

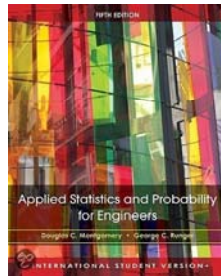
Statistical Software

- InterTU Seminar, 19-6-2013 -



dr. J.J.M. rijpkema
Eindhoven University of Technology,
dept. Mathematics & Computer Science
P.O.Box 513, 5600 MB Eindhoven, NL

in education !



Overview

- What is R ?
- R in Education ?
- Some TU/e-Experiences.....
- Some Educational Challenges.....
- Some References.....
- Perspective....



What is R?

WIKIPEDIA The Free Encyclopedia

Article Talk

R (disambiguation)

From Wikipedia, the free encyclopedia

R is the eighteenth letter of the Latin alphabet

R may also refer to:

- Mathematics** [edit]
- The set of all real numbers, often written as \mathbb{R} or \mathbb{R}
- The Pearson product-moment correlation coefficient r in statistics
- Computing** [edit]
- R (programming language), an environment for statistical computing and graphics
- R, RequestBot, a service in Qualtrics's R/C services
- IBM System R, an IBM database system
- Engineering** [edit]
- R-value (insulation), a non-SI measure of thermal resistance used in home energy audits
- R-value (soils), property of soils used in pavement design
- An electronic resistor, e.g. in circuit diagrams
- The R# refrigerant numbering system
- In finance mathematics, the ratio between the dividend and maximum value

Google R

Over 11 500 000 000 resultaten (0.15 seconden)

the molar gas constant = $8.3144621 \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1} \text{ mol}^{-1}$

The R Project for Statistical Computing

the R project.org - Vertaal deze pagina

R, also called GNU S, is a strongly functional language and environment to statistically explore data sets, make many graphical displays of data from custom...

The Comprehensive R Archive - CRAN - Mirrors - Manuals - Mailing Lists

R (softwarepakket) - Wikipedia

en.wikipedia.org/wiki/R_(softwarepakket) - Vertaal deze pagina

R is a free software programming language and a software environment for statistical computing and graphics. The R language is widely used among...

R (programming language) - Wikipedia, the free encyclopedia

en.wikipedia.org/wiki/R_(programming_language) - Vertaal deze pagina

R is a free software programming language and a software environment for statistical computing and graphics. The R language is widely used among...

[http://en.wikipedia.org/wiki/R_\(disambiguation\)](http://en.wikipedia.org/wiki/R_(disambiguation))

• What is R?

WIKIPEDIA The Free Encyclopedia

Article Talk

R (programming language)

From Wikipedia, the free encyclopedia

R is a free software programming language and a software environment for statistical computing and graphics. The R language is widely used among statisticians and data miners for developing statistical software^[1] and data analysis^[2]. Polls and surveys of data miners are showing R's popularity has increased substantially in recent years.^[3]

R is an implementation of the S programming language combined with lexical scoping semantics inspired by Scheme. S was created by John Chambers while at Bell Labs. It was created by Ross Ihaka and Robert Gentleman^[4] at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team, of which Chambers is a member. R is named partly after the first names of the first two R authors and partly as a play on the name of S.^[5]

R is a GNU project.^[6] The source code for the R software environment is written primarily in C, Fortran,

The R Project for Statistical Computing

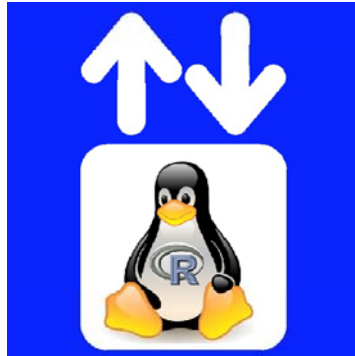
Getting Started:

- R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download](#), [install](#), [learn](#) more about R, please read our [getting started](#) page.
- If you have questions about R, like how to download and install the software, or what the license terms are, please read our [getting started](#) page.

