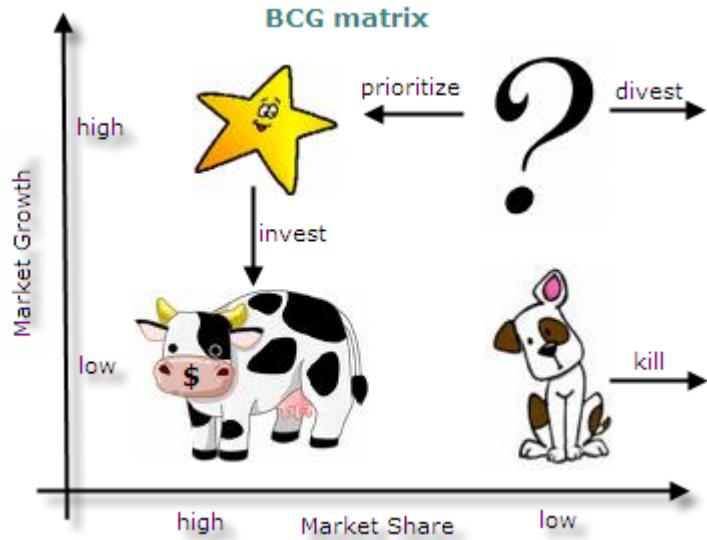


# **Business Analytics**

using the R language

# Strategy Refresher- Business Strategy Models

## BCG Matrix for Product Lines



# Strategy Refresher- Business Strategy Models

## Porter's Model for Industries



# Strategy Refresher- Business Strategy Models

## Grenier's Theory



# More Strategy Anyone?

<http://decisionstats.com/2013/12/19/business-strategy-models/>

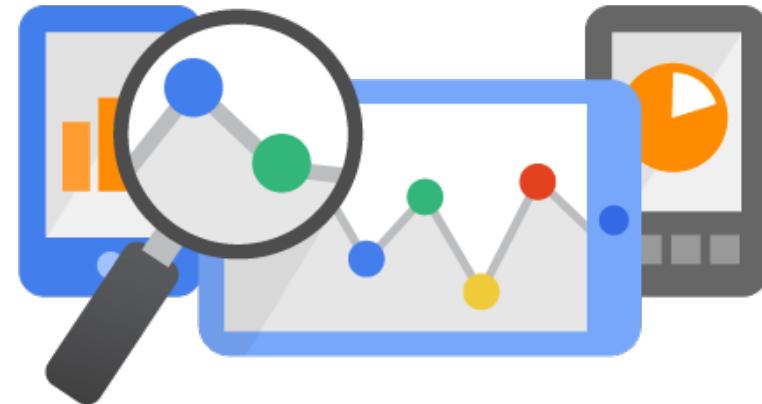
1. Porters 5 forces Model- To analyze industries
2. Business Canvas
3. BCG Matrix- To analyze Product Portfolios
4. Porters Diamond Model- To analyze locations
5. McKinsey 7 S Model- To analyze teams
6. Gernier Theory- To analyze growth of organization
7. Herzberg Hygiene Theory- To analyze soft aspects of individuals
8. Marketing Mix Model- To analyze marketing mix.

# Introduction

The R language is now the leading language for analytics and statistics on this planet. This R training starts with R language basics and covers basics of analytics

This course provides hands-on experience to execute analytics using the R language.

There will be many challenging tasks and focused practicals for the learners.



# Requirements

## Installations

R

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

RStudio

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

R Packages

rattle Rcmdr Deducer

# Episode 1

# Learning Objectives

- learn about R
- install R and it's packages

# **What will you learn from this lesson**

- Installation of R, Rtools, R Studio, R packages and GUIs
- Using RStudio and Using GUIs

# Data Driven Decision Making

- using data and trending historical data
- validating assumptions if any
- using champion challenger to test scenarios
- using experiments
- use baselines
- continuous improvement
  - customer experiences
  - costs
  - revenues

