M8_Exercise_BoxPlot_FirstName_LastName

Installing files and loading datasets

```r
#installing library  and loading  dataset
library(datasets)
data("iris")
summary(iris)
names(iris)
```

```
> library(datasets)
> data("iris")
> summary(iris)
  Sepal.Length    Sepal.Width     Petal.Length    Petal.Width
 Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
 1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
 Median :5.800   Median :3.000   Median :4.350   Median :1.300
 Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
 3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
 Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
       Species
 setosa    :50
 versicolor:50
 virginica :50



> names(iris)
[1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
>
```

R code for the sepal length of each of these three different species of iris

```
#the sepal length of each of these three different species of iris.
boxplot(Sepal.Length~Species ,data = iris ,xlab = "Species",ylab = "Sepal Length" ,
main="Sepal length of different species of iris Dataset")
```
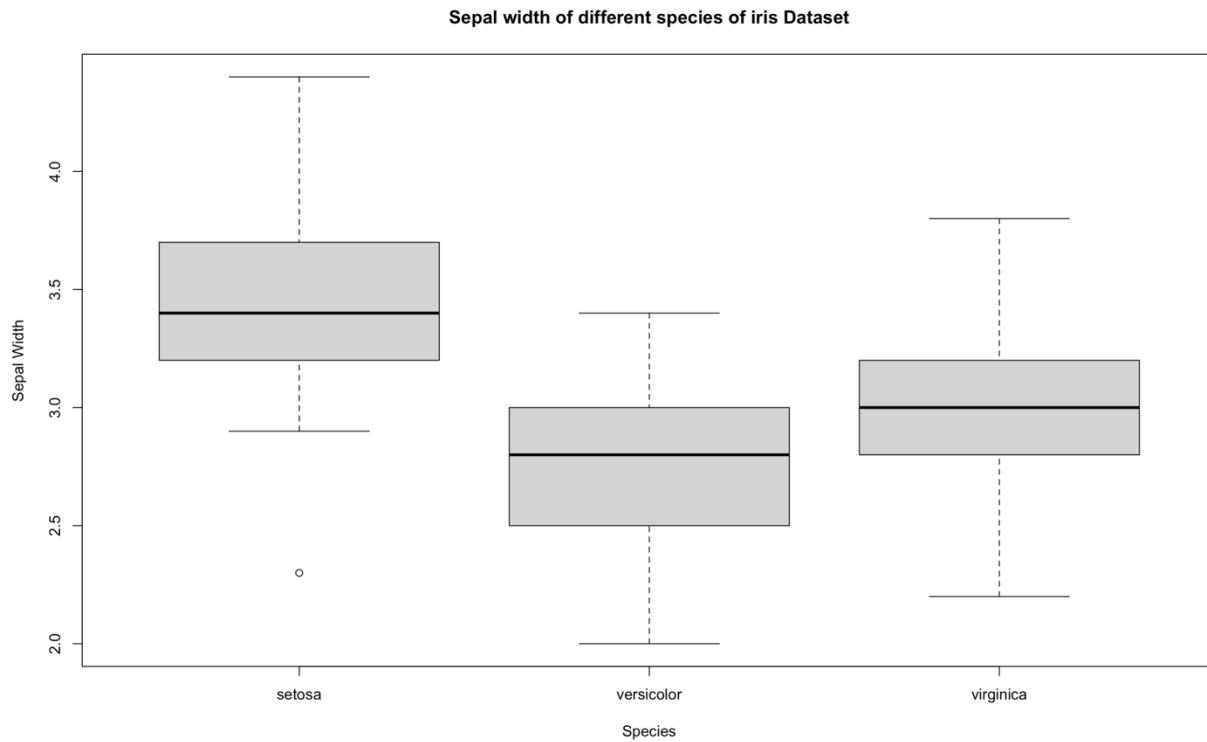
Screenshot