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

[http://en.wikipedia.org/wiki/R_\(programming_language\)](http://en.wikipedia.org/wiki/R_(programming_language))

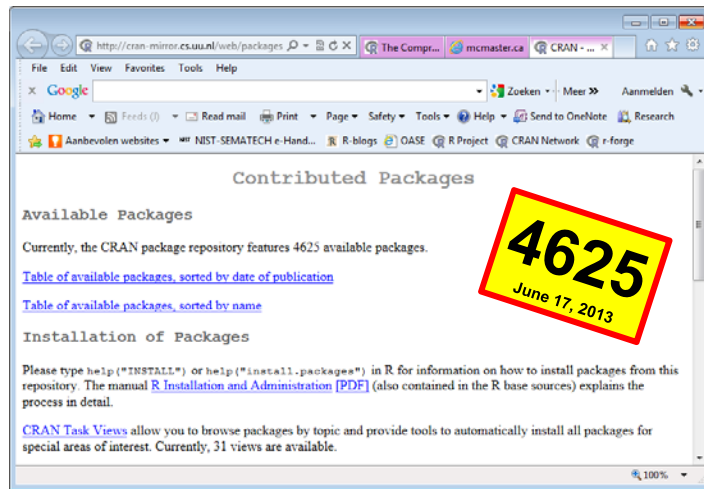
• Elevator Pitch:



- GNU-license
- Open source
- Supports different operating systems
- Active community
- Exponential growing number of packages

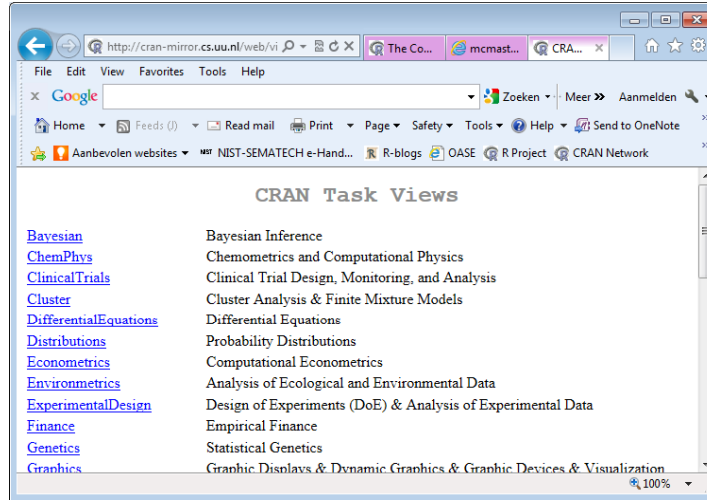
The Lingua Franca for Statistics....

• An Exponential Growing Number of Packages !



<http://cran-mirror.cs.uu.nl/web/packages/>

- For all kind of Tasks:



<http://cran-mirror.cs.uu.nl/web/views/>

- Active Community:



<http://www.edii.uclm.es/~useR-2013/>



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



<http://www.istatsoft.org/>

• Just a Hype ?

Data Analysts Captivated by R's Power

R first appeared in 1996, when the statistics professors Robert Gentleman, left, and Ross Ihaka released the code as a free software package.

By ASHLEE VANCE
Published: January 6, 2009

To some people R is just the 18th letter of the alphabet. To others, it's the rating on racy movies, a measure of an attic's insulation or what pirates in movies say.

Related
Bits: R: You Ready for R?
The R Project for Statistical Computing

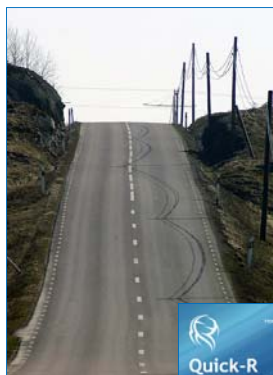
R is also the name of a popular programming language used by a growing number of data analysts inside corporations and academia. It is becoming their lingua franca partly because data mining has entered a golden age, whether being used to set ad prices, find new drugs more quickly or fine-tune financial models. Companies as diverse as [Google](#), [Pfizer](#), [Merck](#), [Bank of America](#), the InterContinental Hotels Group and Shell use it.

Twitter LinkedIn Sign in to Email Print Single Page Reprints Share

Way Back Machine Watch Trailer

<http://www.nytimes.com/2009/01/07/technology/business-computing/07program.html?pagewanted=all>

• Steep Learning Curve:



```
RGui (32-bit)
File Edit View Misc Packages Windows Help

R Console
R version 3.0.0 (2013-04-03) -- "Masked Marvel"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/x86_64 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

>
```

Quick-R

Why R has A Steep Learning Curve

A long answer to a simple question...

