

iris flower classification

Naimish padhan

2024-04-02

```
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 ...

# Get first 5 rows of each subset
subset(iris, Species == "setosa")[1:5,]

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1 5.1 3.5 1.4 0.2 setosa
## 2 4.9 3.0 1.4 0.2 setosa
## 3 4.7 3.2 1.3 0.2 setosa
## 4 4.6 3.1 1.5 0.2 setosa
## 5 5.0 3.6 1.4 0.2 setosa

subset(iris, Species == "versicolor")[1:5,]

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 51 7.0 3.2 4.7 1.4 versicolor
## 52 6.4 3.2 4.5 1.5 versicolor
## 53 6.9 3.1 4.9 1.5 versicolor
## 54 5.5 2.3 4.0 1.3 versicolor
## 55 6.5 2.8 4.6 1.5 versicolor

subset(iris, Species == "virginica")[1:5,]

## 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

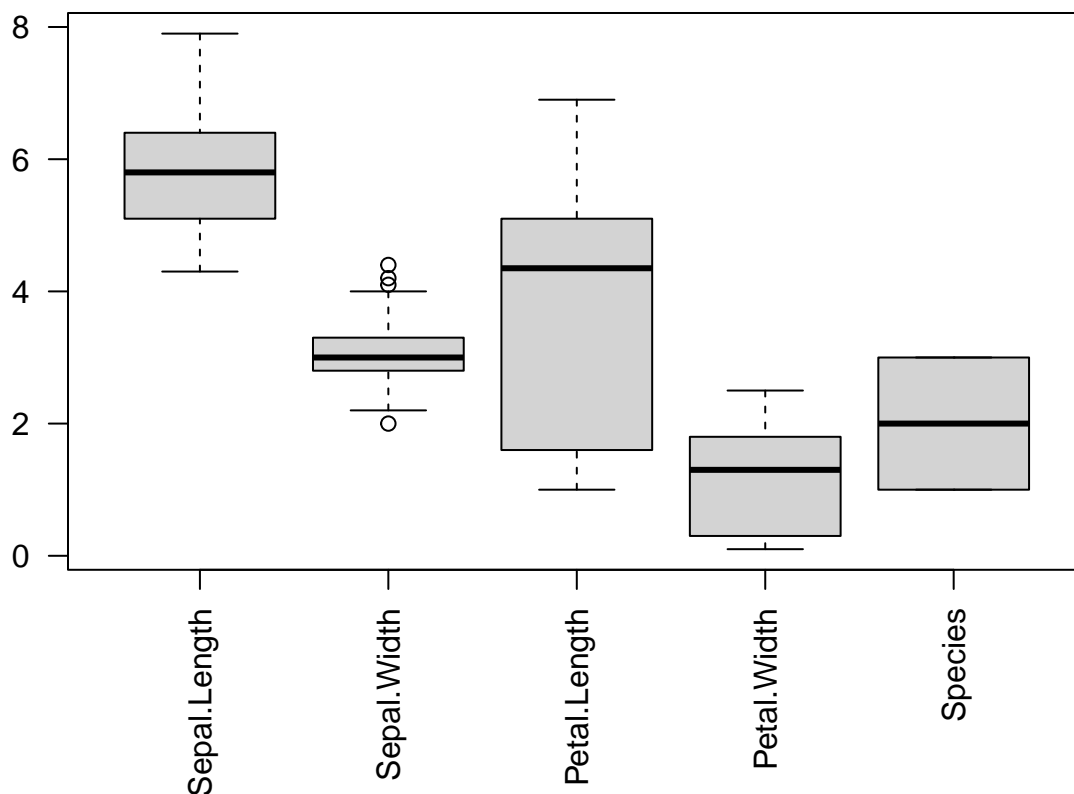
# Get column "Species" for all lines where Petal.Length < 2
subset(iris, Petal.Length < 2)[,"Species"]

## [1] setosa setosa setosa setosa setosa setosa setosa setosa setosa setosa
## [11] setosa setosa setosa setosa setosa setosa setosa setosa setosa setosa
## [21] setosa setosa setosa setosa setosa setosa setosa setosa setosa setosa
## [31] setosa setosa setosa setosa setosa setosa setosa setosa setosa setosa
## [41] setosa setosa setosa setosa setosa setosa setosa setosa setosa setosa
## Levels: setosa versicolor virginica
```

