



R programming

Club Bioinfo - Institut Jacques Monod

Leslie REGAD and Gaëlle LELANDAIS

Mails :

leslie.regad@univ-paris-diderot.fr ;

gaelle.lelandais@univ-paris-diderot.fr



Handling R Objects

Section 1

Vectors

3

- ◉ Example

```
> vect = 12:28  
[1] 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
```

- ◉ Retrieving values from a vector
 - > Symbols « [] »

```
> vect[2]      # the 2nd element of the vector is displayed  
[1] 13
```

Vectors

- Getting multiple values from a vector
 - Consecutive elements

```
> vect[c(5, 6, 7, 8, 9)]  
[1] 16 17 18 19 20
```

```
> vect[5:9]  
[1] 16 17 18 19 20
```

- Non consecutive elements

```
> vect[c(5, 10, 13)]  
[1] 16 21 24
```

- Deleting values

```
> vect[-1]  
[1] 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28  
> vect[-5:-9]  
[1] 12 13 14 15 21 22 23 24 25 26 27 28
```

Tables

5

- Example :

```
> Mat = cbind(1:4, 5:8)
  [,1] [,2]
[1,]  1  5
[2,]  2  6
[3,]  3  7
[4,]  4  8
```

- Getting multiple values from a table
 - > Symbols « [row, column] »

```
> Mat[3, 2]           # one value is selected (row3, column2)
[1] 7
> Mat[1, ]           # all values, row 1
[1] 1 5
> Mat[c(1,3),]       # all values, rows 1 and 3
  [,1] [,2]
[1,]  1  5
[2,]  3  7
> Mat[,1]           # all values, column 1
```

Attribute Names for a Vector

6

- Principle
 - > Give a name (character string) to each element of a vector
 - > Allow to follow and retrieve elements from a vector more easily
- Function
 - > *names()*

```
> notes = c(12,15,8,9,11,15,5,20,17)
> names(notes) = c("Villon", "Polin", "Exfi", "Rotaf", "Zerif",
"Gared", "Neyres", "Ropert", "Saidil")
> notes
Villon Polin Exfi Rotaf Zerif Gared Neyres Ropert Saidil
    12    15    8    9    11    15    5    20    17
> notes[c("Exfi", "Gared")]
Exfi Gared
    8    15
```

Attribute Names for a Table

7

- Principle
 - > Give a name (character string) to each row and each column of a table
- Functions
 - > `row.names()`, `colnames()`

```
> Mat1
      [,1]      [,2]
[1,] -0.1177814 -0.7376553
[2,] -1.1422671 -0.4758635
> row.names(Mat1) = c("Ligne1", "Ligne2")
> colnames(Mat1) = c("Colonne1", "Colonne2")
> Mat1
      Colonne1  Colonne2
Ligne1 -0.1177814 -0.7376553
Ligne2 -1.1422671 -0.4758635
```

Identification of Elements Satisfying a Particular Condition

8

- ◉ Allow to retrieve (with a unique command line) all elements that satisfy a specified constrain
- ◉ Positions of the elements in the vector can be identified
 - > *which(cond)*

```
> vect = c(12,15,8,15,9,5,11,17,19,5,15,12,8)
> vect <= 10
[1] FALSE FALSE TRUE FALSE TRUE TRUE FALSE FALSE FALSE
TRUE FALSE FALSE TRUE
```

```
> vect = 100:110
> vect
[1] 100 101 102 103 104 105 106 107 108 109 110
> which(vect >= 105)
[1] 6 7 8 9 10 11
```