If you can't measure it, you can't manage it -Peter Drucker

# What is R



## The R environment

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes

- an effective data handling and storage facility,
- a suite of operators for calculations on arrays, in particular matrices,
- a large, coherent, integrated collection of intermediate tools for data analysis,
- graphical facilities for data analysis and display either on-screen or on hardcopy, and
- a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

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

# Statistical Software Landscape

SAS

Python (Pandas)

IBM SPSS

R

Julia

Clojure

Octave

Matlab

JMP

E views



# Using R with other software

<https://rforanalytics.wordpress.com/useful-links-for-r/using-r-from-other-software/>

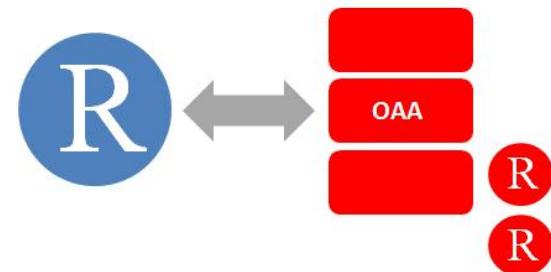
Tableau <http://www.tableausoftware.com/new-features/r-integration>

Qlik <http://qliksolutions.ru/qlikview/add-ons/r-connector-eng/>

Oracle R <http://www.oracle.com/technetwork/database/database-technologies/r/r-enterprise/overview/index.html>

Rapid Miner <https://rapid-i.com/content/view/202/206/lang,en/#r>

JMP <http://blogs.sas.com/jmp/index.php?/archives/298-JMP-Into-R!.html>



# Using R with other software

<https://rforanalytics.wordpress.com/useful-links-for-r/using-r-from-other-software/>

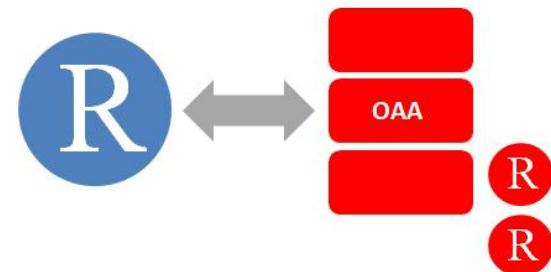
SAS/IML <http://www.sas.com/technologies/analytics/statistics/iml/index.html>

Teradata <http://developer.teradata.com/applications/articles/in-database-analytics-with-ter>

Pentaho <http://bigdatatechworld.blogspot.in/2013/10/integration-of-rweka-with-pentaho-data.html>

IBM SPSS [https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=ibm-analytics&S\\_PKG=ov18855&S\\_TACT=M161003W&dynform=127&lang=en\\_US](https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=ibm-analytics&S_PKG=ov18855&S_TACT=M161003W&dynform=127&lang=en_US)

TIBCO TERR <http://spotfire.tibco.com/discover-spotfire/what-does-spotfire-do/predictive-analytics/tibco-enterprise-runtime-for-r-terr>



# Some Advantages of R

open source

free

large number of algorithms and packages esp for statistics

flexible

very good for data visualization

superb community

rapidly growing

can be used with other software



# Some Disadvantages of R

- in memory (RAM) usage
- steep learning curve
- some IT departments frown on open source
- verbose documentation
- tech support
- evolving ecosystem for corporates



# Solutions for Disadvantages of R

- in memory (RAM) usage → specialized packages, in database computing
- steep learning curve → TRAINING !!!
- some IT departments frown on open source → TRAINING and education!
- verbose documentation → CRAN View , R Documentation
- tech support → expanding pool of resources
- evolving ecosystem for corporates → getting better with MS et al

# R used by Government

- In the early days of the [Deepwater Horizon disaster](#), NIST used uncertainty analysis in R to harmonize spill estimates from various sources, and to provide ranges of estimates to other agencies and the media.
- Before new drugs are allowed on the market, the FDA works with pharmaceutical companies to verify safety and efficacy through clinical trials. Despite a [false perception](#) that only commercial software may be used, many pharmaceutical companies are now using open-source R to [analyze data from clinical trials](#).
- The National Weather Service uses R for research and development of [models to predict river flooding](#).
- The newly-formed [Consumer Financial Protection Bureau](#) -- freed from the restrictions of a legacy IT infrastructure -- is championing the use of open-source technologies in government.
- Local governments are also building data-based applications. The SF Estuary Institute [uses R and Google Maps](#) to provide a [tool to track pollution](#) in the San Francisco Bay area.