I have been a hardware SAs and SPSS programmer for more than 20 years, a system programmer for 18 years and a VBA programmer for 2 years. But when I started learning R recently, I found it frustratingly difficult. Why?

R has a different paradigm: Rather than setting up a complete analysis at once, the process is highly interactive.

Robert I. Kabacoff

<http://www.statmethods.net/about/learningcurve.html>

R in Education?

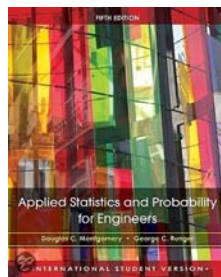


<http://en.wikipedia.org/wiki/3R%27s>



the 4th R ?

- Restriction:



Philosophy: Focus on the principles of Statistics and Data Analysis, not on details of a statistical package. Software is just a tool, which we don't want to become too much dependent on.

However: Anticipate on the R-paradigm of dynamic analyses, on a new way of 'open source' learning and on new approaches of analysis that might become in reach!

Some TU/e-Experiences.....

©jr / department of mathematics and computer science InterTU Seminar

TU/e Elective Bachelor Course 2AS00:



Jim Albert & Mario Rizzo, 'R by Example', ISBN 9781461413646



D. De Veaux et al., 'Stats: Data and Models', ISBN 978-0-321-75372-4

Modules

- Exploratory Data Analysis
- Probability Distributions & More Probability Distributions
- Testing/Estimation & More Testing/Estimation
- Regression Modeling & More Regression Modeling
- ANOVA & More ANOVA



• Flatten the Learning Curve with R Commander !

The screenshot shows the R Commander interface with three highlighted windows:

- Scripts:** Contains R code for loading data, creating a histogram, and a pie chart.
- Numerical Output:** Shows the execution of the R code, including library calls and data processing commands.
- Graphical Output:** Displays a histogram of 'Cars93\$MPG.highway' with frequency on the y-axis (0-40) and highway miles per gallon on the x-axis (20-50).

<http://socserv.mcmaster.ca/fox/Misc/Rcmdr/>

• R-Commander: Statistical Techniques

The screenshot shows the R Commander menu structure with three sub-menus highlighted:

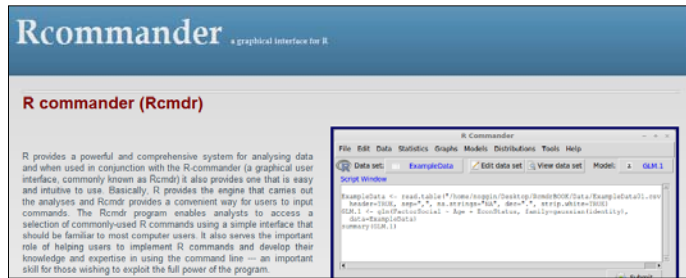
- Statistics:** Summaries, Contingency tables, Means, Proportions, Variances, Nonparametric tests, Dimensional analysis, Fit models.
- Distributions:** Continuous distributions, Discrete distributions.
- Models:** Select active model..., Summarize model, Add observation statistics to data..., Confidence intervals..., Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Stepwise model selection..., Subset model selection..., Hypothesis tests, Numerical diagnostics, Graphs.

• R-Commander: Additional Information



The R Commander: A Basic-Statistics GUI for R
 Current Version: **1.9-x**
 Development Version **2.0-0 (built under R 3.0.1): source package, Windows binary package, Mac OS X binary package**
John Fox

<http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/>



<http://www.rcommander.com/>

Some Educational Challenges



INT. J. MATH. EDUC. SCI. TECHNOL., 1991, VOL. 22, NO. 5, 791-798

Mathematics courses with a PC

by J. J. M. RIJPKEMA, F. H. SIMONS and J. G. M. M. SMITS
 Department of Mathematics and Computing Science,
 Eindhoven University of Technology, P.B. 513,
 5600 MB Eindhoven, The Netherlands

(Received 28 August 1990)

The authors present their views on the impact that computers should have on the contents of service courses in mathematics. They point to the fact that besides analytical techniques, mathematical software-packages are used increasingly in many fields of applications.

1. Introduction

The teaching of service courses in mathematics for first and second year university students is and will be affected in two ways by the large-scale availability of computers and software-packages.