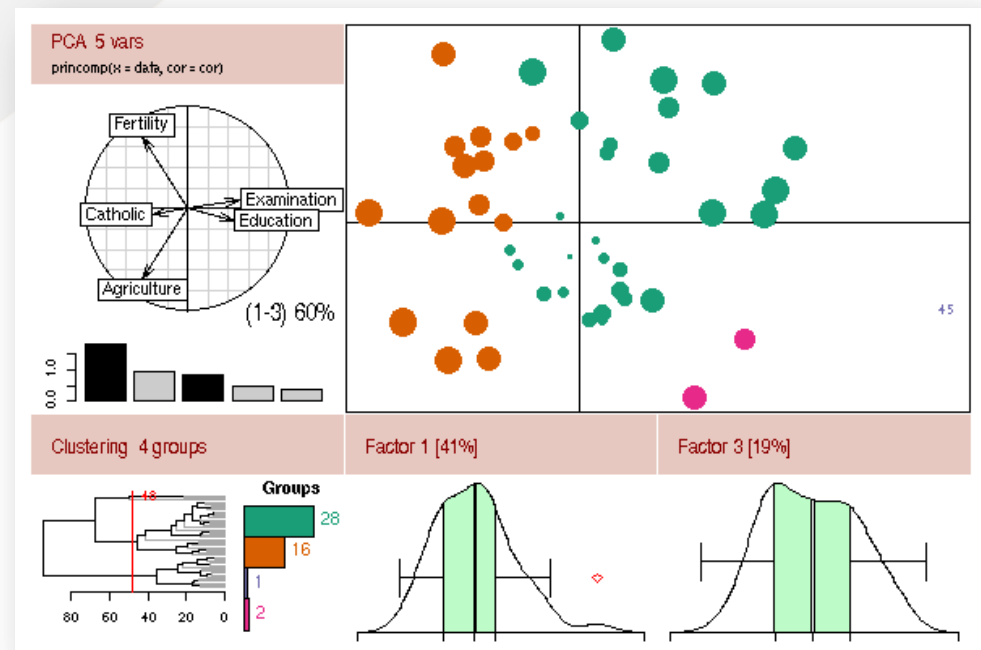
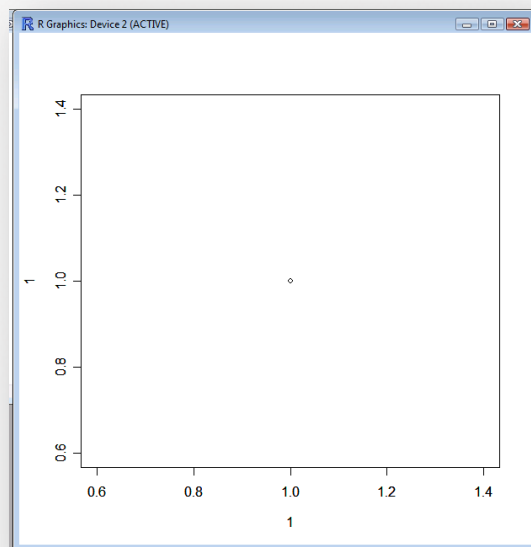

Graphical Representations of Data

Section 2

Graphics with R

10

- Many graphical representations can be drawn
 - > From the simplest to the most sophisticated

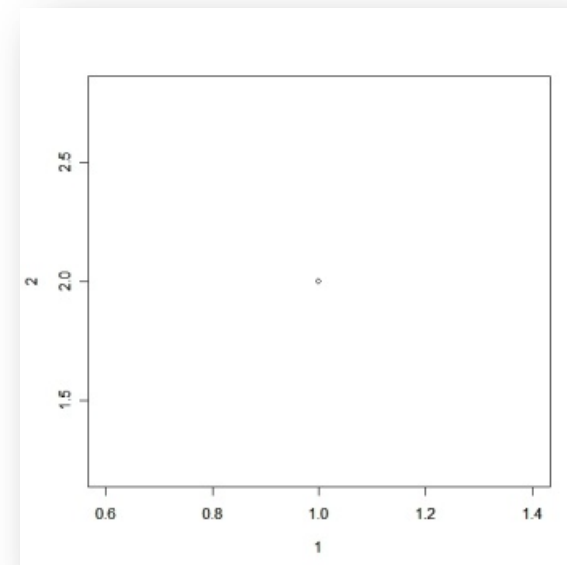


The Function « plot() »

