

Revision

Fulya Gökalg Yavuz

Spring 2024

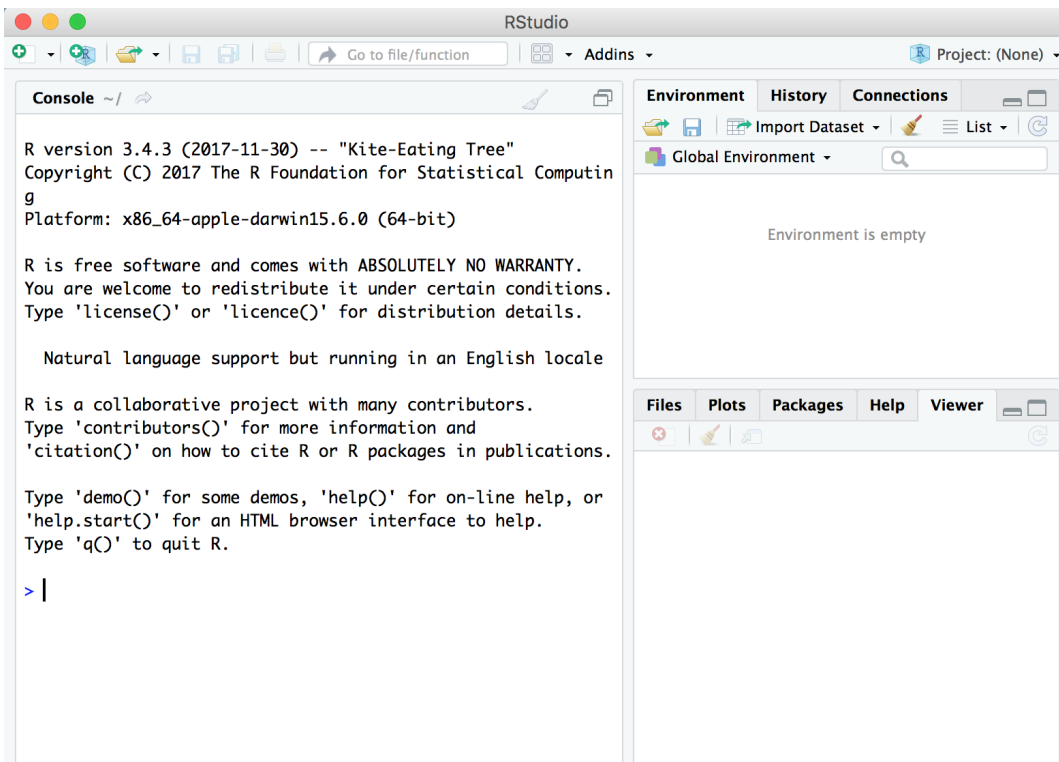
R and RStudio

To be able to use RStudio, please download R from the following link:

<http://cran.r-project.org>.