On the one hand the computer influences the way in which mathematics is taught, or will be taught in future. This has been widely recognized and there is fast-growing literature on the topic. Roughly, and excluding specialized areas such as numerical analysis, most literature concentrates on the following two topics:

- (1) How can a computer be used as a teacher or as an aid in the teaching process (computer aided teaching or computer aided instruction, see e.g. [1])
- (2) How can a computer with software be used to illustrate the concepts taught in existing mathematics courses (see e.g. [2]).


On the other hand, the increasing use of computers and mathematical software-packages will have consequences for the contents of the service courses themselves.




**R has a different paradigm:
 Rather than setting up a complete analysis
 at once, the process is highly interactive.**

International Journal of Mathematical Education in Science and Technology, Vol. 22(1991), No. 5, p. 791-798


• Alternative Interfaces:



Journal of Statistical Software
June 2012, Volume 49, Issue 1.
<http://www.jstatsoft.org/>



R Through Excel
R.M. Heiberger et al.



Graphical User Interfaces for R

<http://www.red-r.org/>

Pedro M. Valero-Mora
Universitat de València

Rubén D. Ledesma
Universidad Nacional de Mar del Plata



JGR
Java GUI for R
Version 1.2-7
Markus Helbig
Simon Urbaniak
Ian Fellows
<http://www.r-suda.org>
(c) 2003 - 2011, RoSuDa, Univ. Augsburg
<http://forge.net/JGR/index.html>

<http://www.istatsoft.org/v49/i01/paper>




DEDUCER
INTUITIVE DATA ANALYSIS
<http://www.deducer.org>

pmg – Poor Man's GUI
an R GUI for Introductory Statistics

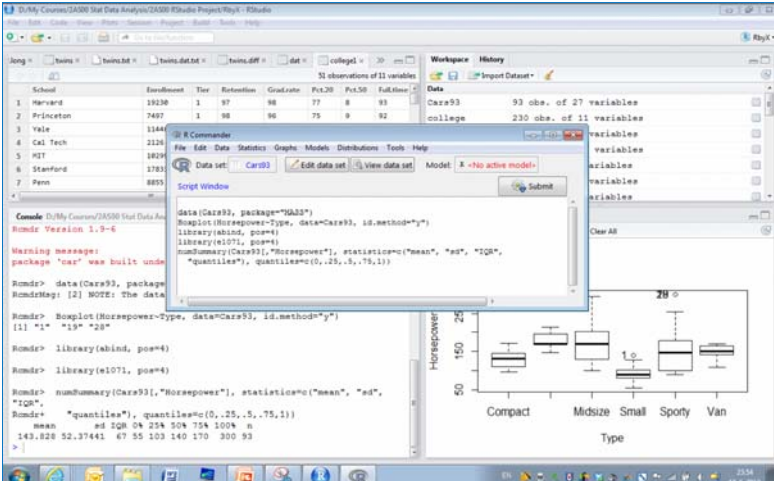
John Verzani

<http://cran.r-project.org/web/packages/pmg/vignettes/pmg.pdf>



R Studio
<http://www.rstudio.com/>

• RStudio + R Commander:

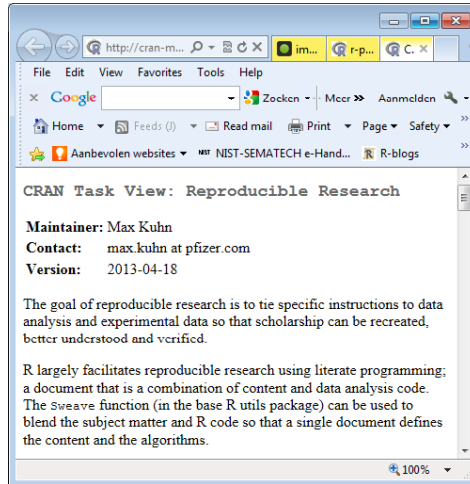


```

Rcmdr> data(Cars93, package="ISL3")
Rcmdr> boxplot(Horsepower~Type, data=Cars93, id.method="y")
Rcmdr> library(skimr, pos=4)
Rcmdr> library(e1071, pos=4)
Rcmdr> summary(Cars93[, "Horsepower"], statistics=c("mean", "sd", "IQR", "quantiles"), quantiles=c(0.25, .5, .75, 1))
Rcmdr> "quantiles"; quantiles=c(0.25, .5, .75, 1)
  mean      sd IQR  n
143.428  52.37442  67 55 103 140 170 300 93
  > |
    
```

