18BCE080_PRAC8

Ishan Tewari

29/04/2021

Loading the iris dataset:

```
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
##
##
##
```

Checking the shape and feature names of the dataset:

```
print("Shape of IRIS dataset")
```

```
## [1] "Shape of IRIS dataset"
```

```
dim(iris)
```

```
## [1] 150   5
```

```
print("Features of IRIS dataset")
```

```
## [1] "Features of IRIS dataset"
```

```
names(iris)
```

```
## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
```

Exploring the structure of IRIS dataset.

```
str(iris)
```

```
## 'data.frame':    150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

Exploring the entries of the dataset for 'virginica' species

```
virginica = iris[iris['Species'] == "virginica", ]
head(virginica)
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 101          6.3         3.3          6.0         2.5 virginica
## 102          5.8         2.7          5.1         1.9 virginica
```

```
## 103               7.1               3.0               5.9               2.1 virginica
## 104               6.3               2.9               5.6               1.8 virginica
## 105               6.5               3.0               5.8               2.2 virginica
## 106               7.6               3.0               6.6               2.1 virginica
```

Finding the mean and standard deviation for 'sepal width' and 'sepal length' for each type of species.

```
print("Mean of sepal length")
```

```
## [1] "Mean of sepal length"
```

```
aggregate(x = iris$Sepal.Length,
          by = list(iris$Species),
          FUN = mean)
```

```
##        Group.1     x
## 1       setosa 5.006
## 2 versicolor 5.936
## 3  virginica 6.588
```

```
print("Standard deviation of sepal length")
```

```
## [1] "Standard deviation of sepal length"
```

```
aggregate(x = iris$Sepal.Length,
          by = list(iris$Species),
          FUN = sd)
```

```
##        Group.1         x
## 1       setosa 0.3524897
## 2 versicolor 0.5161711
## 3  virginica 0.6358796
```

```r
print("Mean of sepal width")
```

```
## [1] "Mean of sepal width"
```

```r
aggregate(x = iris$Sepal.Width,
          by = list(iris$Species),
          FUN = mean)
```

```
##      Group.1     x
## 1     setosa 3.428
## 2 versicolor 2.770
## 3  virginica 2.974
```

```r
print("Standard deviation of sepal width")
```

```
## [1] "Standard deviation of sepal width"
```
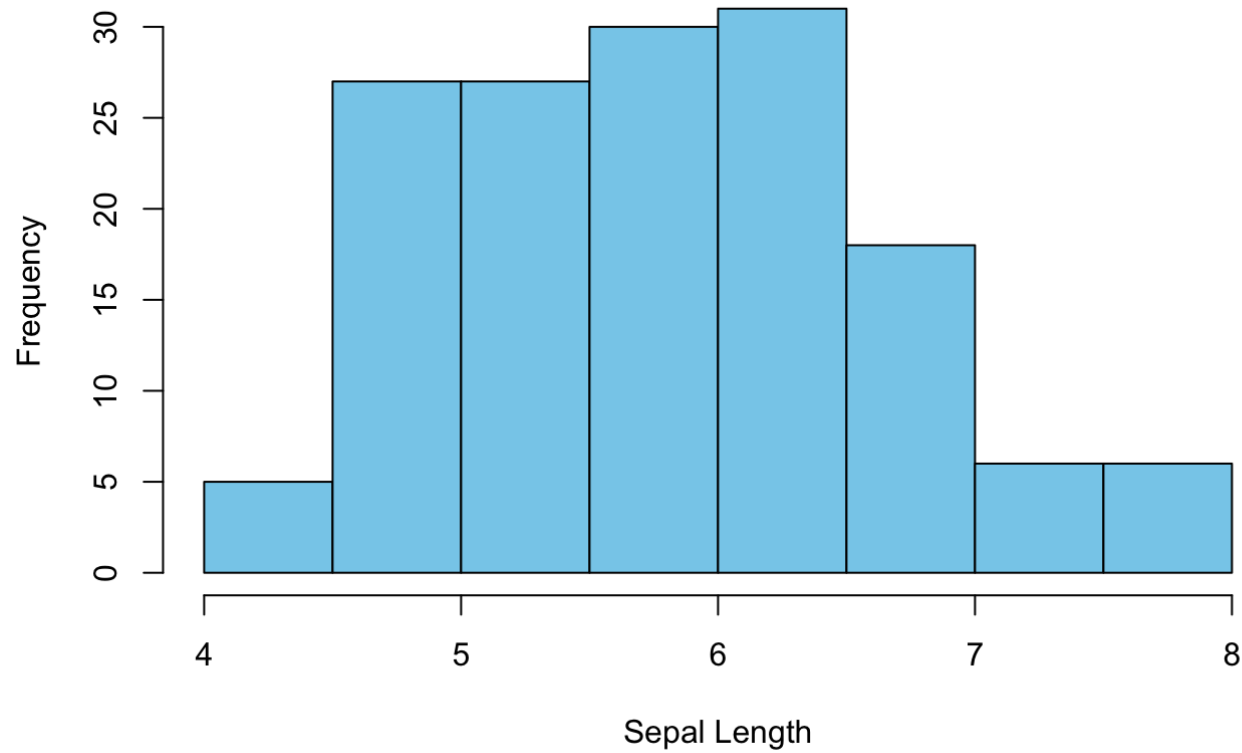
```r
aggregate(x = iris$Sepal.Width,
          by = list(iris$Species),
          FUN = sd)
```

