Last updated October 3, 2016

Introduction to the software

G. Bacaro

Design and Analysis of Environmental Monitoring and Experiments
Master Degree in Global Change Ecology
I Year, I term

Lecture Aims

Introduction to Applied Statistics and R

- *Introduce students to the use of R and Rstudio;*

- *Familiarize with some of the most common statistical analyses performed in R.*

- *Stimulate ideas for using these techniques with your own dataset.*

Some of my works with

Introduction to Applied Statistics and R

OPEN ACCESS Freely available online

PLOS ONE

A New Measure of Functional Evenness and Some of Its Properties

Carlo Ricotta[1]*, Giovanni Bacaro[2,3], Marco Moretti[4]

1 Department of Environmental Biology, University of Rome 'La Sapienza', Rome, Italy, 2 CNR, Istituto di Ricerca per la Protezione Idrogeologica, Perugia, Italy, 3 BIOCONNET, Biodiversity and Conservation Network, Department of Life Science, University of Siena, Siena, Italy, 4 Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

Abstract

Functional evenness is increasingly considered an important facet of functional diversity that sheds light on the complex relationships between community assembly and ecosystem functioning. Nonetheless, in spite of its relevant role for ecosystem functioning, only a few measures of functional evenness have been proposed. In this paper we introduce a new measure of functional evenness that reflects the regularity in the distribution of species abundances, together with the evenness in their pairwise functional dissimilarities. To show how the proposed measure works, we focus on changes in functional evenness calculated from Grime's classification of plant strategies as competitors (C), stress-tolerators (S) and ruderals (R) along a post-fire successional gradient in temperate chestnut forests of southern Switzerland.

Ecological Complexity 12 (2012) 63–69

Contents lists available at SciVerse ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Short note

The spatial domain matters: Spatially constrained species rarefaction in a Free and Open Source environment

Giovanni Bacaro [a,b], Duccio Rocc
Alessandro Chiarucci [a]

[a] BIOCONNET, Biodiversity and Conservation Network, I
[b] CNR-IRPI, Istituto di Ricerca per la Protezione Idrogeo
[c] Department of Biodiversity and Molecular Ecology, Re

```
## Function code

SCR <- function(community, spatial_order) {
library(vegan)
 f <- nrow(spatial_order)
 n <- ncol(spatial_order)
 change <- names(community)
 change[1] <- "com"
 colnames(community) <- change
 result <- array(dim = c(f, n))
 for(i in 1:n) {
  frame<-data.frame(spatial_order[,i])
  colnames(frame) <- "ordered"
  agg <- merge(frame,community, by.x="ordered", by.y="com", sort=FALSE)
  c <- specaccum(agg[,2:ncol(agg)], method="collector")
  result[,i ] <- c$richness
 }
 average <- rowMeans(result)
 IC_plus <- average + (1.96*(sd(t(result))/sqrt(n)))
 IC_neg <- average - (1.96*(sd(t(result))/sqrt(n)))
 SCR <- data.frame(as.matrix(average), IC_neg,IC_plus)
 names(SCR) <- c("SCR", "95%IC Negative", "95%IC Positive")
 return(SCR)
}
```

```
Function Syntax

FeveR<-function(abundances, distances)
{
 rel_abundance<-sweep(abundances, 1, rowSums(abundances), "/")
rel_abu_matrix<-as.matrix(rel_abundance)
n_plot<-nrow(rel_abu_matrix)
n_species<-ncol(rel_abu_matrix)
dista_matrix<-as.matrix(distances)
###BULLA1
index_array<-rep(NA, n_plot)
for(i in 1:n_plot)
{mat_develop<-rep(rel_abu_matrix[i,],n_species)
matrix_two<-(t(matrix(mat_develop, nrow=n_species,nc
matrix_non<-matrix(mat_develop, nrow=n_species,ncol
uni<-matrix(1,nrow=n_species,ncol=n_species)
subtraction<- uni-matrix_non
division<-matrix_two/subtraction
moltiplication<-division*dista_matrix
row_sums<-rowSums(moltiplication)
per_abundance<-row_sums*rel_abu_matrix[i,]
great_sum<-sum(per_abundance)
divised<-per_abundance/great_sum
espress<-which(per_abundance>0)
S<-length(espress)
bulla<-rep(NA, length(divised))
for(l in 1:length(divised))
{
 bulla[l]<-min(divised[l], 1/S)
}
indice<-sum(bulla)
norm_index<-(indice-(1/S))/(1-(1/S))
index_array[i]<-norm_index
}
return(index_array)
}
```