11

- ⦿ Principle
 - > Generic function that allows to draw dots and lines according to X and Y coordinates
- ⦿ Example
 - > A single dot with $x = 1$ and $y = 2$

```
> plot(1,2)
```

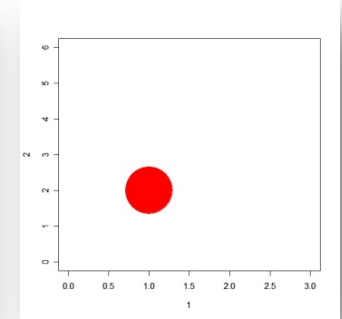
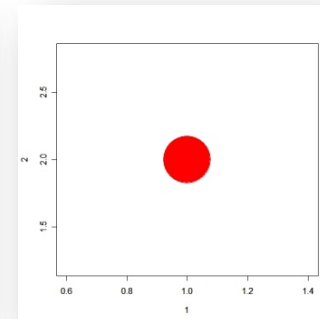
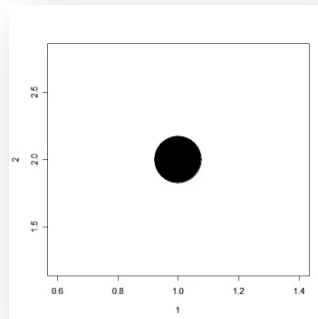
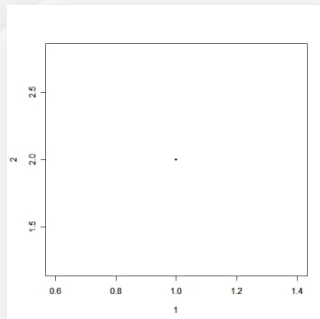
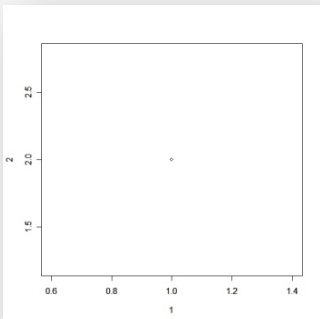


Parameters for the Function « plot() »

12

- Lots of parameters can be specified in order to:
 - > Modify the shape, the color and the size of the dots : *pch*, *col*, *cex*
 - > Modify axes : *xlim*, *ylim*, *axis*, *etc.*
 - > Add legends, titles and labels: *xlab*, *ylab*, *main*, *etc.*
 - > Etc...

```
> plot(1,2) ; plot(1,2, pch = 20) ; plot(1,2, pch = 20, cex = 20) ; plot(1,2, pch = 20, cex = 20, col = "red") ; plot(1,2, pch = 20, cex = 20, col = "red", xlim = c(0,3), ylim = c(0, 6))
```



The Function « hist() »

13

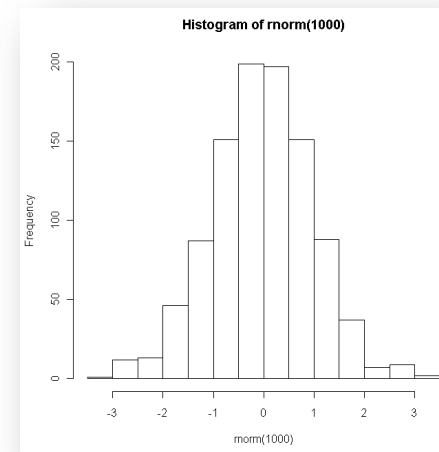
Principle

- > Draw the distribution of a random variable in sample
- > X axis : values for the random variable
- > Y axis : number of observations

Example

- > Histogram of 100 values (chosen according to a normal distribution)

```
> hist(rnorm(100))
```