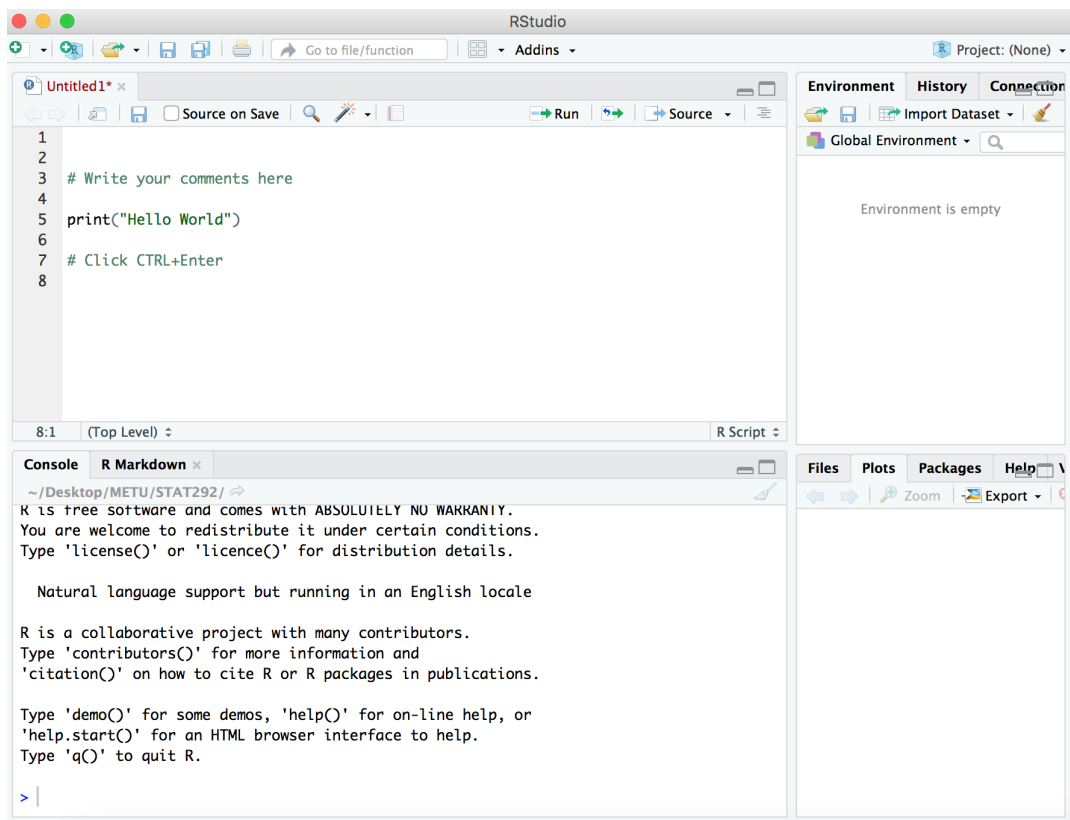
RStudio

<https://www.rstudio.com>.



You may want to write your commands directly on console.

However, it is easier to open a new console and write your codes there.



You are ready to go in R!

Type your code and hit Enter

```
1+1
```

```
## [1] 2
```

Notice the numbers in brackets

```
1:100
```

```
## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
## [19] 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
## [37] 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54
## [55] 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72
## [73] 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
## [91] 91 92 93 94 95 96 97 98 99 100
```

```
2.35:50
```

```
## [1] 2.35 3.35 4.35 5.35 6.35 7.35 8.35 9.35 10.35 11.35 12.35 13.35
## [13] 14.35 15.35 16.35 17.35 18.35 19.35 20.35 21.35 22.35 23.35 24.35 25.35
## [25] 26.35 27.35 28.35 29.35 30.35 31.35 32.35 33.35 34.35 35.35 36.35 37.35
## [37] 38.35 39.35 40.35 41.35 42.35 43.35 44.35 45.35 46.35 47.35 48.35 49.35
```

R tells you when it does not interpret your code

```
1::100
```

```
## Error: <text>:1:2: unexpected '::'
## 1: 1::
##      ^
```

Use R as a calculator

```
1 + 5
```

```
## [1] 6
```

```
3 * 5
```

```
## [1] 15
```

```
100 / 5
```

```
## [1] 20
```

!.. R does not run any line starting with a hashtag “#”. Use it to comment in your code chunk.

Lets have a virtual dice

```
1:6
```

```
## [1] 1 2 3 4 5 6
```

Store dice vector inside an R object

```
mydice <- 1:6
```

```
mydice
```

```
## [1] 1 2 3 4 5 6
```

Notice that the object stored in the environment pane on the upper right.