Appendix I: R function 'BetaDispersion'

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License http://www.gnu.org/licenses/.

Ecol Res
DOI 10.1007/s11284-011-0899-z

ORIGINAL ARTICLE

Giovanni Bacaro · Margherita Gioria
Carlo Ricotta

Testing for differences in beta diversity from plot-to-plot dissimilarities

Please, consider that I'm not a Statistician!!

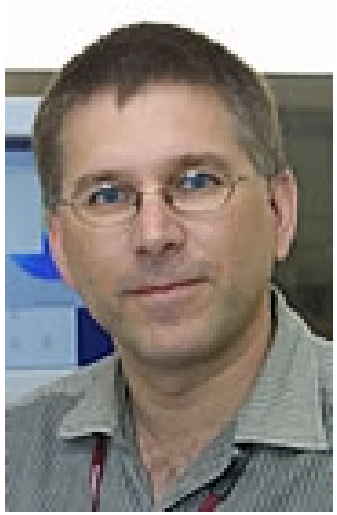
UNIVERSITÀ DEGLI STUDI DI TRIESTE
DIPARTIMENTO DI SCIENZE DELLA VITA

Design and Analysis of environmental monitoring (DAEME)
MD in Global Changes Ecology

A very brief introduction to R

- R started in the early 1990's as a project by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, intended to provide a **statistical environment** in their teaching lab. The lab had Macintosh computers, for which no suitable commercial environment was available.

Ross Ihaka



Robert Gentleman

R: A Language for Data Analysis and Graphics

Ross IHAKA and Robert GENTLEMAN

In this article we discuss our experience designing and implementing a statistical computing language. In developing this new language, we sought to combine what we felt were useful features from two existing computer languages We feel that the new language provides advantages in the areas of portability, computational efficiency, memory management, and scoping.

Key Words: Computer language; Statistical computing.

Ross Ihaka is Senior Lecturer, and Robert Gentleman is Senior Lecturer, Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand, e-mail ihaka@stat auckland ac nz.

Journal of Computational and Graphical Statistics, Volume 5, Number 3, Pages 299–314

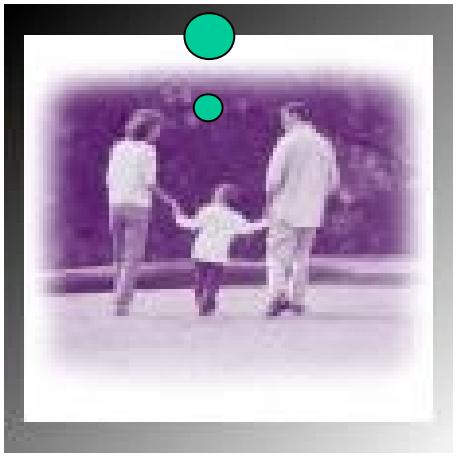
Parents

- The S language

  - S: an ***interactive environment*** for data analysis developed at ***Bell Laboratories*** since 1976
  - Exclusively licensed by *AT&T/Lucent* to *Insightful Corporation*, Seattle WA. Product name: “S-plus”.

