

Introductory course on the R software

P. Lafaye de Micheaux¹

¹Mathematics and Statistics Departement
Université de Montréal
CANADA

New South Wales University
Sydney, Australia
May 6, 2014

<https://biostatisticien.eu/springeR/courseRw1.pdf>

Audience : who should attend this course ?

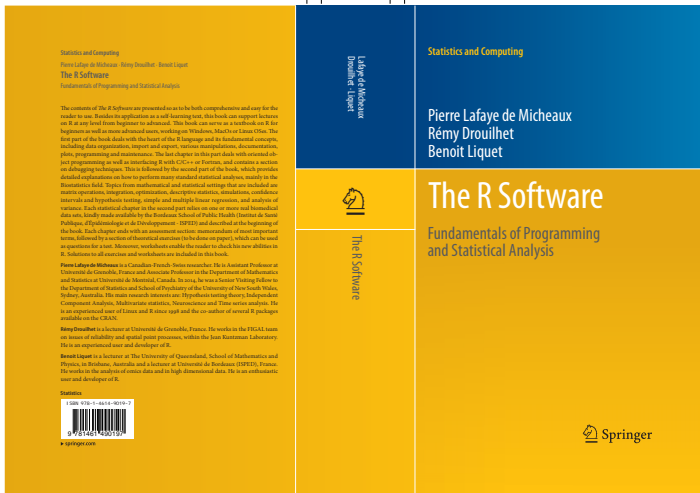
People who expect to :

- learn the basics of R (obviously !)
- want to perform sophisticated data manipulation or non-standard statistical analyses ;
- plan to use it on a regular basis in the future (if not, you will probably forget everything ...).

Level of difficulty

Matlab users (who can read and produce code) will learn something here ... but I will start with the very basics.

Reference for the course



Course format

- Some “theory” and a lot of practice !
- Each people attending the course should bring his/her own laptop.
- “Assignments” will be given for each future session.

Questions ?

Feel free to interrupt me at any time !

Schedule

Today (Week 1), we start with the very basics (Chapter 1 and beginning of Chapter 3).

Then we could plan weekly meetings to cover other chapters :

Week 2 End of Chapter 3 - Basic Concepts and Data Organisation .

Week 3 Chapter 4 -Importing, Exporting and Producing Data.

Week 4 Chapter 5 - Data Manipulation.

Week 5 Chapter 5 - Functions.

Week 6 Chapter 7- Drawing Curves and Plots.

Week 7 Part of Chapters 8 and 9 - Programming in R ; Managing Sessions.

Contact

Please, fill this blank page with your name and email.

Why use R ?

R is :

- free ;
- open-source ;
- cross-platform (Windows, MacOS, Linux, etc.) ;
- rapidly evolving (many packages added each day !).

Warning

R is harder to comprehend than other software on the market. You need to spend time learning the syntax and commands to become an efficient user.

Why use R ?

R is especially powerful for data manipulation, calculations and plots. Its features include :

- an integrated and very well conceived documentation system (in English) ;
- efficient procedures for data treatment and storage ;
- a suite of operators for calculations on tables, especially matrices (but also arrays) ;
- a vast and coherent collection of statistical procedures for data analysis ;
- advanced graphical capabilities ;
- a “simple” and efficient programming language, including conditioning, loops, recursion and input-output possibilities.


Installing R

Let's start by installing the software !


- 1 Download the file `R-x-win.exe` (where x is the number of the latest version) at the address :

`http://cran.r-project.org/bin/windows/base/`

- 2 Save this executable file on the Windows Desktop and

double-click the file `R-x-win.exe` (its icon is ).

- 3 The software then installs. Follow the instructions displayed on your screen and keep the default options.

- 4 When the icon  is added to the Desktop, installation is complete.

Installing R for MacOS or Linux

Download the file `R-x-snowleopard.pkg` (where `x` is the number of the latest version, i.e. 3.1.0) at the address : `http://cran.r-project.org/bin/macosx/`

I guess that if you use Linux, you will know how to do it !

R and Statistics

Many **classical** and **modern** statistical techniques are implemented in R. The most common methods for statistical analysis, such as :

- descriptive statistics ;
- hypothesis testing ;
- analysis of variance ;
- linear regression methods (simple and multiple) ;
- and so on

are directly included at the **core** of the system.

Third part of my book covers the following notions : basic mathematics, descriptive statistics, generation of random values, confidence intervals and hypothesis testing, simple and multiple linear regression, elementary analysis of variance.

Extending R

Most **advanced or recent statistical methods** are available through external **packages**, easy to install from R (see Section A2. from my Book).

They are all grouped and can be browsed on the website of the *Comprehensive R Archive Network* (CRAN) :

http://cran.r-project.org/web/packages/available_packages_by_name.html

See also the Task Views (on the CRAN) that group packages related to some domains of interest :

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

Official website of R : The Comprehensive R Archive Network

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

R and plots

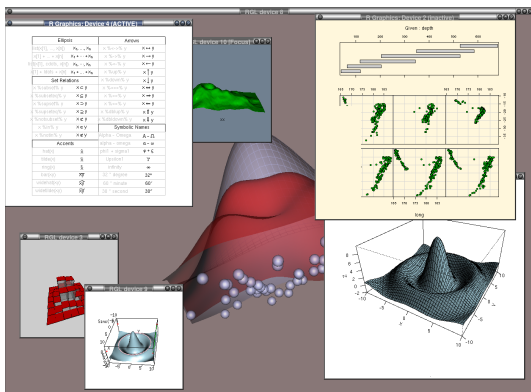


FIGURE : A few of the graphical possibilities offered by R.

R and other software

R can interact with the following software :

- use R from within Excel :
`http://rcom.univie.ac.at/download.html#RExcel`
- R from within SAS : `http://support.sas.com/rnd/app/studio/Rinterface2.html`
- R from within SPSS : `http://www.ibm.com/developerworks/library/ba-call-r-spss`
- R from Matlab : `http://www.mathworks.com/matlabcentral/fileexchange/5051-matlab-r-link`
- Matlab from R : install package `R.matlab`

R and plots

Let's play a little bit ! Try these commands :

```
demo(image)
example(contour)
demo(graphics)
demo(persp)
demo(plotmath)
demo(Hershey)
install.packages("lattice")
require("lattice") # Load package, previously installed.
demo(lattice)
example(wireframe)
install.packages("rgl")
require("rgl")
demo(rgl) # Interact using your mouse.
example(persp3d)
```

The R Graphical User Interface

The R Graphical User Interface (*i.e.* its set of menus) is very limited, and completely nonexistent on some operating systems, when compared to other standard software (SPSS say). This minimality can set back some new users. However, this drawback is limited since :

- it has the didactic advantage that it incites users to know well the statistical procedures they wish to use ;
- there are additional tools which extend the GUI

Please, download and install RStudio :

<https://www.rstudio.com/ide/download/desktop>

We can discuss quickly the RStudio frames.

Your first steps in R

Please, do by yourself the contents of Section 1.5 of my Book.

<http://biostatisticien.eu/springeR/Rbook-chap1.pdf>

I will stay around and answer your questions.

Warning

- R Commander (called using `CommandeR()`) is a tool mostly dedicated for beginners.
- R is case sensitive.

Thank you for your attention !