Sepal length of different species of iris Dataset
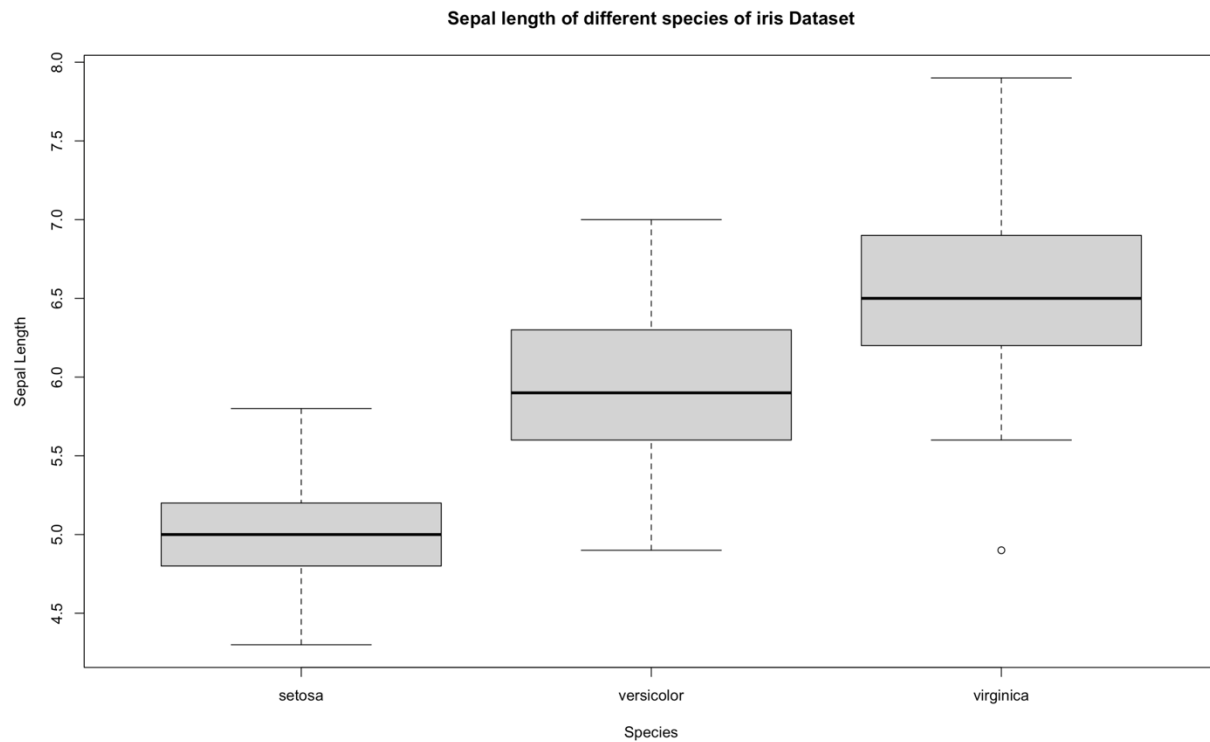
R code for the sepal width of each of these three different species of iris.

```r
#the sepal length of each of these three different species of iris.