Version 1	1976-1980	Honeywell GCOS, Fortran-based
Version 2	1980-1988	Unix; Macros, Interface Language
	1981-1986	QPE (Quantitative Programming Environment)
	1984-	General Outside Licensing, books
Version 3	1988-1998	C-based; S functions & objects
	1991; rev. 1992	Statistical Models; informal Classes & Methods
	1993-	Exclusive license to StatSci (later MathSoft) for S-Plus.
Version 4	1998-	Formal class-method model; Connections; Large Objects
	1999-	Interfaces to Java, Corba?

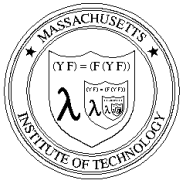
My father is S, mother is Scheme, but why my name is “R”?



You can learn more from:
http://cm.bell-labs.com/cm/ms/departments/sia/S/history.html

R Parents

- The Scheme language



Scheme is a statically scoped and properly tail-recursive dialect of the Lisp programming language invented by Guy Lewis Steele Jr. and Gerald Jay Sussman.

Learn more: http://swiss.csail.mit.edu/projects/scheme/

- Scheme's underlying semantics + S'syntax = R

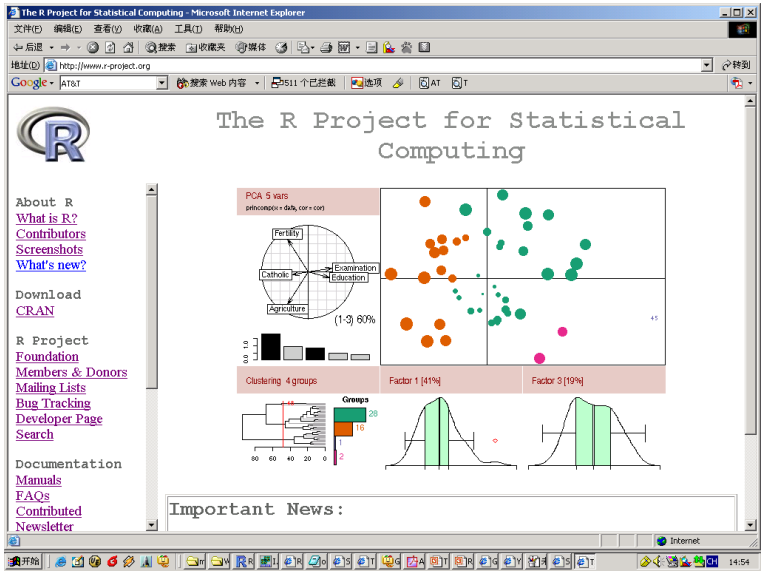
" We have named our language R –in part to acknowledge the influence of S and in part to celebrate our own efforts."

-- **R**. Ihaka
R. Gentleman

R Now

- Since mid-1997 there has been a core group who can modify the R source code CVS archive.
- The R package system

  CRAN (the Comprehensive R Archive Network)



http://www.r-project.org

The Characters of R

- **R is "GNU S" — A language and environment for data manipulation, calculation and graphical display.**
  - That is R is a Free Software (or Open source software). (Here, Free refers to freedom, not price, although R is free in that sense as well)

- **The core of R is an interpreted computer language.**
  - A mosaic of procedure-based programming and object-oriented programming
  - Good interface to procedures written in C, C++, FORTRAN and other languages
  - A flexible data exchange mechanism accessing relational databases -ODBC, PostgreSQL, MySQL and so on.

Open Source Philosophy

Let the four freedoms paradigm apply to ecology

Duccio Rocchini and Markus Neteler

Fondazione Edmund Mach, Research and Innovation Centre, Department of Biodiversity and Molecular Ecology, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

The famous 'four freedoms' expounded by Stallman [1] are: (i) the freedom to run the program for any purpose; (ii) the freedom to study how the program works and adapt it to one's own needs; (iii) the freedom to redistribute copies; and (iv) the freedom to make improvements to the program and release them to the public. Thus, the whole (scientific) community benefits from software development. These freedoms are also inherent in several free software licenses, the GNU General Public License (GPL) being one of the most popular.