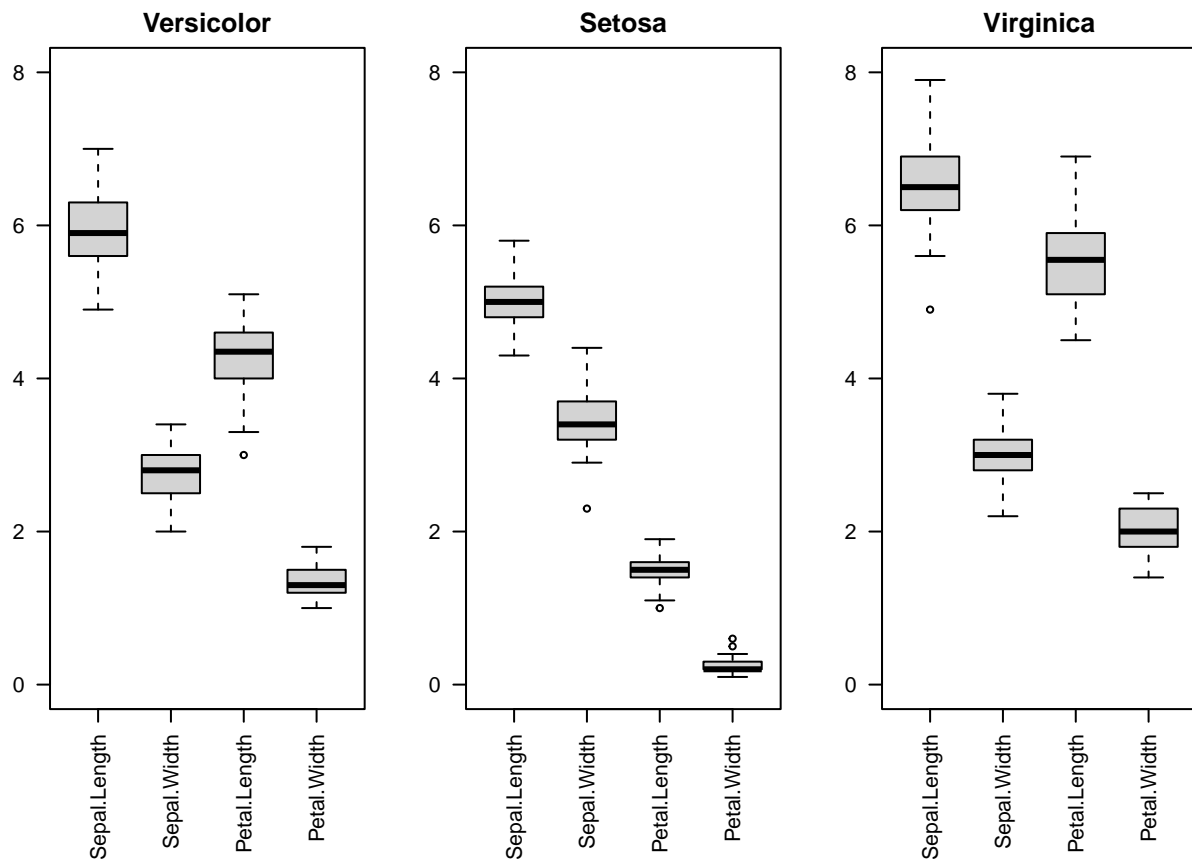
```
summary(iris)
```

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

```
par(mar=c(7,5,1,1)) # more space to labels
boxplot(iris,las=2)
```

