

Introduction to R



Tiago Nardi

University of Pavia

What is R?

R is a language and environment for
statistical computing and graphics

R provides tools for statistical analysis
and data visualisation



Pros

- Free (as in "free speech", but also as in "free beer")
- It runs on all platforms (Windows, Mac or Linux)
- Can be extended with packages
- Large community: easy to get help
- R supports fairly large datasets (assuming you have enough RAM)

Cons

- You have to learn how to use a scripting language
- R natively supports only fairly large datasets

Aim: coding in R
without being too
scared



Installation

- Installing R
- Installing RStudio, the graphical interface

Installing R

Download R from the <http://cran.mirror.garr.it/mirrors/CRAN/> web page and install it

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux \(Debian, Fedora/Redhat, Ubuntu\)](#)
- [Download R for macOS](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.





[Source Code for all Platforms](#)

After the installation you should have two executable: *R_x86_64* and a *R_i386*

RStudio

For this course we will use **RStudio** . It is a graphical interface for R, that helps the programmer making it more user-friendly.

After installing R, download RStudio from
<https://www.rstudio.com/products/rstudio/download/#download>

OS	Download	Size	SHA-256
Windows 10/8/7	 RStudio-1.2.5019.exe	149.82 MB	7c6a943c
macOS 10.12+	 RStudio-1.2.5019.dmg		
Ubuntu 14/Debian 8	 rstudio-1.2.5019-amd64.deb		
Ubuntu 16	 rstudio-1.2.5019-amd64.deb	104.91 MB	24fad367

Select your version and click to download

Plan

4 Lessons

- Introduction
- Basic functions
- Plotting
- Hypothesis testing

+

1 Lesson

- Simulation of exam

Files

The *R* folder in the gdrive *bioinformatics_2021_2022* contains all the required material:

- *patric_redux.csv*, the table we will use for the lessons
- the course slides in pdf (including this one)
- the *R_commands.pdf*, a cheatsheet with the commands to use for the exam (it will be also provided during the exam)
- zip files for the lessons

The course material for R is also available online at
https://tiagonardi.github.io/R_intro_2022/

Other resources

A complete cheatsheet for R is available at

<https://github.com/rstudio/cheatsheets/blob/main/base-r.pdf>