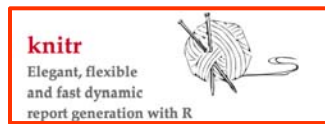
<http://www.rstudio.com/>

• Reproducible Research:



Sweave: Dynamic Generation of Statistical Reports Using Literate Data Analysis

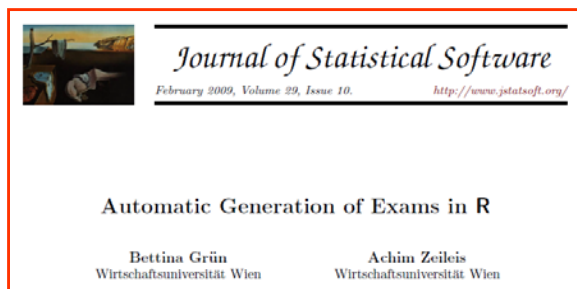
<http://www.stat.uni-muenchen.de/~leisch/Sweave/>



<http://yihui.name/knitr/>

<http://cran-mirror.cs.uu.nl/web/views/ReproducibleResearch.html>

• Automatic Generation of Exams:



<http://www.jstatsoft.org/v29/i10/paper>

Package 'exams'

June 6, 2013

Version 1.9-6

Date 2013-06-06

Title Automatic Generation of Exams in R

Description Sweave-based automatic generation of exams including multiple-choice questions and arithmetic problems. Exams can be produced in various formats, including PDF, HTML, Moodle XML, QTI 1.2 (for OLAT).

<http://cran-mirror.cs.uu.nl/web/packages/exams/exams.pdf>

Some References

- The  Manuals:



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

- Introduction to -Commander:

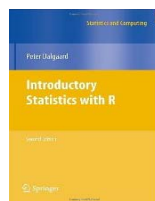


<http://socserv.mcmaster.ca/fox/Misc/Rcmdr/Getting-Started-with-the-Rcmdr.pdf>

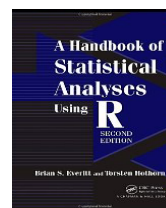
- Basic  :



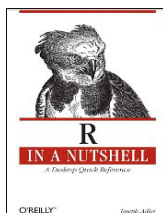
Mike Allerhand,
A Tiny Handbook of R



Peter Dalgaard,
Introductory Statistics using R



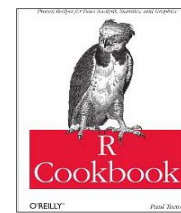
Brian Everitt et al.,
A Handbook of Statistical Analyses using R



Joseph Adler,
R in a Nutshell



<http://www.r-project.org/doc/bib/R-books.html>



Paul Teator,
R Cookbook

• Applied :

 Springer	springer.com
	<p>Use R!</p> <p>Series Editors: Gentleman, Robert, Hornik, Kurt, Parmigiani, Giovanni</p> <p>This series of inexpensive and focused books on R will publish shorter books aimed at practitioners. Books can discuss the use of R in a particular subject area (e.g., epidemiology, econometrics, psychometrics) or as it relates to statistical topics (e.g., missing data, longitudinal data). In most cases, books will combine LaTeX and R so that the code for figures and tables can be put on a website. Authors should assume a background as supplied by Dalgaard's Introductory Statistics with R or other introductory books so that each book does not repeat basic material.</p> <p>Recently published:</p>

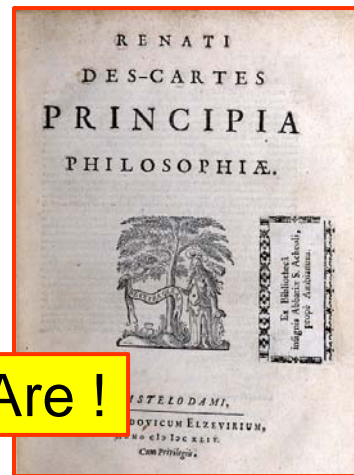
<http://www.springer.com/series/6991?detailsPage=titles>

Perspective.....

R ergo Es !



You , so You Are !



- Contact Details:



dr. J.J.M. (Koo) Rijkema
TU/e-Mathematics & Computer Science
Chair Probability & Statistics
Room: Metaforum MF 4.071A
Tel.: 040-2473170 (040-2472753)
E-mail j.j.m.rijpkema@tue.nl