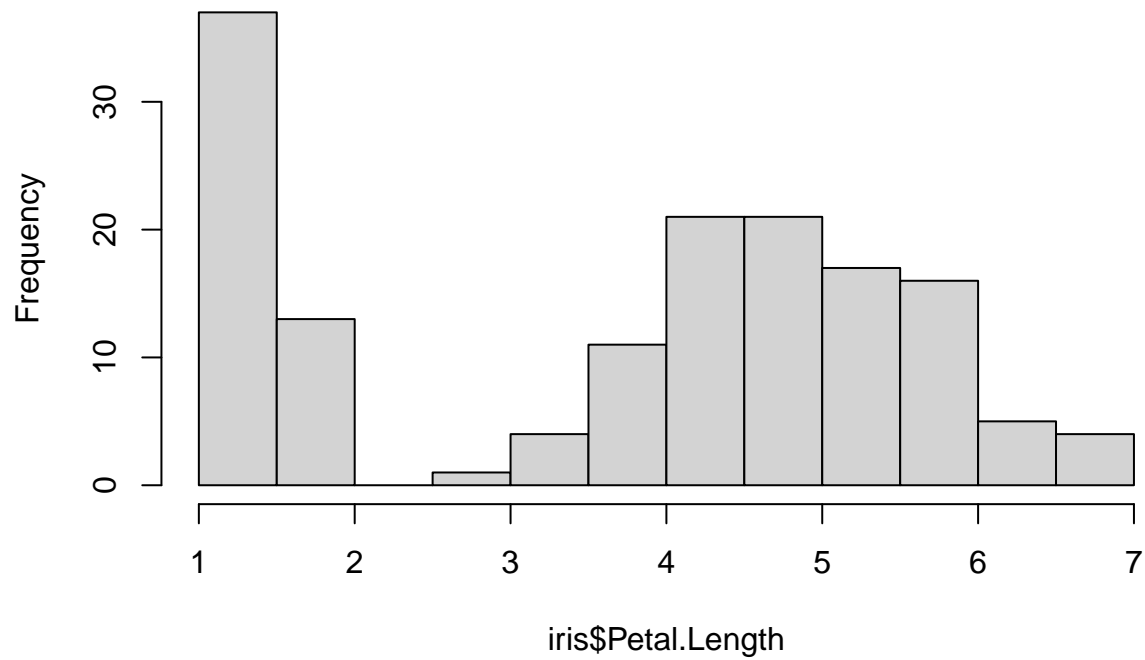


```
irisVer <- subset(iris, Species == "versicolor")
irisSet <- subset(iris, Species == "setosa")
irisVir <- subset(iris, Species == "virginica")
par(mfrow=c(1,3),mar=c(6,3,2,1))
boxplot(irisVer[,1:4], main="Versicolor",ylim = c(0,8),las=2)
boxplot(irisSet[,1:4], main="Setosa",ylim = c(0,8),las=2)
boxplot(irisVir[,1:4], main="Virginica",ylim = c(0,8),las=2)
```