```
##      Group.1         x
## 1     setosa 0.3790644
## 2 versicolor 0.3137983
## 3  virginica 0.3224966
```

Plotting the histogram and box plot of 'sepal.length' variable.
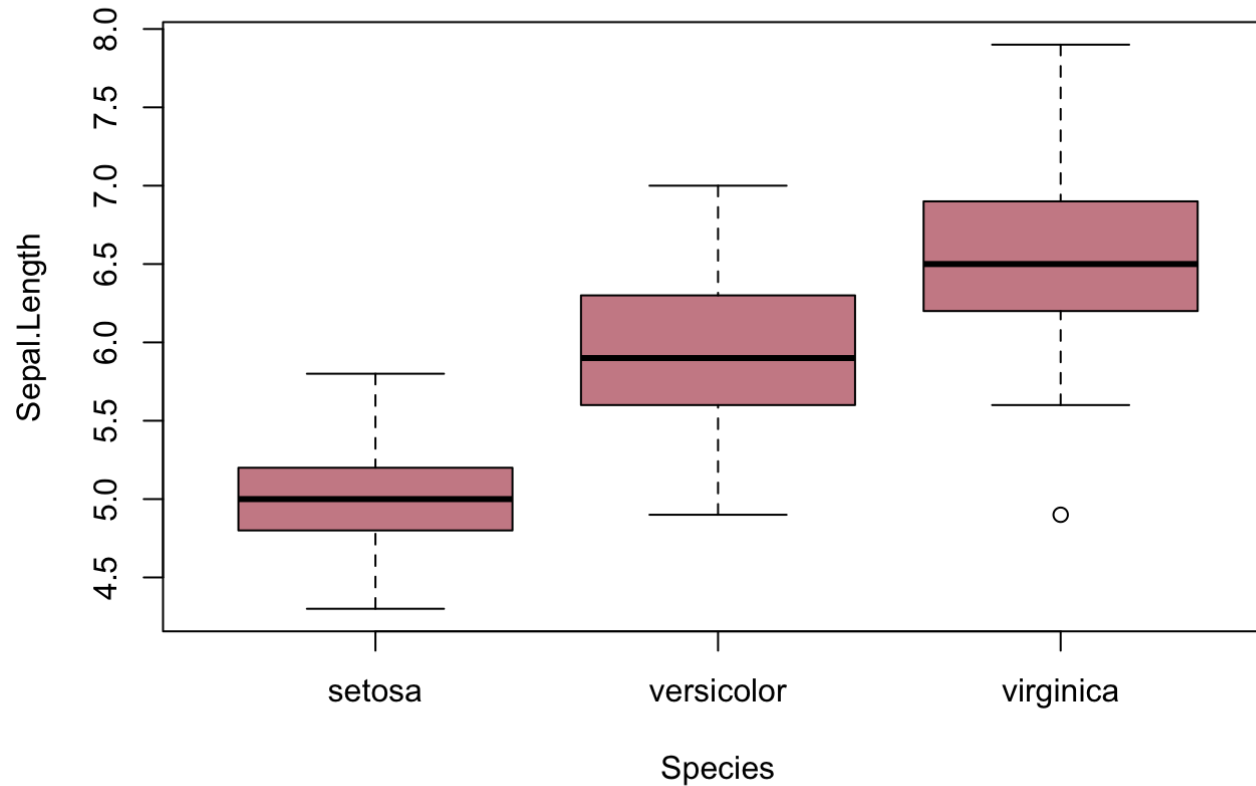
```r
hist(iris$Sepal.Length, col = "skyblue", xlab = "Sepal Length", main = "Histogram of Sepal Length of Iris Dataset")
```

Histogram of Sepal Length of Iris Dataset


