Edgar Anderson's Iris Data

Contents

Description

This famous (Fisher's or Anderson's) iris data set gives the measurements in centimeters of the variables sepal length and width and petal length and width, respectively, for 50 flowers from each of 3 species of iris. The species are *Iris setosa*, *versicolor*, and *virginica*.

Usage

```
iris
```

Format

`iris` is a data frame with 150 cases (rows) and 5 variables (columns) named:

- `Sepal.Length`
- `Sepal.Width`
- `Petal.Length`
- `Petal.Width`
- `Species`

Source

Anderson, Edgar (1935). "The irises of the Gaspe Peninsula." *Bulletin of the American Iris Society*, **59**: 2–5.
Fisher, Ronald A. (1936). "The use of multiple measurements in taxonomic problems." *Annals of Eugenics*, **7** (Part II): 179–188.

Examples

We investigate the Sepal and Petal leaves for the three species in the Iris data:

```
summary(iris)
```

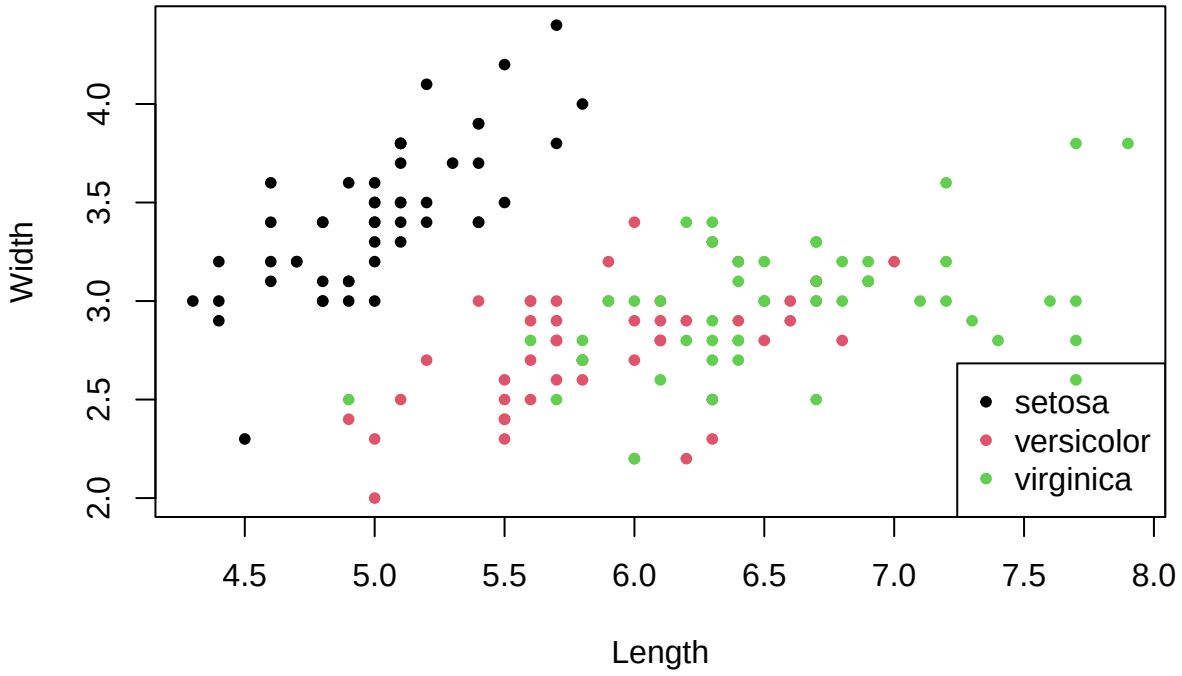
```
##   Sepal.Length    Sepal.Width     Petal.Length    Petal.Width          Species
##  Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100   setosa    :50
##  1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300   versicolor:50
##  Median :5.800   Median :3.000   Median :4.350   Median :1.300   virginica :50
##  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
```

To examine the Sepal leaves, we select the length and the width:

```
llen <- iris$Sepal.Length
lwid <- iris$Sepal.Width
```

Then we plot the data:

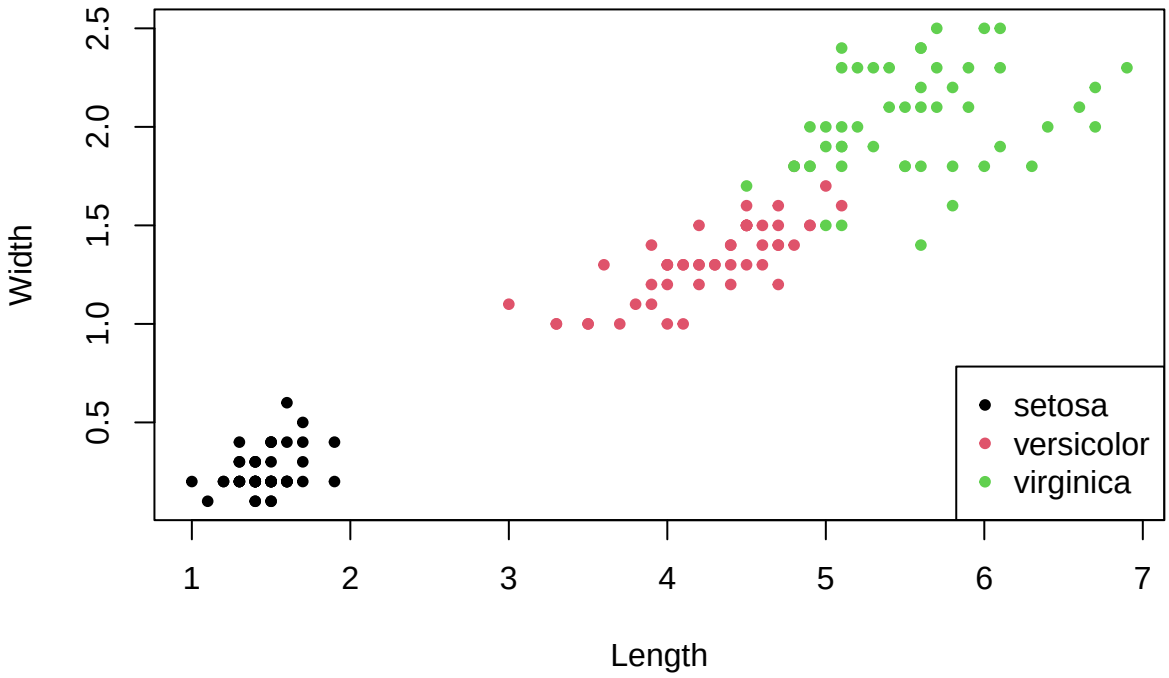
```
plot(llen, lwid, xlab = "Length", ylab = "Width",
     pch = 20, col = as.numeric(iris$Species))
legend("bottomright", legend = levels(iris$Species), col = 1:3, pch = 20)
```



We can also select the Petal leaves:

```
llen <- iris$Petal.Length
lwid <- iris$Petal.Width
```

This gives us the following plot:


Width
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
7
Length
setosa
versicolor
virginica