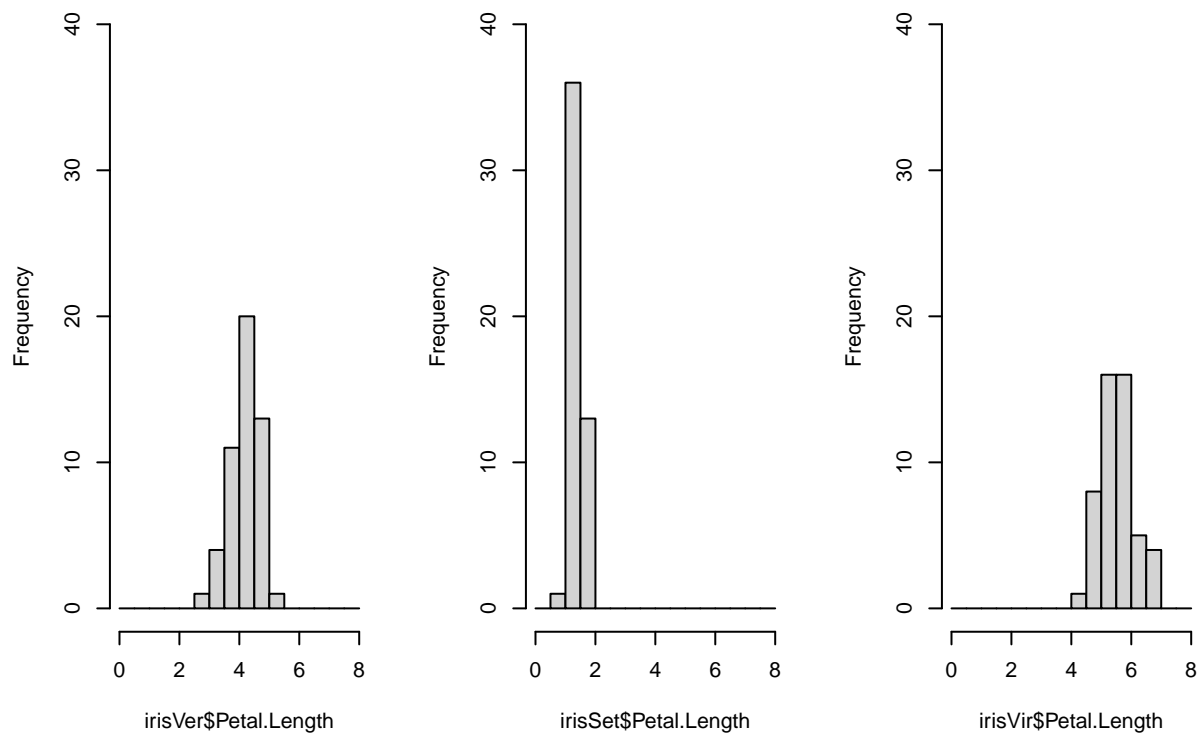
```
hist(iris$Petal.Length)
```

Histogram of iris\$Petal.Length



```
par(mfrow=c(1,3))
hist(irisVer$Petal.Length,breaks=seq(0,8,l=17),xlim=c(0,8),ylim=c(0,40))
hist(irisSet$Petal.Length,breaks=seq(0,8,l=17),xlim=c(0,8),ylim=c(0,40))
hist(irisVir$Petal.Length,breaks=seq(0,8,l=17),xlim=c(0,8),ylim=c(0,40))
```

Histogram of irisVer\$Petal.Length Histogram of irisSet\$Petal.Length Histogram of irisVir\$Petal.Length



```
library(beanplot)
```

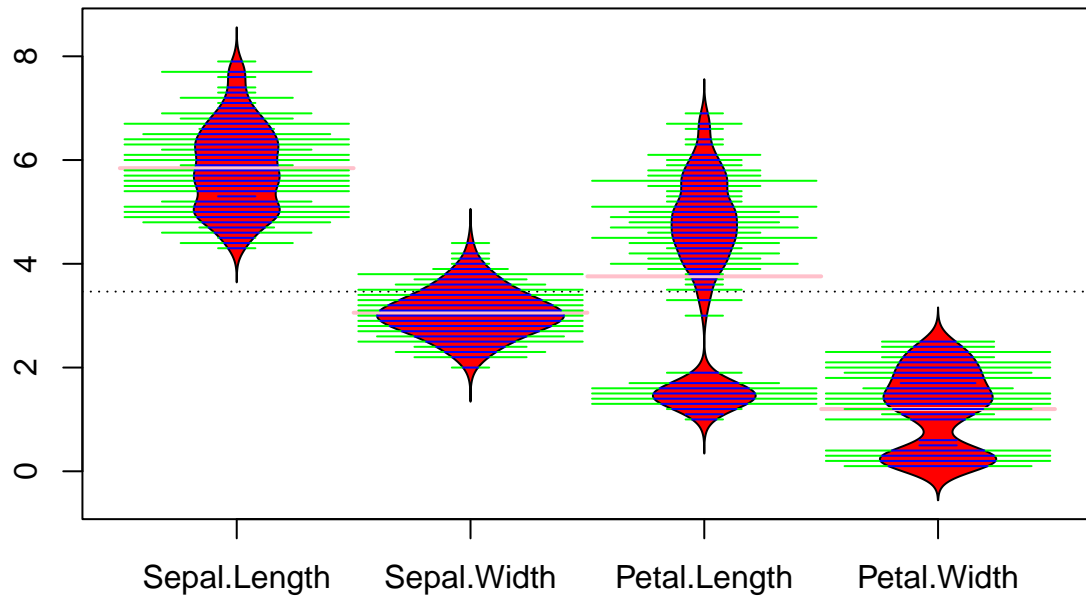
```
## Warning: package 'beanplot' was built under R version 4.3.3
```

```
xiris <- iris
```