An object name cannot start with a number or special symbols such as ^,!,\$,@,+,-,/, or*:

R is case-sensitive, so please be careful with object names

```
aaa <- 555
```

```
Aaa <- 666
```

```
aaa
```

```
## [1] 555
```

```
Aaa
```

```
## [1] 666
```

R overwrite

```
aaa
```

```
## [1] 555
```

```
aaa <- 777
```

```
aaa
```

```
## [1] 777
```

Which objects we used so far?

```
ls()
```

```
## [1] "aaa"    "Aaa"    "mydice"
```

Remember that we have a dice

```
mydice
```

```
## [1] 1 2 3 4 5 6
```

Some math with our dice

```
mydice - 1
```

```
## [1] 0 1 2 3 4 5
```

```
mydice * mydice
```

```
## [1] 1 4 9 16 25 36
```

Inner and outer multiplication

```
mydice %*% mydice
```

```
##      [,1]
```

```
## [1,] 91
```

```
mydice %o% mydice
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
```

```
## [1,] 1 2 3 4 5 6
```

```
## [2,] 2 4 6 8 10 12
```

```
## [3,] 3 6 9 12 15 18
```

```
## [4,] 4 8 12 16 20 24
```

```
## [5,] 5 10 15 20 25 30
```

```
## [6,] 6 12 18 24 30 36
```

Convert a vector into n-dimensional matrix

```
dim(mydice) <- c(2,3)
```

```
mydice
```

```
##      [,1] [,2] [,3]
```

```
## [1,] 1 3 5
```

```
## [2,] 2 4 6
```

Convert a vector into arrays

```
dim(mydice) <- c(1, 2, 3)
```

```
mydice
```

```
## , , 1
```

```
##
```

```
##      [,1] [,2]
```

```
## [1,] 1 2
```

```
##
```

```
## , , 2
```

```
##
```

```
##      [,1] [,2]
```

```
## [1,] 3 4
```

```
##
```

```
## , , 3
```

```
##
```

```
##      [,1] [,2]
```

```
## [1,] 5 6
```

Another command to generate matrix

```
m <- matrix(mydice, nrow = 2)
```

```
m
```

```
##      [,1] [,2] [,3]
```

```
## [1,] 1 3 5
## [2,] 2 4 6
```

```
m <- matrix(mydice, nrow = 2, byrow = TRUE)
m
```

```
##      [,1] [,2] [,3]
## [1,] 1 2 3
## [2,] 4 5 6
```

Another command to generate arrays

```
myarray <- array(c(1:12), dim = c(2, 2, 3))
myarray
```

```
## , , 1
##
##      [,1] [,2]
## [1,] 1 3
## [2,] 2 4
##
## , , 2
##
##      [,1] [,2]
## [1,] 5 7
## [2,] 6 8
##
## , , 3
##
##      [,1] [,2]
## [1,] 9 11
## [2,] 10 12
```

Time and date

Computer time

```
now <- Sys.time()
now
```

```
## [1] "2024-02-19 11:09:22 +03"
```

```
class(now)
```

```
## [1] "POSIXct" "POSIXt"
```

POSIXct shows the seconds between now and 1 Jan 1970 (in the Universal Time Coordinated (UTC) zone).

```
unclass(now)
```

```
## [1] 1708330162
```

What was the date and time 1000000 seconds ago?

```
mil <- 1000000
now - mil
```

```
## [1] "2024-02-07 21:22:42 +03"
```

Remember one of the most original! data in R:

```

data(iris)
?iris
head(iris)

##   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
## 6          5.4          3.9          1.7          0.4  setosa

class(iris)

## [1] "data.frame"

class(iris$Sepal.Length)

## [1] "numeric"

class(iris$Species)

## [1] "factor"

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 ...

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
##
##
##

iris$Sepal.Length

##   [1] 5.1 4.9 4.7 4.6 5.0 5.4 4.6 5.0 4.4 4.9 5.4 4.8 4.8 4.3 5.8 5.7 5.4 5.1
##  [19] 5.7 5.1 5.4 5.1 4.6 5.1 4.8 5.0 5.0 5.2 5.2 4.7 4.8 5.4 5.2 5.5 4.9 5.0
##  [37] 5.5 4.9 4.4 5.1 5.0 4.5 4.4 5.0 5.1 4.8 5.1 4.6 5.3 5.0 7.0 6.4 6.9 5.5
##  [55] 6.5 5.7 6.3 4.9 6.6 5.2 5.0 5.9 6.0 6.1 5.6 6.7 5.6 5.8 6.2 5.6 5.9 6.1
##  [73] 6.3 6.1 6.4 6.6 6.8 6.7 6.0 5.7 5.5 5.5 5.8 6.0 5.4 6.0 6.7 6.3 5.6 5.5
##  [91] 5.5 6.1 5.8 5.0 5.6 5.7 5.7 6.2 5.1 5.7 6.3 5.8 7.1 6.3 6.5 7.6 4.9 7.3