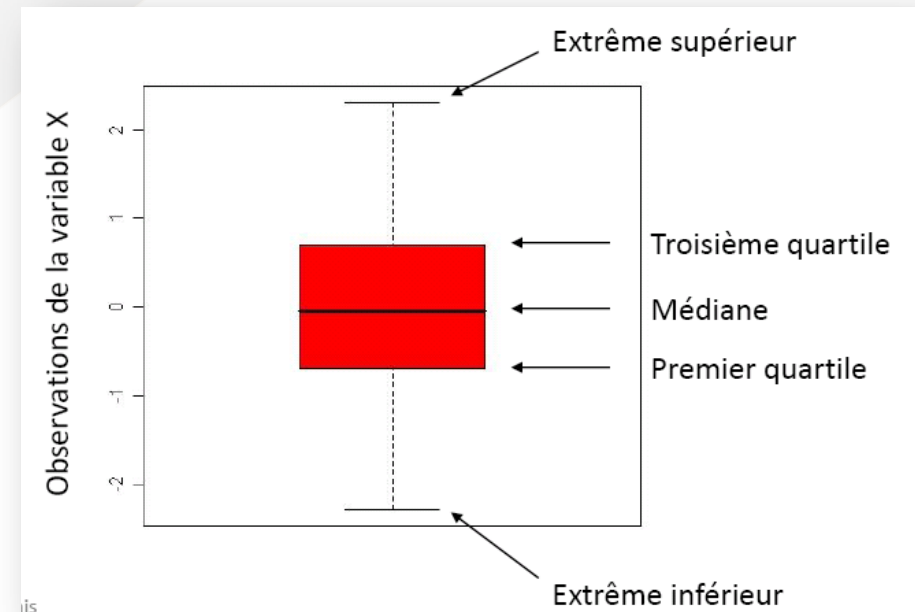


The Function « boxplot() »

14

- Principle
 - > Graphical summary of a set of numbers. Informative values are calculated and represented (median, quantiles, outliers)
- Example

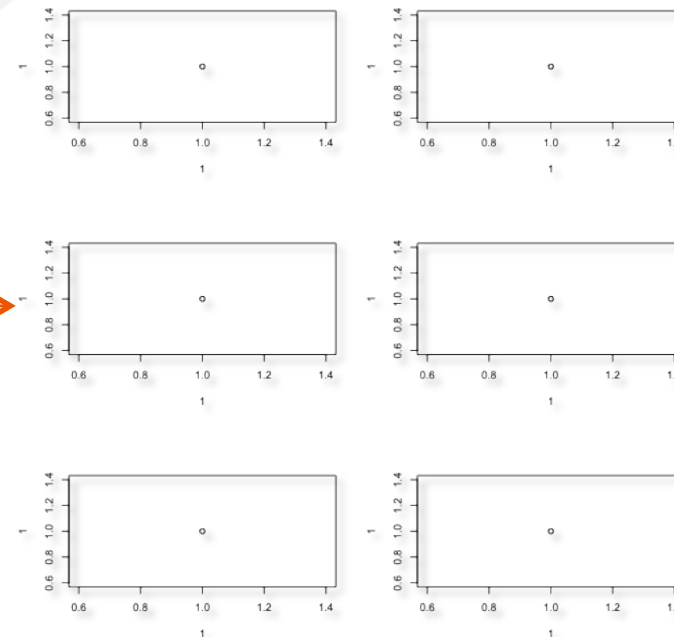
```
> boxplot(data)
```



Multiple Graphics on a Single Page

- ◉ Split the graphical window to represent simultaneously different graphs
 - > `par(mfrow=c(RowNumber, ColNumber))`

```
> par(mfrow=c(3,2))  
> plot(1,1)  
> plot(1,1)  
> plot(1,1)  
> plot(1,1)  
> plot(1,1)  
> plot(1,1)
```



Saving Graphics

16

- To create PDF files or image files
 - > `jpeg()`, `png()`, `bmp()`,
 - > `pdf()`
- To open / close of a graphical window
 - > `x11()`
 - > `dev.off()`
- To specify graphical window settings
 - > `par()`

```
> pdf("MonGraphique.pdf")  
> boxplot(data)  
> dev.off()
```




🎯 Practical session

(working with gene expression data)