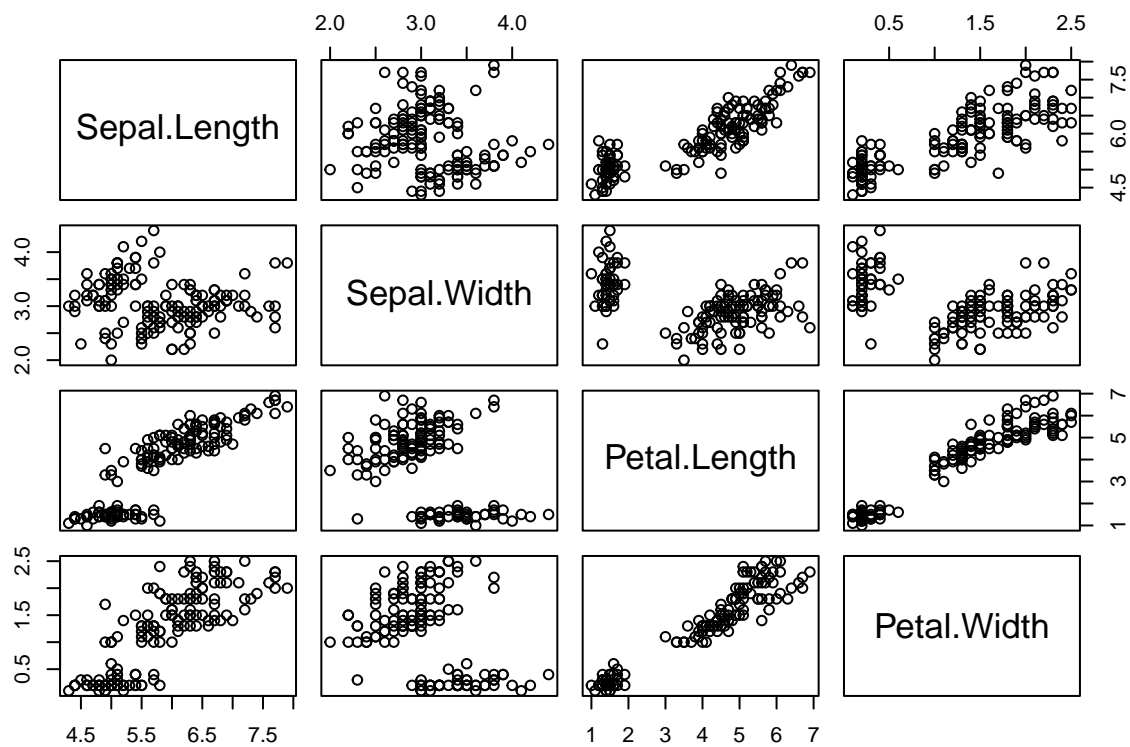
```
xiris$Species <- NULL
```

```
beanplot(xiris, main = "Iris flowers", col=c('red','blue','green','pink'), border = "black")
```

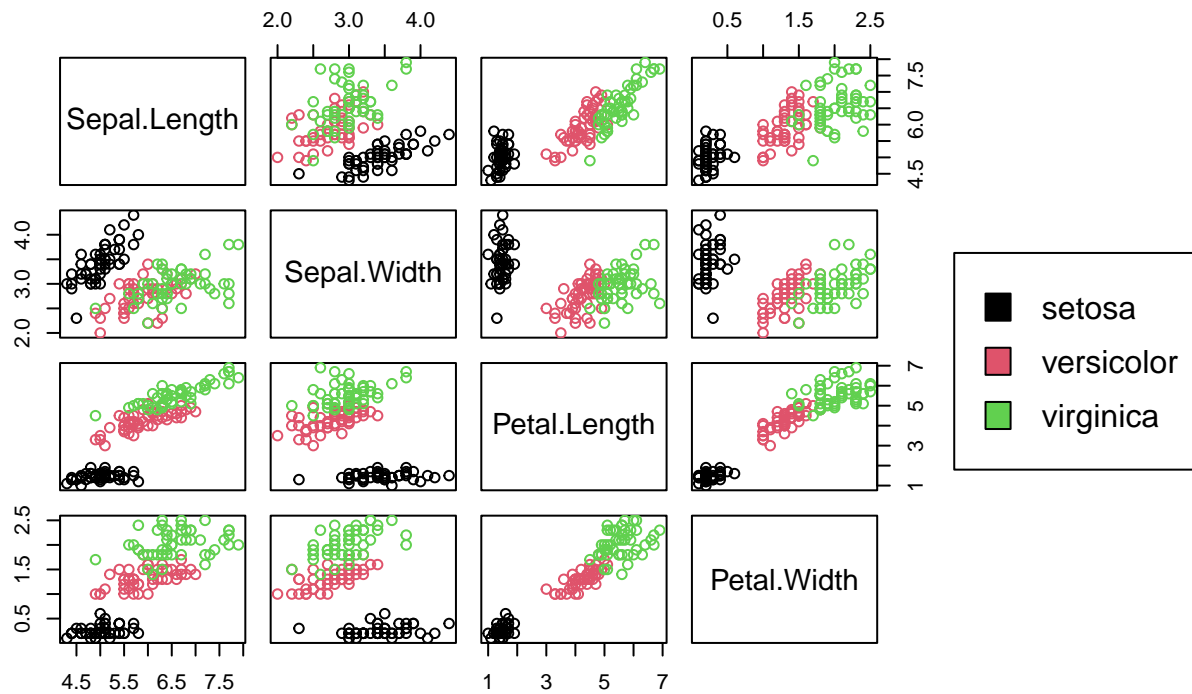
Iris flowers



```
pairs(iris[,1:4])
```



```
pairs(iris[,1:4],col=iris[,5],oma=c(4,4,6,12))
par(xpd=TRUE)
legend(0.85,0.6, as.vector(unique(iris$Species)),fill=c(1,2,3))
```



