from each sample: the length and the width of the sepals and petals, in centimeters.

**Purpose:** This analysis will help to find out the species of Iris by the length and width of the sepals and petals.

**Questions:**

1. Import iris.csv file from folder iris\_dataset.

2. Exploratory Analysis.

A. Explore / Print first 3 Records from Dataset.

B. Find Dimension of Dataset.

C. Find Names, Class of features in the Dataset.

D. Find missing values (if any) & make the data consistent by removing it.

E. Find Structure of Data.

F. Find mean, median, quartile, max, min data for every feature.

G. Plot a Boxplot Graph, Pie chart respective to their Species.

H. Subset tuples based on their Species in different R-Object.

I. Plot a BoxPlot Graph for Individual R-Object.

J. Plot a Histogram on feature Petal lengths of iris dataset.

K. Plot a Histogram for Petal Lengths of Different Species on different Graph.

L. Find correlation between multiple features also plot a scatter plot for correlation.

4. Classify Data based on iris Species and plot a Decision Tree.

**Methods:**

Read.csv() was used to read and import the dataset from the system to do the analysis.

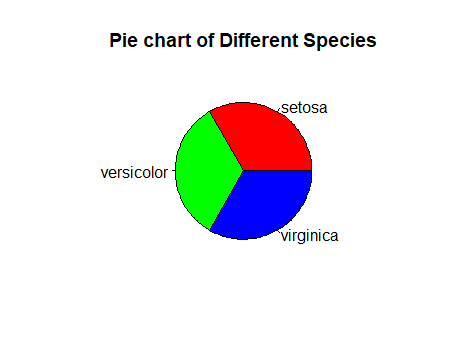
Exploratory analysis was done using head(),dim(),names(),class(),str(),summary(),boxplot(),hist(),subset(), and cor().

Data was classified on the basis of species and a decision tree machine learning algorithm was used.

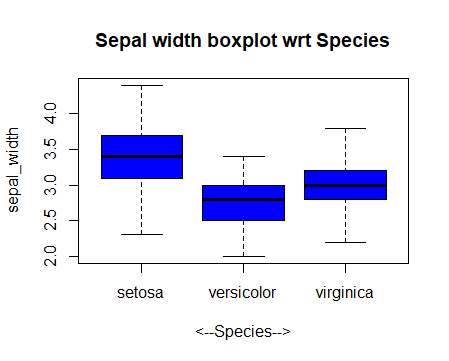
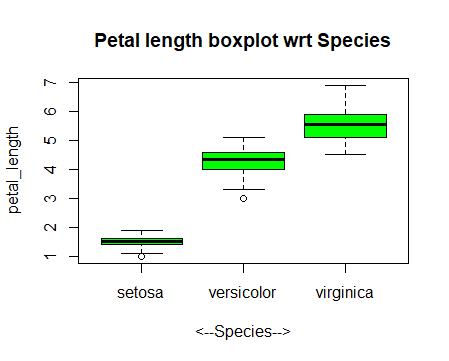
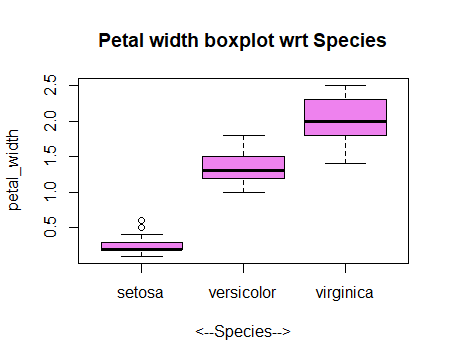
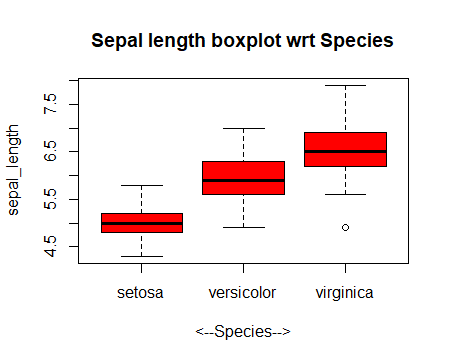
**Instrumentation:** The analysis was done on RStudio with R-programming