```
boxplot(Sepal.Length  ~ Species, iris, main = "Boxplot of Sepal Length of iris dataset", col = "lightpink3")
```

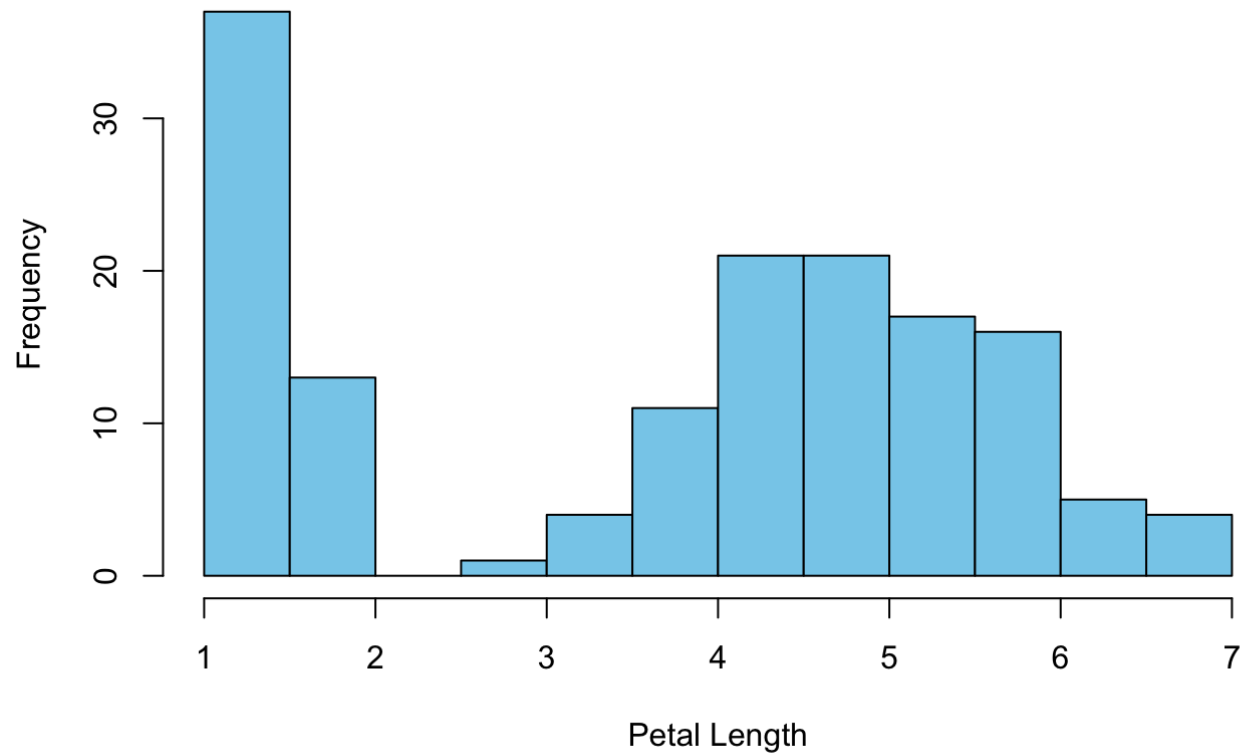**Boxplot of Sepal Length of iris dataset**

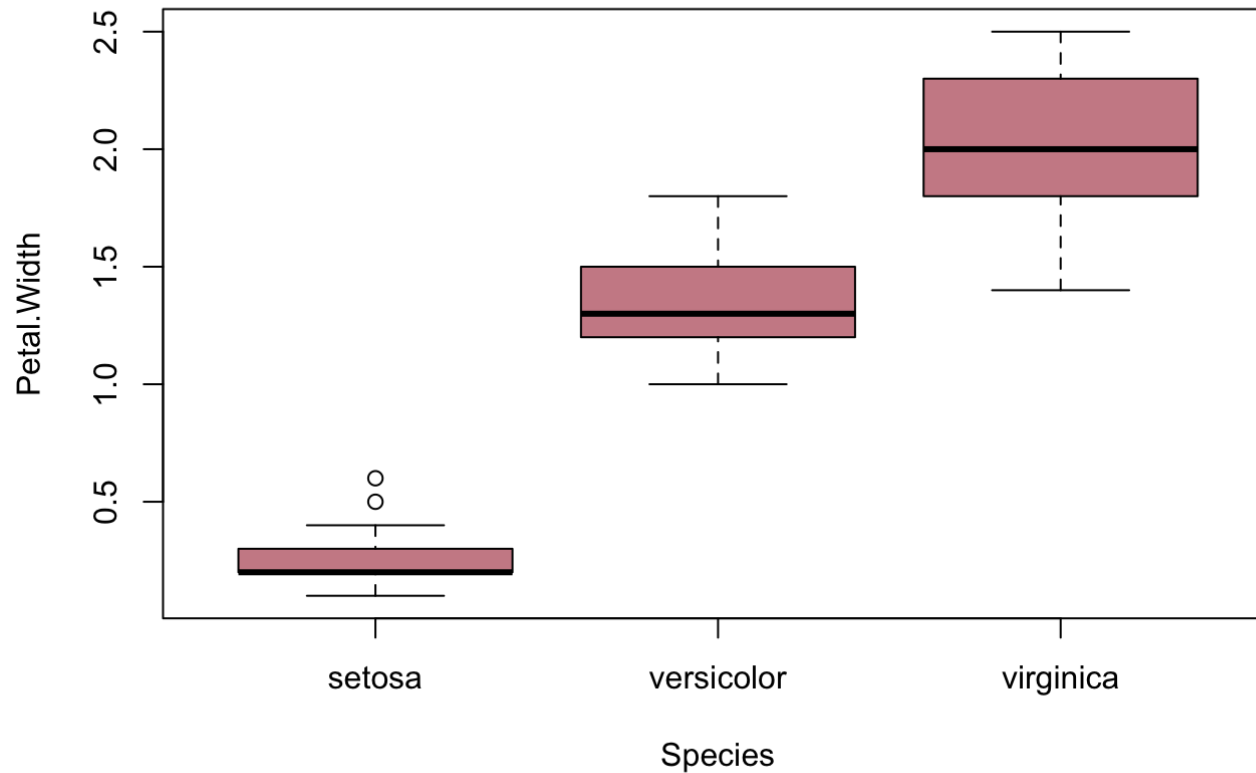Plotting the histogram and box plot of 'petal.length' variable.