```

```
## [109] 6.7 7.2 6.5 6.4 6.8 5.7 5.8 6.4 6.5 7.7 7.7 6.0 6.9 5.6 7.7 6.3 6.7 7.2
## [127] 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.8
## [145] 6.7 6.7 6.3 6.5 6.2 5.9
```

Indexing

Select the first row:

```
iris[1,]
```

```
## Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1          3.5          1.4          0.2 setosa
```

Select the first column:

```
iris[,2]
```

```
## [1] 3.5 3.0 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 3.7 3.4 3.0 3.0 4.0 4.4 3.9 3.5
## [19] 3.8 3.8 3.4 3.7 3.6 3.3 3.4 3.0 3.4 3.5 3.4 3.2 3.1 3.4 4.1 4.2 3.1 3.2
## [37] 3.5 3.6 3.0 3.4 3.5 2.3 3.2 3.5 3.8 3.0 3.8 3.2 3.7 3.3 3.2 3.2 3.1 2.3
## [55] 2.8 2.8 3.3 2.4 2.9 2.7 2.0 3.0 2.2 2.9 2.9 3.1 3.0 2.7 2.2 2.5 3.2 2.8
## [73] 2.5 2.8 2.9 3.0 2.8 3.0 2.9 2.6 2.4 2.4 2.7 2.7 3.0 3.4 3.1 2.3 3.0 2.5
## [91] 2.6 3.0 2.6 2.3 2.7 3.0 2.9 2.9 2.5 2.8 3.3 2.7 3.0 2.9 3.0 3.0 2.5 2.9
## [109] 2.5 3.6 3.2 2.7 3.0 2.5 2.8 3.2 3.0 3.8 2.6 2.2 3.2 2.8 2.8 2.7 3.3 3.2
## [127] 2.8 3.0 2.8 3.0 2.8 3.8 2.8 2.8 2.6 3.0 3.4 3.1 3.0 3.1 3.1 3.1 2.7 3.2
## [145] 3.3 3.0 2.5 3.0 3.4 3.0
```

Select the column by name:

```
iris[, "Species"]
```

```
## [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 versicolor versicolor
## [73] versicolor versicolor versicolor versicolor versicolor versicolor
## [79] versicolor versicolor versicolor versicolor versicolor versicolor
## [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 virginica 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
```