**Results:** The species was successfully classified by the analysis. Following are the findings and the plots.

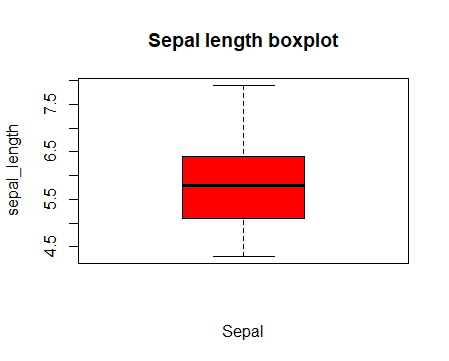
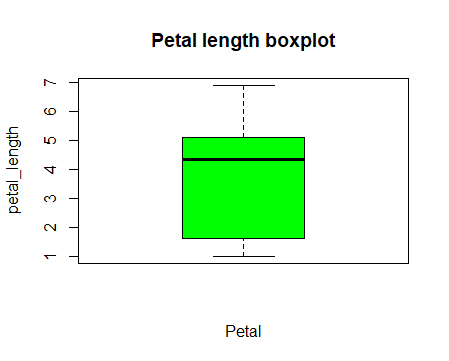
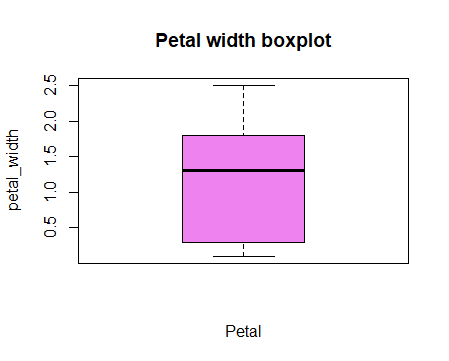
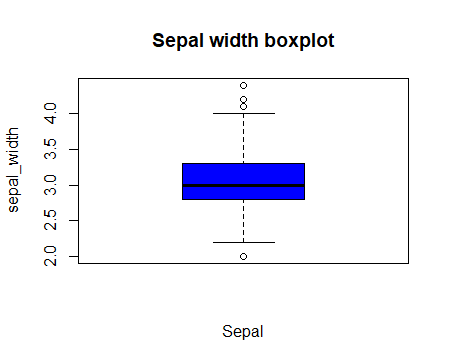
Pie chart of Different Species

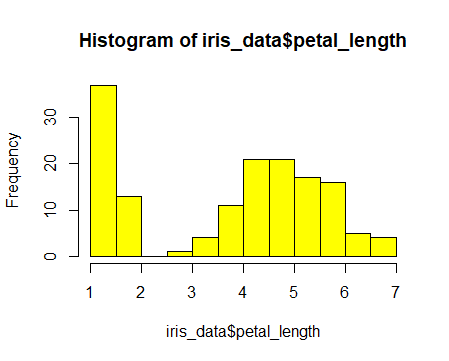


Boxplot Graph respective to their Species

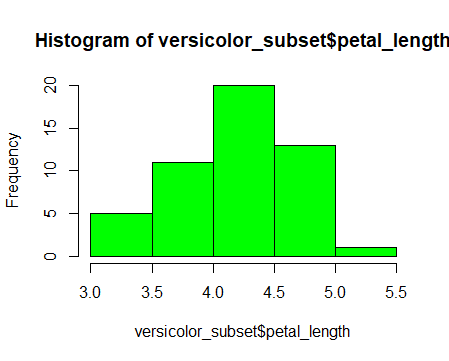
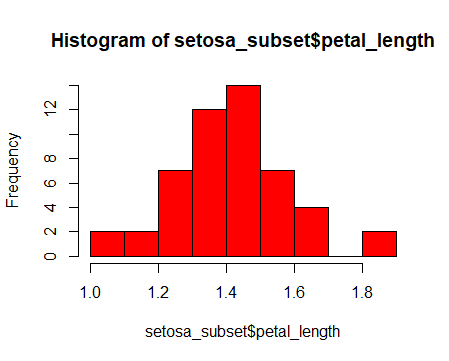
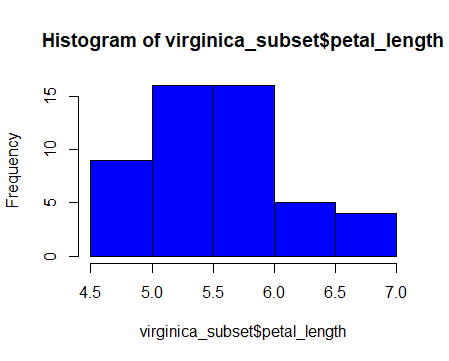
  