And Statistics

- Most packages deal with statistics and data analysis.
- Powerful statistical graphics.
- Well crosstalking with other statistical softwares.
- Most R user are statistical experts. You can learn more modern analysis method from they by email.
- You can do it when you come across a thing no body do it before.

Learning R

Introduction to Applied Statistics and R

UNIVERSITÀ
DEGLI STUDI DI TRIESTE
DIPARTIMENTO DI
SCIENZE DELLA VITA

Design and Analysis of environmental monitoring (DAEME)
MD in Global Changes Ecology

Pros and Cons

- Fast and free.
- State of the art: Statistical researchers provide their methods as R packages. SPSS and SAS are years behind R!
- 2nd only to MATLAB for graphics.
- Mx, WinBugs, and other programs use or will use R.
- Active user community
- Excellent for simulation, programming, computer intensive analyses, etc.
- Forces you to *think* about your analysis.
- Interfaces with database storage software (SQL)
- Not user friendly @ start - steep learning curve, minimal GUI.
- No commercial support; figuring out correct methods or how to use a function on your own can be frustrating.
- Easy to make mistakes and not know.
- Working with large datasets is limited by RAM
- Data prep & cleaning can be messier & more mistake prone in R vs. SPSS or SAS
- Some users complain about hostility on the R listserve

vs Commercial Software

Topic	*SAS Product*	*SPSS Product*	*R Package*
Advanced Models	SAS/STAT®	SPSS Advanced Models™	stat, MASS, *many* others
Basics	SAS®	SPSS Base™	R
Conjoint Analysis	SAS/STAT®: Transreg	SPSS Conjoint™	homals, psychoR , bayesm
Correspondence Analysis	SAS/STAT®: Corresp	SPSS Categories™	homals, MASS, FactoMineR, ade4, PTAk, cocorresp, vegan, made4, PsychoR
Custom Tables	SAS Base® Report, SQL, Tabulate	SPSS Custom Tables™	reshape
Data Access	SAS/ACCESS®	SPSS Data Access Pack™	DBI, foreign, RODBC
⋮	⋮	⋮	⋮

For the full comparison chart, see http://rforsasandspssusers.com/ by Bob Muenchen

There are over 800 add-on packages

(http://cran.r-project.org/src/contrib/PACKAGES.html)

- This is an enormous advantage - new techniques available without delay, and they can be performed using the R language you already know.
- Allows you to build a customized statistical program suited to your own needs.
- Downside = as the number of packages grows, it is becoming difficult to choose the best package for your needs, & QC is an issue.

News!

Introduction to Applied Statistics and R

The R Journal

RSS Feed
ISSN: 2073-4859

About The R Journal

The R Journal is the open access, refereed journal of the R project for statistical computing. It features short to medium length articles covering topics that might be of interest to users or developers of R, including

Add-on packages:	short introductions to R extension packages.
Programmer's Niche:	hints for programming in R.
Help Desk:	hints for newcomers explaining aspects of R that might not be so obvious from reading the manuals and FAQs.
Applications:	demonstrating how a new or existing technique can be applied in an area of current interest using R, providing a fresh view of such analyses in R that is of benefit beyond the specific application.

The R Journal intends to reach a wide audience and have a fast-track but thorough review process. Papers are expected to be reasonably short, clearly written, not too technical, and of course focused on R. Authors of refereed articles should take care to:

- put their contribution in context, in particular discuss related R functions or packages;
- explain the motivation for their contribution;
- provide code examples that are reproducible.

Continuing from *R News*, *The R Journal* also has a news section, including information on:

Changes in R:	new features of the latest release.
Changes on CRAN:	new add-on packages, manuals, binary distributions, mirrors,...
Upcoming conferences:	announcements of conferences related to R.
Conference reports:	overviews of past conferences related to R.

Help and specific mailing lists

Introduction to Applied Statistics and R

[Home]

Download

CRAN

R Project

About R
Contributors
What's New?
Mailing Lists
Bug Tracking
Conferences
Search

R Foundation

Foundation
Board
Members
Donors
Donate

Mailing Lists

Please read the instructions below and the posting guide *before* sending anything to any mailing list!