boxplot(Sepal.Width~Species ,data = iris ,xlab = "Species",ylab = "Sepal Width" ,
main="Sepal width of different species of iris Dataset")
```
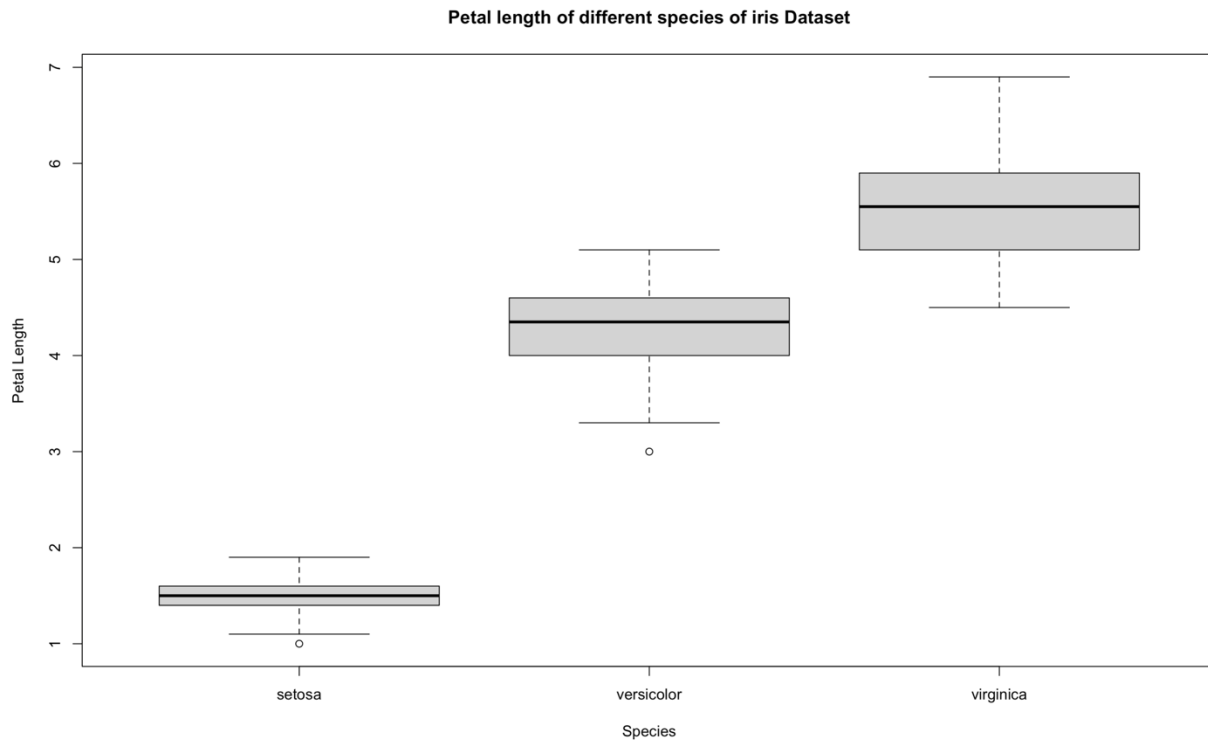
Screenshot