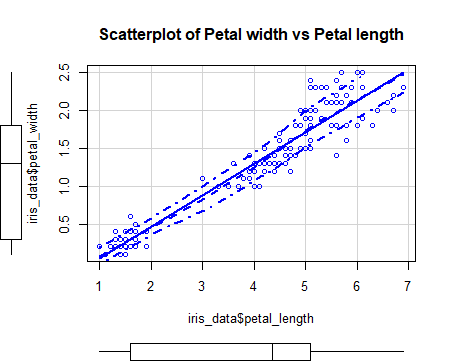
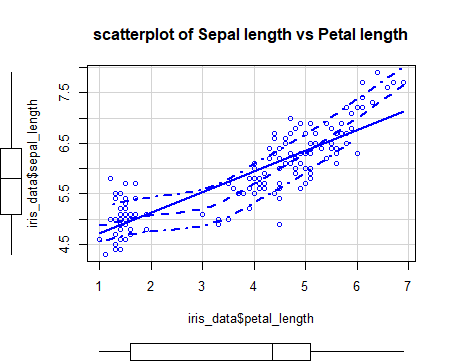
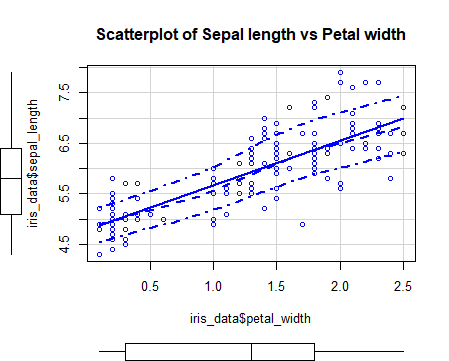
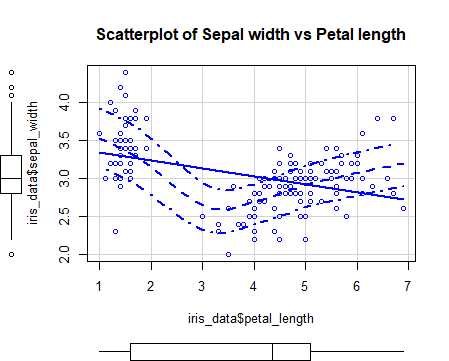
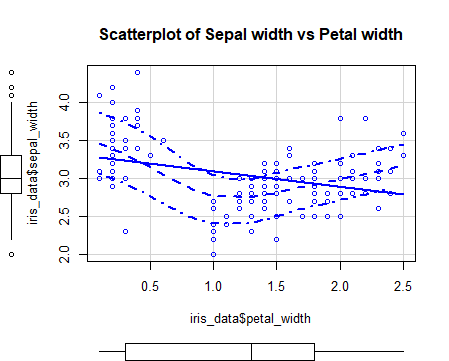
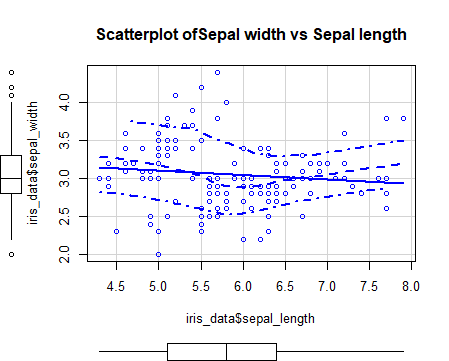
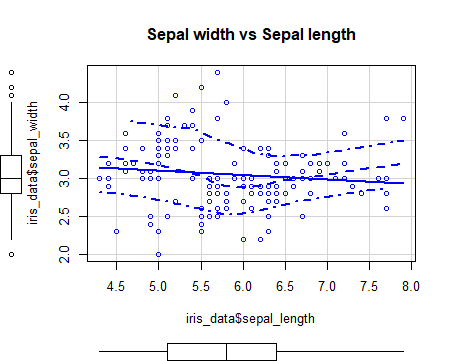
BoxPlot Graph for Individual R-Object

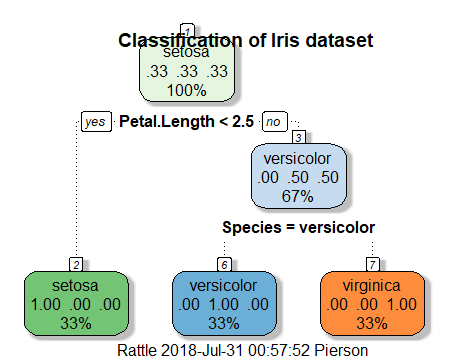
Histogram on feature Petal lengths

Histogram for Petal Lengths of Different Species

  …………………………………………

A scatter plot for correlation between multiple features…………………….

Classification of dataset:



**Summary:**

The iris dataset was analyzed. The data given had no missing values and hence nothing was omitted. The data had 3 species of iris namely, iris setosa, iris virginica and iris versicolor each having 50 sets of petal and sepal length. The mean, median, quartile, max, min data for every feature was observed. Various boxplots were plotted to show the same. It was observed that there were differences between the lengths and width of the sepal and petal of the flowers. Decision tree algorithm was used to show the classification of the species. Thus the classification of the species was done

References:

1) <https://www.neuraldesigner.com/learning/examples/iris_flowers_classification>

2)Steven J Schiff - The Dynamical Evolution of Seizures Discriminating Spatiotemporal Patterns in Neurons

3) <https://en.wikipedia.org/wiki/Iris_flower_data_set>