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

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

Logical Indexing

```
LogicIndex <- iris[, "Petal.Length"] > 5.5
LogicIndex
```

```
##      [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##     [97] FALSE FALSE FALSE FALSE  TRUE  FALSE  TRUE  TRUE  TRUE  TRUE  FALSE  TRUE
##    [109]  TRUE  TRUE  FALSE  FALSE  FALSE  FALSE  FALSE  FALSE  FALSE  TRUE  TRUE  FALSE
##    [121]  TRUE  FALSE  TRUE  FALSE  TRUE  TRUE  FALSE  FALSE  TRUE  TRUE  TRUE  TRUE
##    [133]  TRUE  FALSE  TRUE  TRUE  TRUE  FALSE  FALSE  FALSE  TRUE  FALSE  FALSE  TRUE
##    [145]  TRUE  FALSE  FALSE  FALSE  FALSE  FALSE  FALSE
```

```
iris[LogicIndex, ]
```

```
##      Sepal.Length Sepal.Width Petal.Length Petal.Width  Species
## 101          6.3         3.3         6.0         2.5 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
## 108          7.3         2.9         6.3         1.8 virginica
## 109          6.7         2.5         5.8         1.8 virginica
## 110          7.2         3.6         6.1         2.5 virginica
## 118          7.7         3.8         6.7         2.2 virginica
## 119          7.7         2.6         6.9         2.3 virginica
## 121          6.9         3.2         5.7         2.3 virginica
## 123          7.7         2.8         6.7         2.0 virginica
## 125          6.7         3.3         5.7         2.1 virginica
## 126          7.2         3.2         6.0         1.8 virginica
## 129          6.4         2.8         5.6         2.1 virginica
## 130          7.2         3.0         5.8         1.6 virginica
## 131          7.4         2.8         6.1         1.9 virginica
## 132          7.9         3.8         6.4         2.0 virginica
## 133          6.4         2.8         5.6         2.2 virginica
## 135          6.1         2.6         5.6         1.4 virginica
## 136          7.7         3.0         6.1         2.3 virginica
## 137          6.3         3.4         5.6         2.4 virginica
## 141          6.7         3.1         5.6         2.4 virginica
## 144          6.8         3.2         5.9         2.3 virginica
## 145          6.7         3.3         5.7         2.5 virginica
```


Summary Statistics

```
mean(iris$Sepal.Length)
```

```
## [1] 5.843333
```

```
sd(iris$Sepal.Length)
```

```
## [1] 0.8280661
```

Functions

Let's say we want to find mean and standard deviation at the same time:

```
MeanAndStd <- function(x) {  
  c(mean=mean(x), std=sd(x))  
}
```

```
MeanAndStd(iris$Sepal.Length)
```

```
##      mean      std
```

```
## 5.8433333 0.8280661
```

```
Stats <- aggregate(Sepal.Length ~ Species, data=iris,  
                   FUN=MeanAndStd)
```

```
Stats
```

```
##      Species Sepal.Length.mean Sepal.Length.std
```

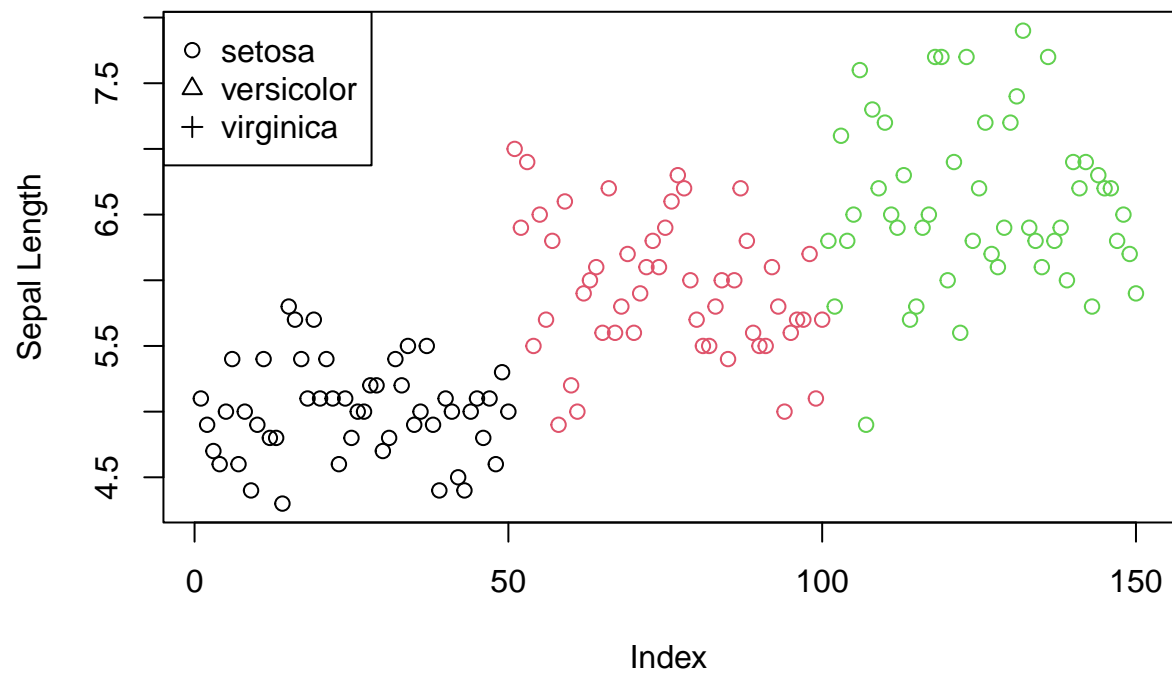
```
## 1      setosa      5.0060000      0.3524897
```

```
## 2 versicolor      5.9360000      0.5161711
```

```
## 3  virginica      6.5880000      0.6358796
```