```
hist(iris$Petal.Length, col = "skyblue", xlab = "Petal Length", main = "Histogram of Petal Length of Iris Dataset")
```

Histogram of Petal Length of Iris Dataset


```
boxplot(Petal.Width  ~ Species, iris, main = "Boxplot of Petal Width of iris dataset", col = "lightpink3")
```

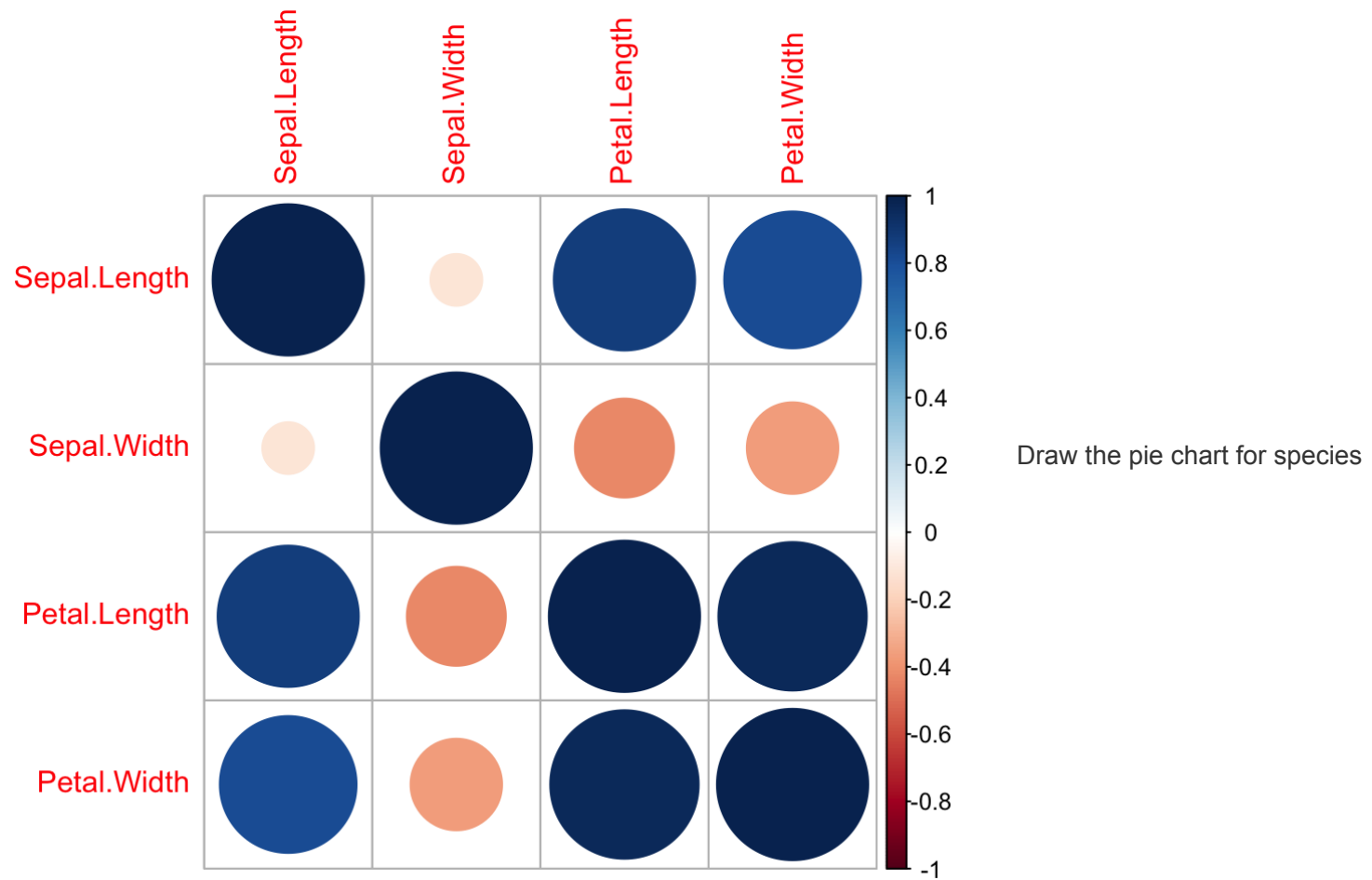# Boxplot of Petal Width of iris dataset

Plotting the correlation amongst the attributes in the dataset.

```r
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```r
corelation = cor(iris[, 1:4])
corrplot(corelation, method='circle')
```

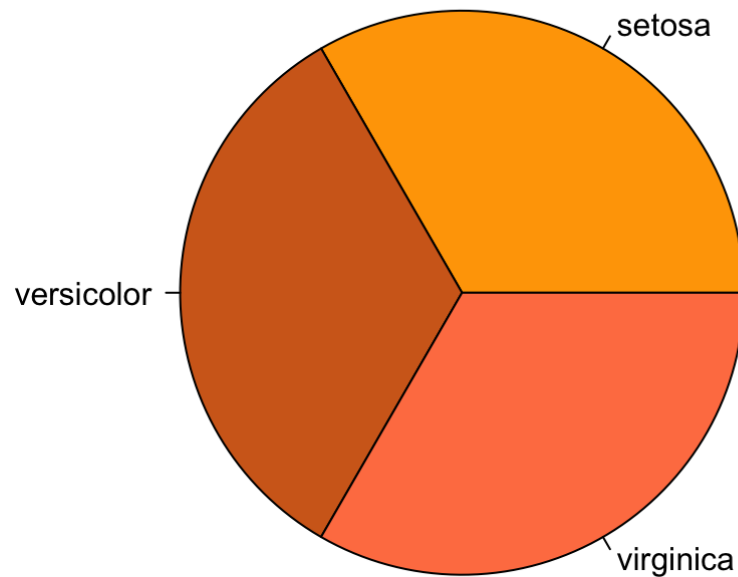Draw the pie chart for species attribute.

```
table(iris$Species)
```

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

```
pie(table(iris$Species), main = "Pie Chart of the Iris data set Species",
    col = c("orange1", "chocolate", "coral"), radius = 1)
```

Pie Chart of the Iris data set Species

Export the plot created in one of the cells as a .png file.

```
jpeg("prac8_1.jpg")
hist(iris$Sepal.Length, col = "skyblue", xlab = "Sepal Length", main = "Histogram of Sepal Length of Iris Datase
t")
dev.off()
```

```
## quartz_off_screen
##                 2
```

```
jpeg("prac8_2.jpg")
boxplot(Sepal.Length  ~ Species, iris, main = "Boxplot of Sepal Length of iris dataset", col = "lightpink3")
dev.off()
```

```
## quartz_off_screen
##                 2
```