Thanks to Martin Maechler (and ETH Zurich), there are five general mailing lists devoted to *R*.

R-announce

This list is for *major* announcements about the development of *R* and the availability of new code. It has a *low volume* (typically only a few messages a month) and everyone mildly interested should consider subscribing, but note that R-help gets everything from R-announce as well, so you don't need to subscribe to both of them.

Note that the list is *moderated* to be used for announcements mainly by the R Core Development Team. Use the web interface for information, subscription, archives, etc.

R-help

The 'main' *R* mailing list, for discussion about problems and solutions using *R*, announcements (not covered by 'R-announce' or 'R-packages', see above), about the availability of new functionality for *R* and documentation of *R*, comparison and compatibility with *S-plus*, and for the posting of nice examples and benchmarks. Do read the posting guide *before* sending anything!

This has become quite an active list with dozens of messages per day. An alternative is to subscribe and choose daily digests (in plain or MIME format). Use the web interface for information, subscription, archives, etc.

Special Interest Groups

Additionally, there are several specific ***S**pecial **I**nterest **G**roup* (=: SIG) mailing lists; however do post to *only one* list at time ('SIG' or general one), cross-posting is considered to be impolite.

- `R-SIG-Mac` : R Special Interest Group on Mac ports of R
- `R-SIG-DB` : R SIG on Database Interfaces
- `R-SIG-Debian` : R Special Interest Group for Debian ports of R
- `R-SIG-dynamic-models` : Special Interest Group for Dynamic Simulation Models in R
- `R-SIG-ecology` : Using R in ecological data analysis
- `R-SIG-Epi` : R for epidemiological data analysis
- `R-SIG-Fedora` : R Special Interest Group for Fedora and Redhat ports of R
- `R-SIG-Finance` : Special Interest Group for 'R in Finance'
- `R-SIG-Geo` : R Special Interest Group on using Geographical data and Mapping
- `R-SIG-gR` : R SIG on gRaphical models
- `R-SIG-GUI` : R Special Interest Group on GUI Development
- `R-SIG-HPC` : R SIG on High-Performance Computing
- `R-SIG-Insurance` : Special Interest Group on using R in actuarial science and insurance
- `R-SIG-Jobs` : R SIG List for Announcements of Jobs where R is used
- `R-SIG-mediawiki` : R SIG on the R Extension for Mediawiki
- `R-SIG-mixed-models` : R SIG on Mixed Effect Models, notably lmer() related
- `R-SIG-networks` : R SIG for users and developers of network- or graph-related software within R
- `R-SIG-phylo` : R SIG on phylogenetic and comparative methods and analyses
- `R-SIG-QA` : R SIG on Quality Assurance & Validation
- `R-SIG-Robust` : R SIG on Robust Statistics
- `R-SIG-teaching` : SIG on Teaching Statistics (and more) using R
- `R-SIG-Wiki` : SIG on the Development of an "R Wiki"

To satisfy geographic or regional (or subject) needs, some R users have formed *"R User Groups"* for which there are mailing lists. Information about (some of) these groups and their lists can be found at the RUG web page, maintained by John C. Nash.

UNIVERSITÀ DEGLI STUDI DI TRIESTE
DIPARTIMENTO DI SCIENZE DELLA VITA

Design and Analysis of environmental monitoring (DAEME)
MD in Global Changes Ecology

Free Documents

Introduction to Applied Statistics and R

http://cran.r-project.org/manuals.html

The R Manuals

edited by the R Development Core Team.

The following manuals for R were created on Debian Linux and may differ from the manuals for Mac or Windows on platform-specific pages, but most parts will be identical for all platforms. The correct version of the manuals for each platform are part of the respective R installations. The manuals change with R, hence we provide versions for the most recent released R version (R-release), a very current version for the patched release version (R-patched) and finally a version for the forthcoming R version that is still in development (R-devel).

