Installing R and RStudio

Bachelor in Computer Science and Engineering

2020/21

1. Statistical programming language R

During this course, we will use the open source programming language for statistical computing R, see `http://www.r-project.org/`. It is the most widely used programming language and software environment for statistics nowadays. The first R version was released in 2000 and, as of yet, the latest version is 4.0.2 *Taking Off Again* which has been released on 2020-06-22. It is available for Windows, MacOS, and UNIX platforms. You can find R in all the university computers, but it can be useful that you download it and install it either at your laptop or home computer.

Downloading R

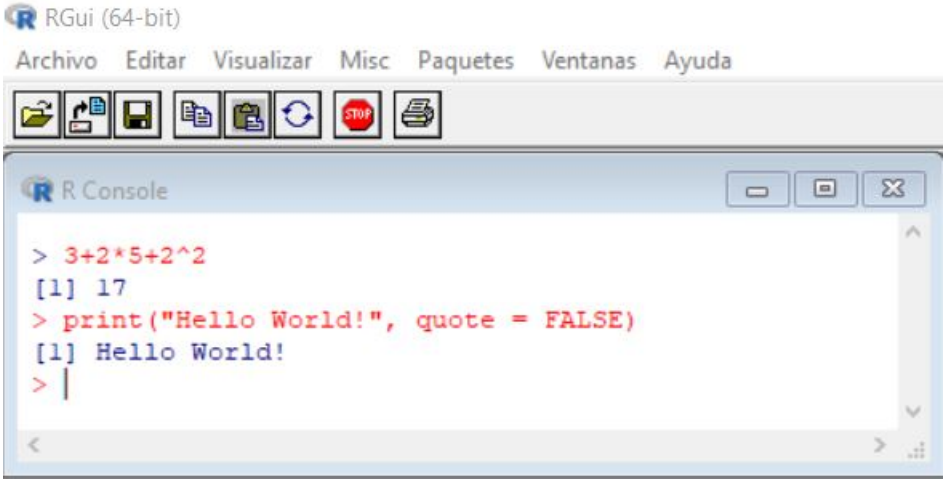
Enter the web page `http://www.r-project.org/` and click on the link `download R`. There you will find a list of mirrors to download R from. Select one located near your current location (in Spain) and download the precompiled binary distribution of R for your platform.

Installing R

By clicking on the executable file that you have just downloaded, you will install R on your computer. The installation process is very fast. You will only have to answer some questions about the basic configuration of the program (language, folder to allocate R, direct link...).

The R console

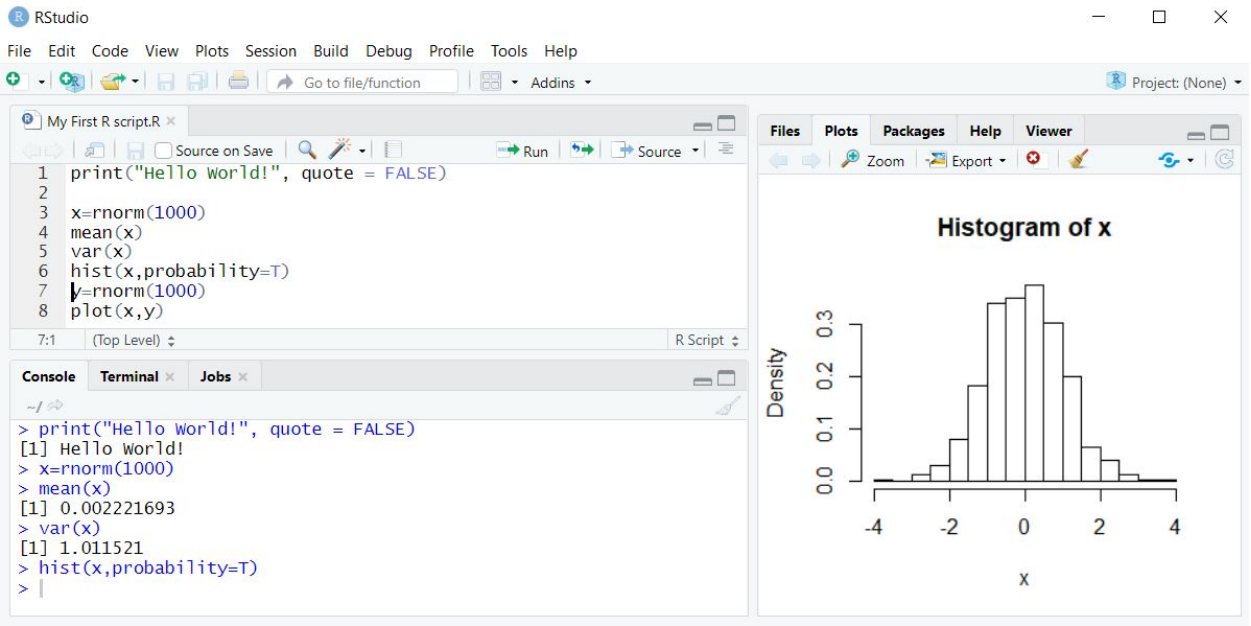
Once you have installed R on your computer, you can start a session. A window with the R session will appear on your screen. Inside that window, you will see a smaller one with the R console. You will see the prompt symbol `>` awaiting for any command you want to insert.



