aboutR

Christine Dillmann

04/12/2020

This document compiles information from the R project website, wikipaedia, and the www.

1 The R project

1.1 R is a free software

R is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.

R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, . . .) and graphical techniques, and is highly extensible. R provides an Open Source route to participation in research in statistical methodology.

One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control.

R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

For more details, see https://www.r-project.org/

1.2 R core team

The current R is the result of a collaborative effort with contributions from all over the world. R was initially written by Robert Gentleman and Ross Ihaka—also known as "R & R" of the Statistics Department of the University of Auckland. Since mid-1997 there has been a core group, the R Core Team (https://www.r-project.org/contributors.html), working on major developments.

1.3 The R foundation

The R Foundation is a not for profit organization working in the public interest. It has been founded by the members of the R Development Core Team in order to:

- Provide support for the R project and other innovations in statistical computing. We believe that R has become a mature and valuable tool and we would like to ensure its continued development and the development of future innovations in software for statistical and computational research.
- Provide a reference point for individuals, instititutions or commercial enterprises that want to support or interact with the R development community.
- Hold and administer the copyright of R software and documentation.

R is an official part of the Free Software Foundation's GNU project, and the R Foundation has similar goals to other open source software foundations like the Apache Foundation or the GNOME Foundation.

Among the goals of the R Foundation are the support of continued development of R, the exploration of new methodology, teaching and training of statistical computing and the organization of meetings and conferences with a statistical computing orientation. We hope to attract sufficient funding to make these goals realities.

The R Foundation Statutes can be downloaded as PDF file in English (https://www.r-project.org/foundation/Rfoundation-statutes.pdf) or German.

1.4   R-CRAN : The Comprehensive R Archive Network

CRAN is a network of ftp and web servers around the world that store identical, up-to-date, versions of code and documentation for R. Each server is called a *mirror*. Those servers are mostly maintained by university staffs. R can be downloaded from any mirror belonging to the CRAN.

2   R architecture

- *R-base* is the Application Programming Interface (API), defined as a normalized set of *classes*, *methods*, *functions* and *constants* that serve as an interface between libraries/softwares.

- *Packages* are collections of R functions, data, and compiled code in a well-defined format. The directory where packages are stored is called the library. R comes with a standard set of packages (*R-base*). Others are available for download and installation. Once installed, they have to be loaded into the session to be used (using the *library* R function). R packages are released in *repositories* (web servers mirrors) from where they can be downloaded.. Common repositories for R-packages are :

  - **RCRAN**. Packages can be installed using the `install.packages(`*"name"*`)` function.

  - **BioClite**. Packages can be installed using `BiocManager::install(`*"name"*`)`.

  - **git**. Git is a distributed version-control system for tracking changes in any set of files, originally designed for coordinating work among programmers cooperating on source code during software development. Git was created by Linus Torvalds in 2005 for development of the Linux kernel. As with most other distributed version-control systems, and unlike most client–server systems, every Git directory on every computer is a full-fledged repository with complete history and full version-tracking abilities, independent of network access or a central server. Git is free and open-source software distributed under GNU General Public License Version 2. Sets of R sripts can be installed from git using `git clone `*`adress`* as online command from a terminal.

3   Versioning

Versioning is the creation and management of multiple releases of a product, all of which have the same general function but are improved, upgraded or customized. The term applies especially to operating systems (OSs), software and Web services.
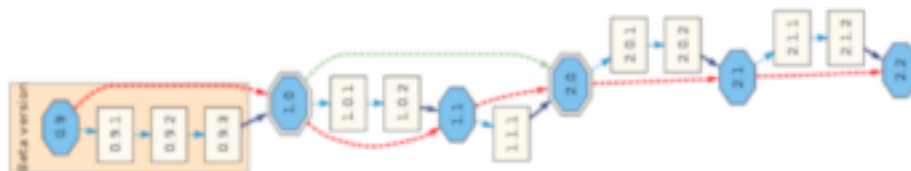

Figure 1: Versioning

Software upgrade versioning is the process of assigning either unique version names or unique version numbers to unique states of computer software. Within a given version number category (major, minor), these

numbers are generally assigned in increasing order and correspond to new developments in the software. At a fine-grained level, revision control is often used for keeping track of incrementally different versions of information, whether or not this information is computer software.

Semantic versioning is a formal convention for specifying compatibility using a three-part version number: major version; minor version; and patch. The patch number is incremented for minor changes and bug fixes which do not change the software's application programming interface (API). The minor version is incremented for releases which add new, but backward-compatible, API features, and the major version is incremented for API changes which are not backward-compatible. For example, software which relies on version 2.1.5 of an API is compatible with version 2.2.3, but not necessarily with 3.2.4.

4 Rstudio

RStudio is an integrated development environment (IDE) for R, a programming language for statistical computing and graphics. It is available in two formats: RStudio Desktop is a regular desktop application while RStudio Server runs on a remote server and allows accessing RStudio using a web browser.

The RStudio IDE is available with the GNU Affero General Public License version 3. The AGPL v3 is an open source license that guarantees the freedom to share the code.RStudio is available in free and fee-based (commercial) editions.

In addition to the RStudio IDE, RStudio PBC and its employees develop, maintain, and promote a number of R packages (including Rmarkdown and knitr, that allow to mix text with code from different languages, includig R).