Sepal width of different species of iris Dataset

R code for the petal length of each of these three different species of iris

```r
#the petal length of each of these three different species of iris
boxplot(Petal.Length~Species ,data = iris ,xlab = "Species",ylab = "Petal Length" ,
 main="Petal length of different species of iris Dataset")
```

Screenshot

Petal length of different species of iris Dataset

R code for the petal width of each of these three different species of iris

```
#the petal width of each of these three different species of iris
boxplot(Petal.Width~Species ,data = iris ,xlab = "Species",ylab = "Petal Width" ,
main="Petal width of different species of iris Dataset")
```

Screenshot

Petal width of different species of iris Dataset
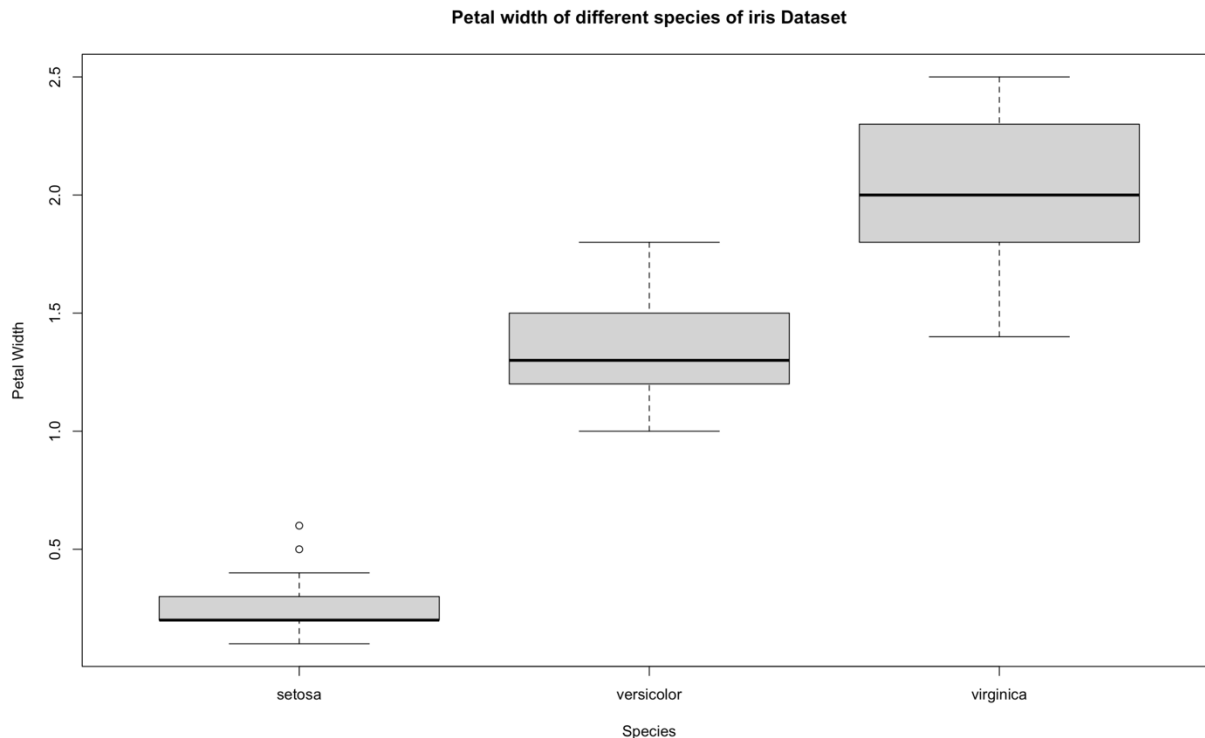
How to identify the species of the iris flower based on the plots

Observe the distribution of the samples in the plot. Look for patterns or clusters that can be associated with a particular species. For example, you may notice that the samples with longer Sepal Length and wider Sepal Widths are predominantly of a specific species. Repeat the process for the other variables in the dataset, such as Petal Length and Petal Width, to see if any different patterns can help identify the species. Once you have identified potential patterns, use them to make an educated guess about the species of the new sample based on where it falls in the plots. For example, if the new sample has a Sepal Length and Sepal Width that fall in the cluster associated with a certain species, it is likelier to belong to that species.However, it is essential to note that this method is less reliable than using machine learning algorithms to classify the samples, as it relies on visual inspection and interpretation of the plots. Machine learning algorithms can more accurately classify samples by using statistical models and learning patterns from the data

Based on the given data, if we are given a iris flower, we can take measurements of the petals and sepals.We see that petal length and width measurements of the setosa species do not overlap with other species.So, if the petal length is in the interval of [1.0, 1.9] and the petal width is in the interval of [0.1, 0.6], then the iris is of the setosa species. There is a lot of difference between the petal length/width of other species soeven if the sample measurement is out of the intervals, it could just be an outlier. Their petal measurements are smaller and less spread-out than those of the other two species as well. If the sepal width is a lot, it could be setosa. We could also say that if petal length is less than 2.5 or petal width is less than 1, it is of setosa species. While only using sepal measurements is good to determine if it is of setosa pecies, for the other two we will have to consider all the measurements. In between the other two species, on average, virginica has higher values than versicolor. Since their intervals overlap, there is a higher chance to make a wrong guess. We can say that if petal length is more than 2.5 i.e. not overlapping with setosa, and less than or equal to 5.1 cm (max petal length for versicolor) and petal width less than 1.8 (max petal width for versicolor), then it could be of the versicolor species. Similarly, if the petal length is greater than 5.1 cm and petal width is greater than 1.8 cm, it could be of virginica species. But here since the petal lengths are overlapping too much and look positively associated, some virginica samples could have petal width more than 1.8 and petal length between 2.5 cm and 5.1 cm. After this, we will have to manually check if the results are right. To more accurately make classifications, we would need more data or different visualizations.