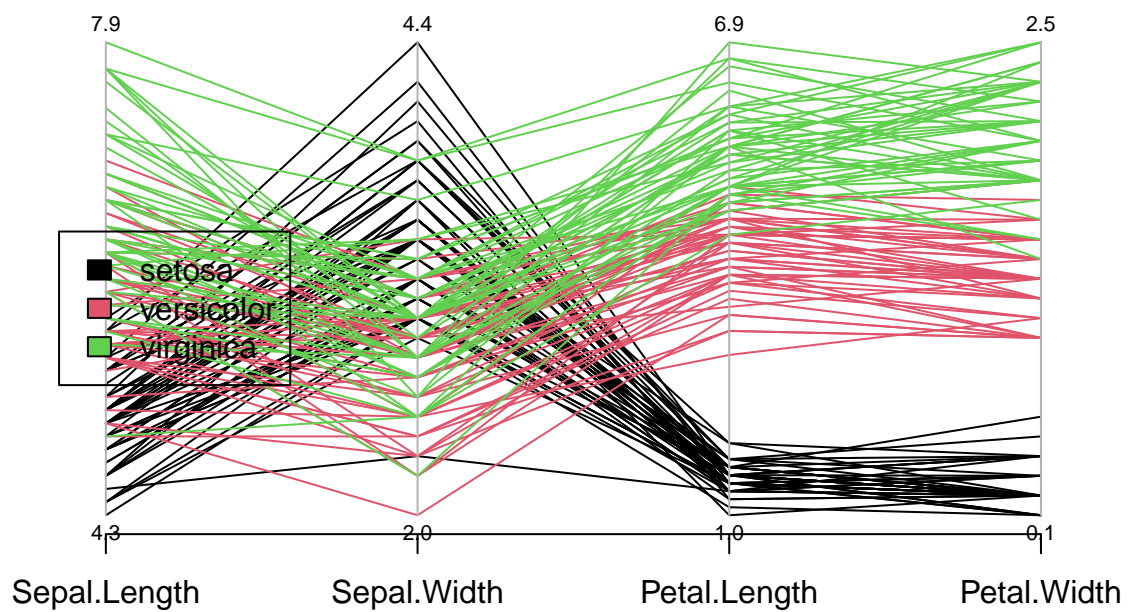
```
library(MASS)
```

```
## Warning: package 'MASS' was built under R version 4.3.3
```

```
parcoord(iris[,1:4], col=iris[,5],var.label=TRUE,oma=c(4,4,6,12))
```

```
par(xpd=TRUE)
```

```
legend(0.85,0.6, as.vector(unique(iris$Species)),fill=c(1,2,3))
```

```
library(C50)
```

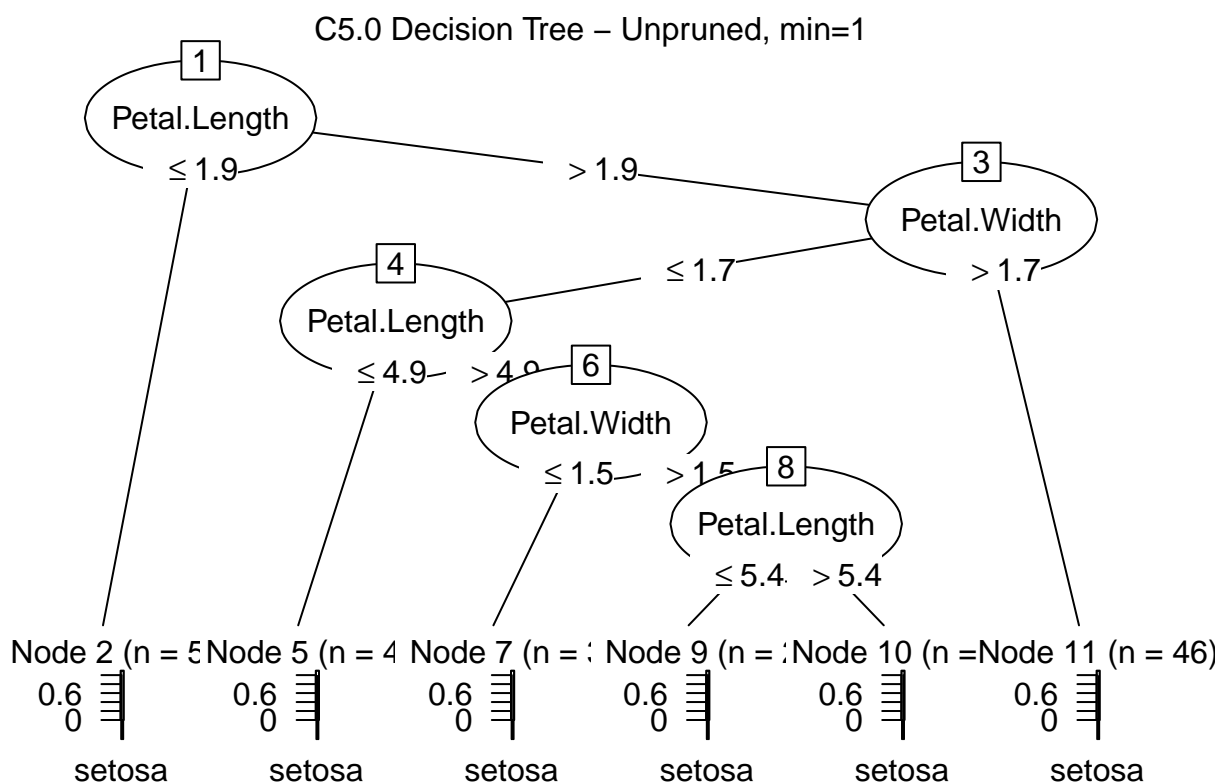
```
## Warning: package 'C50' was built under R version 4.3.3
```

```
input <- iris[,1:4]
```

```
output <- iris[,5]
```

```
model1 <- C5.0(input, output, control = C5.0Control(noGlobalPruning = TRUE, minCases=1))
```

```
plot(model1, main="C5.0 Decision Tree - Unpruned, min=1")
```