If you type an algebraic expression like `3+2*5+2^2` the program acts like a calculator returning 17.

2. IDE RStudio

RStudio is a free Integrated Development Environment (IDE) for R. Visit `https://www.rstudio.com/` to download it and install it on your laptop or home computer.



Type the piece of code below either on the RStudio console (line by line) or on a script (open a new R script and type everything as in the example) and run it, line after line.

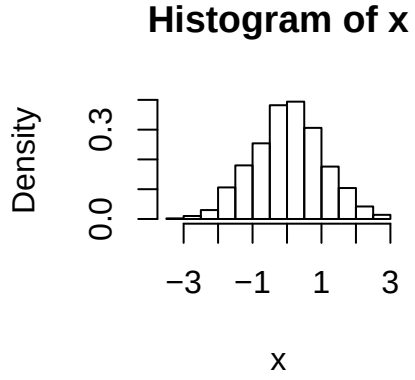
```
x=rnorm(1000)
mean(x)
```

```
## [1] 0.0299009
```

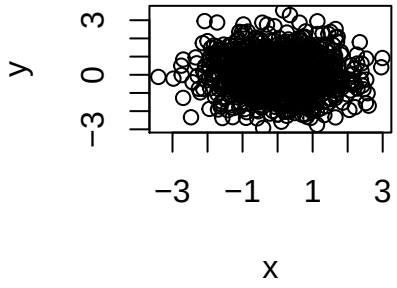
```
var(x)
```

```
## [1] 1.024791
```

```
hist(x,probability=T)
```



```
y=rnorm(1000)
plot(x,y)
```



In the first line we assign 1000 observations drawn from a standard normal distributiopn to `x`, in the second line we calculate the mean of `x` (value close to 0), in the third line its variance (value close to 1), in the fourth line we plot a histogram of `x`, in the fifth line we assign another 1000 observations drawn from the same normal model to `y`, and finally in the last line we obtain a scatter plot of `x` and `y`.

Some basic commands to explore

General:

```
> x=c(-1.1,2.2,5.3,4.7,1.6,2.2,4.3,2.2,1.1)
> x
> length(x)
> abs(x)
> x^2
> sum(x)
> help(sum)
> table(x)
> mean(x)
> median(x)
> mean(x,trim=1/9)
> var(x)
> quantile(x)
> quantile(x,.25)
> stem(x)
> x>2
> sum(x>2)/length(x)
> y=(1:10)
> y=seq(0,7,0.5)
> y+1
> y=matrix(c(1,2,4,3,7,9),ncol=2,byrow=T)
> y
> y[1,]
> y[1,2]
> y=data.frame(y)
> y$X2
> mean(y$X2)
```

```
> attach(y)
> mean(X2)
> X1[X2==9]
> detach(y)
```

Graphical:

```
> win.graph()
> dev.off()
> boxplot(x)
> hist(x)
> hist(x,breaks=seq(1,5.5,.5))
> aux1=seq(-2,2,.2)
> aux2=aux1^2
> plot(aux1,aux2)
> points(aux1[c(1,2)],aux2[c(1,2)],col="red")
> points(aux1[c(3,4)],aux2[c(3,4)],pch=3)
> var(aux1,aux2)
> cor(aux1,aux2)
> y=c(2,3,5,5,2,3,4,2,2)
> lsfit(x,y)
> plot(x,y)
> abline(lsfit(x,y))
> cor(x,y)
```

3. R Markdown

This document was obtained using R Markdown. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents, among others. For more details on using R Markdown see `http://rmarkdown.rstudio.com`.

See `https://github.com/rstudio/cheatsheets/raw/master/rmarkdown-2.0.pdf` for a cheatsheets.

C©pyleft

It is an implementation of copyright. It was born with the idea of protecting the free circulation of computer code and the knowledge it contains. It involves using copyright law to protect rather than restrict freedom of copying, modification, and redistribution (including sale). The author grants all the rights that he/she has over his/her work to the rest of the people. Everyone. The only condition is that any work derived from something protected with Copyleft must maintain the same freedoms.

Many people have contributed to the writing of these practices, and the list below is not intended to be exhaustive, although I would like to. If you think a name is missing, please do not hesitate to contact me at `<El.Aprendiz.de.Getafe@gmail.com>`.

Ignacio Cascos, Isabel Molina, Carlos Montes, Carlos Ruiz, **Ismael Sánchez**, Iñaki Ucar, Andrés M. Alonso.