[http://gsnmagazine.com/node/26483?c=cyber\\_security](http://gsnmagazine.com/node/26483?c=cyber_security)

# R used by Telecom

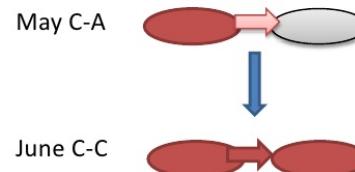
- Churn using

## Social Network Analysis

<http://www.slideshare.net/dataspora/social-network-a>

**Results: A Customer With a Canceller in Their Network Churns at Twice the Rate**

Types of Connections (Edges)



reality	expected by chance	delta
X	Y	2.0

In essence, we are asking whether being connected to another canceller has any effect on one's rate of cancellation. It turns out that it does.

And if we only look at voluntary port-outs, we see that customers churn at 3x the rate.

# R used by Insurance

a few more insurance related packages:

- [ChainLadder](#) – Reserving methods in R. The package provides Mack-, Munich-, Bootstrap, and Multivariate-chain-ladder methods, as well as the LDF Curve Fitting methods of Dave Clark and GLM-based reserving models.
- [cplm](#) – Monte Carlo EM algorithms and Bayesian methods for fitting Tweedie compound Poisson linear models
- [lossDev](#) – A Bayesian time series loss development model. Features include skewed-t distribution with time-varying scale parameter, Reversible Jump MCMC for determining the functional form of the consumption path, and a structural break in this path; by Christopher W. Laws and Frank A. Schmid
- [actuar](#): Loss distributions modelling, risk theory (including ruin theory), simulation of compound hierarchical models and credibility theory check out the [actuar](#) package by C. Dutang, V. Goulet and M. Pigeon.
- [favir](#): Formatted Actuarial Vignettes in R. FAViR lowers the learning curve of the R environment. It is a series of peer-reviewed Sweave papers that use a consistent style.
- [mondate](#): R packackge to keep track of dates in terms of months
- [lifecontingencies](#) – Package to perform actuarial evaluation of life contingencies

and

[Introduction to R for Actuaries](#) by Nigel de Silva

and <http://www.rininsurance.com/>

# R in Finance

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

R/Finance    [home](#)    [agenda](#)    [register](#)    [travel](#)    [committee](#)

**Friday, May 29th, 2015**

08:00 - 09:00 Optional Pre-Conference Tutorials

Ross Bennett: PortfolioAnalytics: Advanced Moment Estimation & Optimization ([pdf](#))

Kris Boudt: High-frequency Price Data Analysis in R ([pdf](#))

Dirk Eddelbuettel: Hands-on Introduction to Rcpp ([pdf](#))

Guy Yollin: Getting Started with Quantstrat

Maria Belianina: An Introduction to OneTick

09:00 - 09:30 Registration (2nd floor Inner Circle) & Continental Breakfast (3rd floor by Sponsor Tables)

Transition between seminars

09:30 - 09:35 Kickoff

09:35 - 09:40 Sponsor Introduction

09:40 - 10:30 **Emanuel Derman:** Understanding the World

10:30 - 10:54 **John Burkett:** Portfolio Optimization: Price Predictability, Utility Functions, Computational Methods, and Applications ([pdf](#))

**Kyle Balkissoon:** A Framework for Integrating Portfolio-level Backtesting with Price and Quantity Information ([html](#))

**Anthoney Tsou:** Implementation of Quality Minus Junk

**Ilya Kipnis:** Flexible Asset Allocation With Stepwise Correlation Rank ([pptm](#))

10:54 - 11:20 Break

11:20 - 11:40 **Sanjiv Das:** Efficient Rebalancing of Taxable Portfolios ([pdf](#))