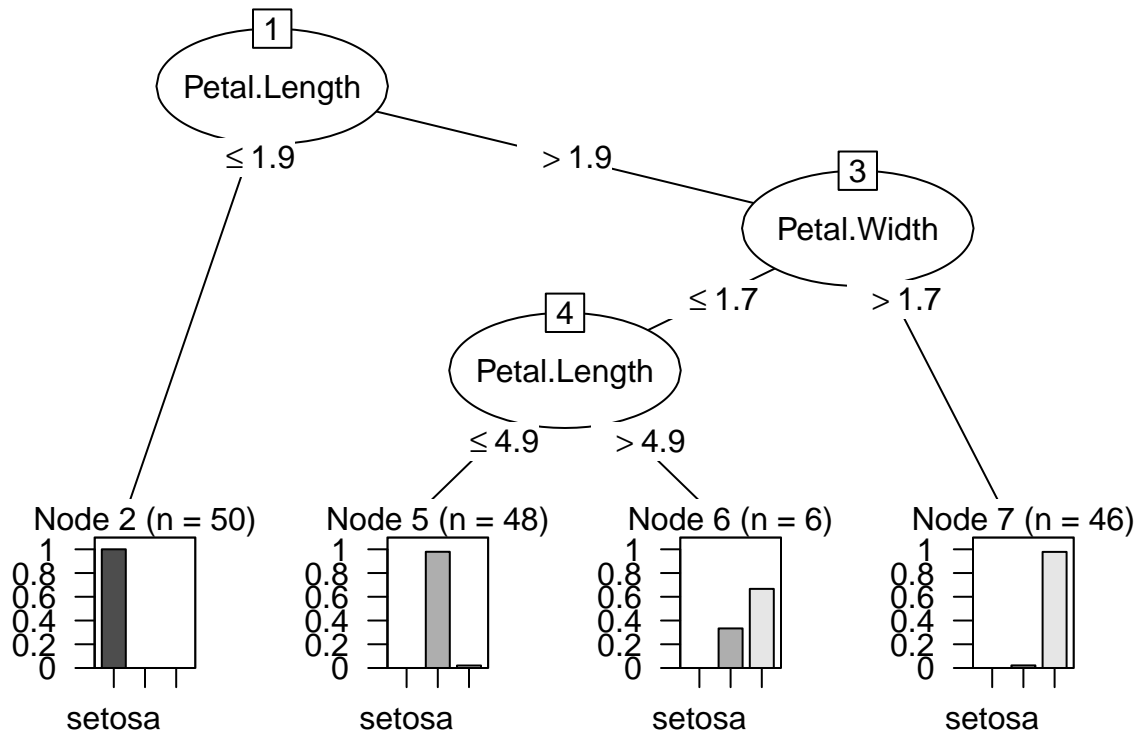


```

model2 <- C5.0(input, output, control = C5.0Control(noGlobalPruning = FALSE))
plot(model2, main="C5.0 Decision Tree - Pruned")