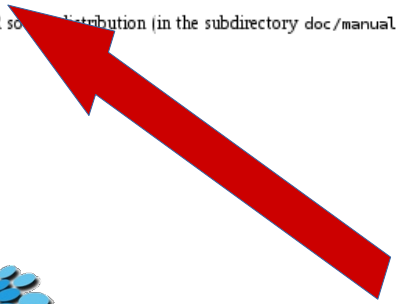
Here they can be downloaded as PDF files, EPUB files, or directly browsed as HTML:

Manual	R-release	R-patched	R-devel
An Introduction to R is based on the former "Notes on R", gives an introduction to the language and how to use R for doing statistical analysis and graphics.	HTML | PDF | EPUB	HTML | PDF | EPUB	HTML | PDF | EPUB
R Data Import/Export describes the import and export facilities available either in R itself or via packages which are available from CRAN.	HTML | PDF | EPUB	HTML | PDF | EPUB	HTML | PDF | EPUB
R Installation and Administration	HTML | PDF | EPUB	HTML | PDF | EPUB	HTML | PDF | EPUB
Writing R Extensions covers how to create your own packages, write R help files, and the foreign language (C, C++, Fortran, ...) interfaces.	HTML | PDF | EPUB	HTML | PDF | EPUB	HTML | PDF | EPUB
A draft of **The R language definition** documents the language *per se*. That is, the objects that it works on, and the details of the expression evaluation process, which are useful to know when programming R functions.	HTML | PDF | EPUB	HTML | PDF | EPUB	HTML | PDF | EPUB
R Internals: a guide to the internal structures of R and coding standards for the core team working on R itself.	HTML | PDF | EPUB	HTML | PDF | EPUB	HTML | PDF | EPUB
The R Reference Index: contains all help files of the R standard and recommended packages in printable form. (9MB, approx. 3500 pages)	PDF	PDF	PDF

Translations of manuals into other languages than English are available from the contributed documentation section (only a few translations are available).

The LaTeX or Texinfo sources of the latest version of these documents are contained in every R so distribution (in the subdirectory `doc/manual` of the extracted archive). Older versions of the manual can be found in the respective archives of the R sources. The HTML versions of the manuals are also part of most R installations (accessible using function `help.start()`).

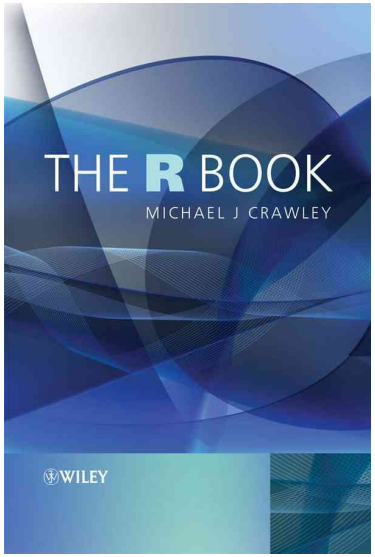
Please check the manuals for R-devel before reporting any issues with the released versions.