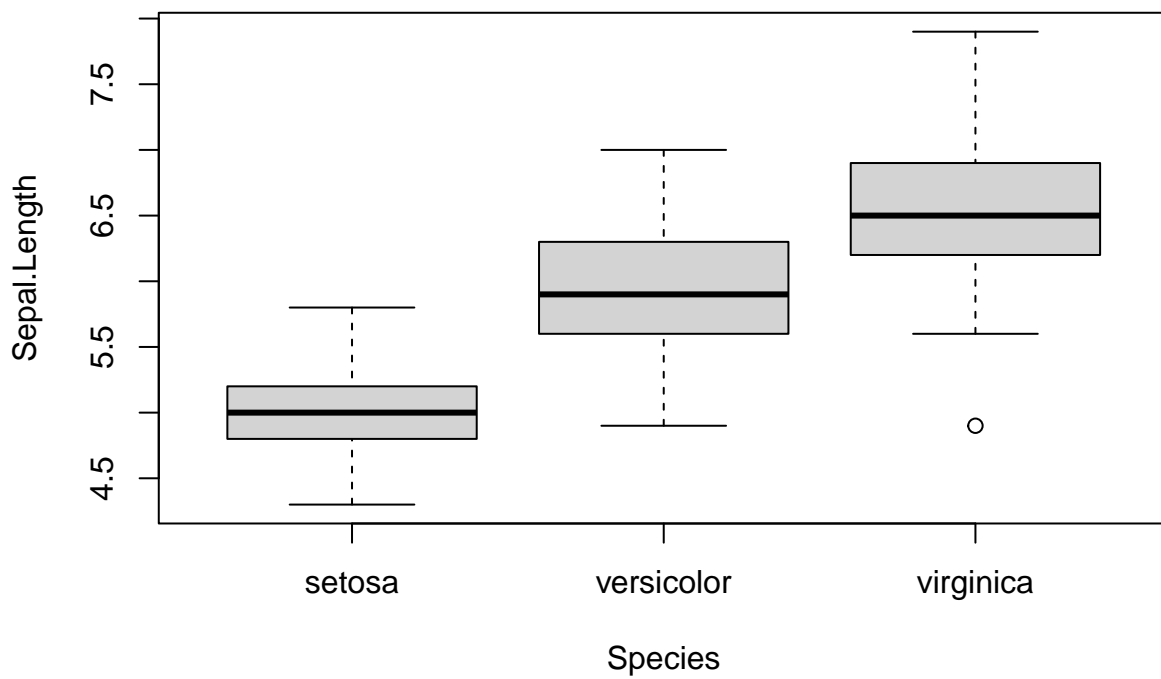
Visualization

```
plot(iris$Sepal.Length,  
     col=as.numeric(iris$Species),  
     ylab="Sepal Length")  
legend('topleft', legend = levels(iris$Species), pch = 1:3)
```



Of course the scatter plot is not the best option in here. For a factor, boxplot is always a better option:

```
boxplot(Sepal.Length ~ Species, data=iris)
```



FGY.