```

C5.0 Decision Tree – Pruned



```
summary(model2)
```

```
##
## Call:
## C5.0.default(x = input, y = output, control = C5.0Control(noGlobalPruning
## = FALSE))
##
##
## C5.0 [Release 2.07 GPL Edition]      Tue Apr  2 22:08:28 2024
## -----
##
## Class specified by attribute `outcome'
##
## Read 150 cases (5 attributes) from undefined.data
##
## Decision tree:
##
## Petal.Length <= 1.9: setosa (50)
## Petal.Length > 1.9:
##   ...Petal.Width > 1.7: virginica (46/1)
##   Petal.Width <= 1.7:
##     ...Petal.Length <= 4.9: versicolor (48/1)
##     Petal.Length > 4.9: virginica (6/2)
##
##
## Evaluation on training data (150 cases):
```

```
##
##      Decision Tree
##      -----
##      Size      Errors
##
##      4      4( 2.7%)  <<
##
##
##      (a)  (b)  (c)  <-classified as
##      ----  ----  ----
##      50                (a): class setosa
##                47      3  (b): class versicolor
##                1      49  (c): class virginica
##
##
## Attribute usage:
##
## 100.00% Petal.Length
## 66.67% Petal.Width
##
##
## Time: 0.0 secs
C5imp(model2,metric='usage')

##              Overall
## Petal.Length 100.00
## Petal.Width  66.67
## Sepal.Length  0.00
## Sepal.Width   0.00

newcases <- iris[c(1:3,51:53,101:103),]
newcases

##      Sepal.Length Sepal.Width Petal.Length Petal.Width  Species
## 1              5.1          3.5          1.4          0.2    setosa
## 2              4.9          3.0          1.4          0.2    setosa
## 3              4.7          3.2          1.3          0.2    setosa
## 51             7.0          3.2          4.7          1.4 versicolor
## 52             6.4          3.2          4.5          1.5 versicolor
## 53             6.9          3.1          4.9          1.5 versicolor
## 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

predicted <- predict(model2, newcases, type="class")
predicted

## [1] setosa  setosa  setosa  versicolor versicolor versicolor virginica
## [8] virginica virginica
## Levels: setosa versicolor virginica

predicted <- predict(model2, iris, type="class")
predicted

## [1] setosa  setosa  setosa  setosa  setosa  setosa
## [7] setosa  setosa  setosa  setosa  setosa  setosa
```

```
## [13] setosa      setosa      setosa      setosa      setosa      setosa
## [19] setosa      setosa      setosa      setosa      setosa      setosa
## [25] setosa      setosa      setosa      setosa      setosa      setosa
## [31] setosa      setosa      setosa      setosa      setosa      setosa
## [37] setosa      setosa      setosa      setosa      setosa      setosa
## [43] setosa      setosa      setosa      setosa      setosa      setosa
## [49] setosa      setosa      versicolor  versicolor  versicolor  versicolor
## [55] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [61] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [67] versicolor  versicolor  versicolor  versicolor  virginica   versicolor
## [73] versicolor  versicolor  versicolor  versicolor  versicolor  virginica
## [79] versicolor  versicolor  versicolor  versicolor  versicolor  virginica
## [85] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [91] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [97] versicolor  versicolor  versicolor  versicolor  virginica   virginica
## [103] virginica   virginica   virginica   virginica   versicolor  virginica
## [109] virginica   virginica   virginica   virginica   virginica   virginica
## [115] virginica   virginica   virginica   virginica   virginica   virginica
## [121] virginica   virginica   virginica   virginica   virginica   virginica
## [127] virginica   virginica   virginica   virginica   virginica   virginica
## [133] virginica   virginica   virginica   virginica   virginica   virginica
## [139] virginica   virginica   virginica   virginica   virginica   virginica
## [145] virginica   virginica   virginica   virginica   virginica   virginica
## Levels: setosa versicolor virginica

iris$predictedC501 <- predicted

iris[iris$Species != iris$predictedC501,]

##      Sepal.Length Sepal.Width Petal.Length Petal.Width   Species predictedC501
## 71           5.9         3.2         4.8         1.8 versicolor   virginica
## 78           6.7         3.0         5.0         1.7 versicolor   virginica
## 84           6.0         2.7         5.1         1.6 versicolor   virginica
## 107          4.9         2.5         4.5         1.7  virginica   versicolor
```