11:40 - 12:00 **Marjan Wauters:** Characteristic-based equity portfolios: economic value and dynamic style allocation ([pdf](#))

12:00 - 12:20 **Bernhard Pfaff:** The sequel of cccp: Solving cone constrained convex programs

12:20 - 13:40 Lunch

13:40 - 14:00 **Markus Gesmann:** Communicating risk - a perspective from an insurer ([pdf](#))

14:00 - 14:20 **Doug Martin:** Nonparametric vs Parametric Shortfall: What are the Differences?

# R in Finance

<http://cran.r-project.org/web/views/Finance.html>

This CRAN Task View contains a list of packages useful for empirical work in Finance, grouped by topic.

- The Rmetrics suite of packages comprises [fArma](#), [fAsianOptions](#), [fAssets](#), [fBasics](#), [fBonds](#), [timeDate](#) (formerly: fCalendar), [fCopulae](#), [fExoticOptions](#), [fExtremes](#), [fGarch](#), [fImport](#), [fNonlinear](#), [fOptions](#), [fPortfolio](#), [fRegression](#), [timeSeries](#) (formerly: fSeries), [fTrading](#), [fUnitRoots](#) and contains a very large number of relevant functions for different aspect of empirical and computational finance.
- The [RQuantLib](#) package provides several option-pricing functions as well as some fixed-income functionality from the QuantLib project to R.
- The [quantmod](#) package offers a number of functions for quantitative modelling in finance as well as data acquisition, plotting and other utilities.
- The [portfolio](#) package contains classes for equity portfolio management; the [portfolioSim](#) builds a related simulation framework. The [backtest](#) offers tools to explore portfolio-based hypotheses about financial instruments. The [stockPortfolio](#) package provides functions for single index, constant correlation and multigroup models. The [pa](#) package offers performance attribution functionality for equity portfolios.
- The [PerformanceAnalytics](#) package contains a large number of functions for portfolio performance calculations and risk management.

# R in Pharma

<http://blog.revolutionanalytics.com/2013/08/r-drug-development-and-the-fda.html>

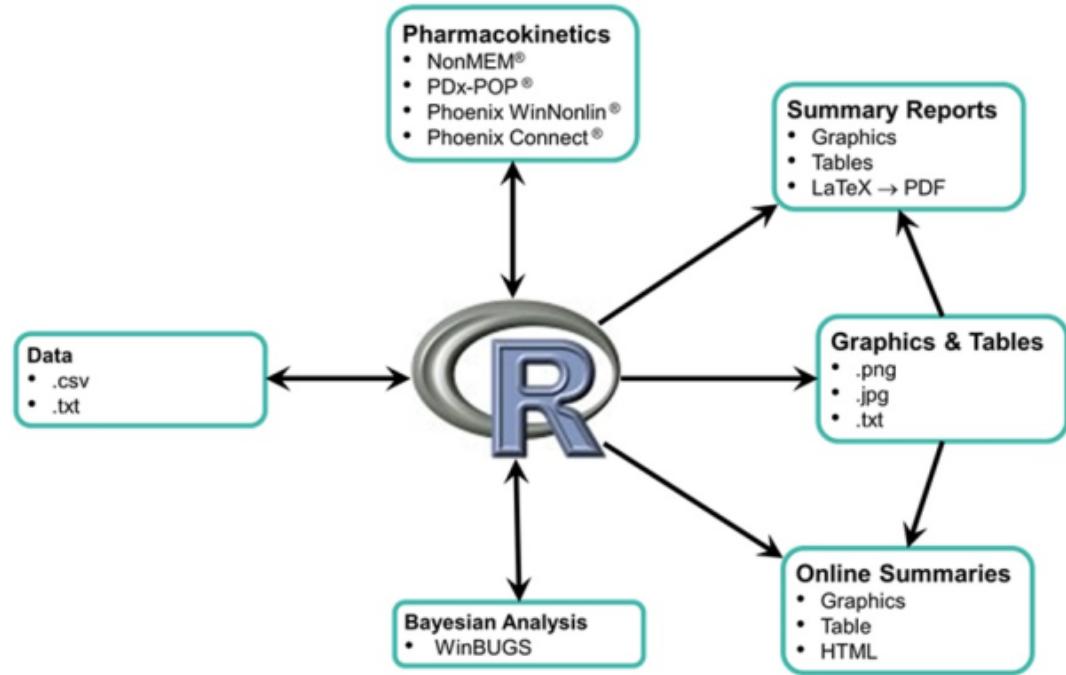
[Opening the Doors to Open Source Programming in Drug Development.](#)