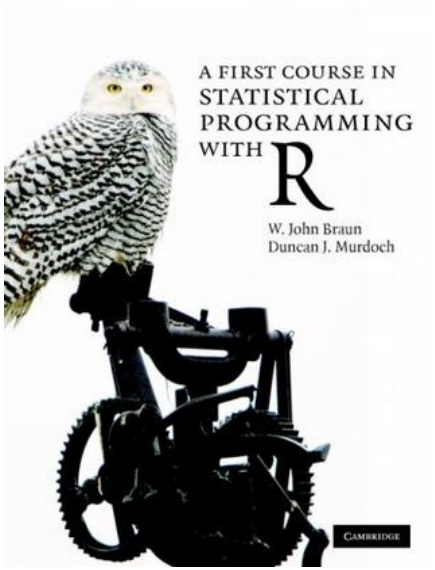


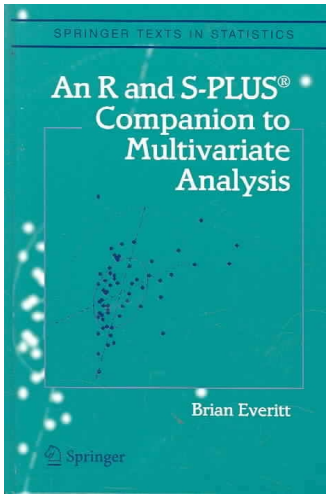
Books

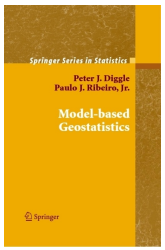
Introduction to Applied Statistics and R

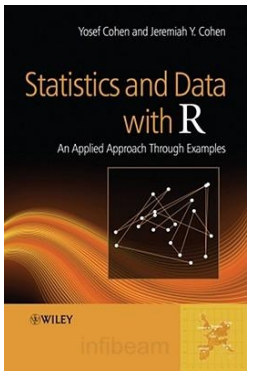
http://cran.r-project.org/manuals.html











https://drive.google.com/drive/folders/0BzxxHDYur0-ffmdqc1hFNjNjUExXM0xUcVFEVlVoeWRWYWRtMGswYkFDOHp1d0NNVURDQ00?usp=sharing

Related Projects

Introduction to Applied Statistics and R

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Search:

Home | Install | Help | Developers | About

ASU GeoDa Center
R Spatial Projects

Home | About Project | People | Presentations | Publications

R GUI Projects

Why a GUI?

R is provided with a command line interface (CLI), which is the preferred user interface for power users because it allows direct control on calculations and it is flexible. However, good knowledge of the language is required. CLI is thus intimidating for beginners. The learning curve is typically longer than with a graphical user interface (GUI), although it is recognized that the effort is profitable and leads to better practice (finer understanding of the analysis; command easily saved and replayed). The user interface remains the biggest difference between R and S-PLUS, since the latter implements a very sophisticated GUI. A fraction of the R users are asking for, and would probably benefit from a R GUI, mainly, occasional users and some teachers. R is open and communicating. Several projects develop or offer the opportunity to develop alternate user interfaces. They are presented here. A Special Interest Group mailing list (R-SIG-GUI) exists also to freely discuss concerned issues. You can subscribe here. The R (GUI) Wiki is also there to exchange information and ideas related to the use of R GUIs and to start using R.

Last update: 16 Augustus 2010 by Philippe Grosjean

R GUI
Why a GUI?
Types of GUIs
Evaluating GUIs
R and GUIs
R (GUI) Wiki

R-Forge

Project ▼ Search | Log In | New Acc

Home | My Page | Projects

What are R and R-Forge?

R is `GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc. Please consult the R-project homepage for further information.
R-Forge offers a central platform for the development of R packages, R-related software and further projects. It is based on FusionForge offering easy access to the best in SVN, daily built and checked packages, mailing lists, bug tracking, message boards/forums, site hosting, permanent file archival, full backups, and total web-based administration.

A Platform for the Whole R Community

In order to get the most out of R-Forge you'll need to register as a site user and then login. This will allow you to participate fully in all we have to offer, e.g., you may register your project. Of course, you may also browse the site without registration, but will only have limited access to some features. For details see the documentation.

Tag Cloud

Bayesian Bioinformatics C++ Clustering Ecology Multivariate Analysis Multivariate Regression R biostatistics classification clustering data mining ecology finance multivariate optimization spatial spatial data spatial methods time series

R-Forge Statistics

Hosted Projects: **1,916**
Registered Users: **9,824**

Most Active This Week

Final Words of Warnings!

"Using R is a bit akin to smoking. The beginning is difficult, one may get headaches and even gag the first few times. But in the long run,it becomes pleasurable and even addictive. Yet, deep down, for those willing to be honest, there is something not fully healthy in it."
--Francois Pinard