[R: Regulatory Compliance and Validation Issues A Guidance Document for the Use of R in Regulated Clinical Trial Environments](#) in which he concluded that useR 2012 FDA statistician Jea Brodsky presented a [poster](#) described how FDA scientists “use R on a daily basis” and have themselves written R packages for use at various stages in the drug submission process.

[Open Source Software in the Biopharma Industry: Challenges and Opportunities,](#)

# R in Pharma

<http://web.quanticate.com/bid/102741/Using-the-Statistical-Programming-Language-R-in-the-Pharma-Industry>



# R in Pharma

<http://cran.r-project.org/web/views/ClinicalTrials.html>

This task view gathers information on specific R packages for design, monitoring and analysis of data from clinical trials. It focuses on including packages for clinical trial design and monitoring in general plus data analysis packages for a specific type of design.

## Design and Monitoring

- [TrialSize](#) This package has more than 80 functions from the book *Sample Size Calculations in Clinical Research* (Chow & Wang & Shao, 2007, 2nd ed., Chapman & Hall CRC).
- [asd](#) This Package runs simulations for adaptive seamless designs using early outcomes for treatment selection.
- [bcm](#) This package implements a wide variety of one and two-parameter Bayesian CRM designs. The program can run interactively, allowing the user to enter outcomes after each cohort has been recruited, or via simulation to assess operating characteristics.
- [blockrand](#) creates randomizations for block random clinical trials. It can also produce a PDF file of randomization cards.
- [contDesign](#) This small package contains a series of simple tools for constructing and manipulating confounded and fractional factorial designs.
- [CRTsize](#) This package contains basic tools for the purpose of sample size estimation in cluster (group) randomized trials. The package contains traditional power-based methods, empirical smoothing (Rotondi and Donner, 2009), and updated meta-analysis techniques (Rotondi and Donner, 2011).
- [dfrm](#) This package provides functions to run the CRM and TITE-CRM in phase I trials and calibration tools for trial planning purposes.
- [experiment](#) contains tools for clinical experiments, e.g., a randomization tool, and it provides a few special analysis options for clinical trials.
- [FFFD](#) This package creates regular and non-regular Factorial Factorial designs. Furthermore, analysis tools for Fractional Factorial designs with 2-level factors are offered (main effects and interaction plots for all factors simultaneously, cube plot for looking at the simultaneous effects of three factors, full or half normal plot, alias structure in a more readable format than with the built-in alias function alias). The package is currently subject to intensive development. While much of the intended functionality is already available, some changes and improvements are still to be expected.
- [GroupSeq](#) performs computations related to group sequential designs via the alpha spending approach, i.e., interim analyses need not be equally spaced, and their number need not be specified in advance.
- [gsDesign](#) derives group sequential designs and describes their properties.
- [lbd](#) and [fmin](#) computes and plots group sequential stopping boundaries from the Lan-DeMets method with a variety of a-spending functions using the ld98 program from the Department of Biostatistics, University of Wisconsin written by DM Rebbapragada, DL DeMets, KM Kim, and KKG Lan.
- [ldbounds](#) uses Lan-DeMets Method for group sequential trial, its functions calculate bounds and probabilities of a group sequential trial.
- [longpower](#) The longpower package contains functions for computing power and sample size for linear models of longitudinal data based on the formula due to Liu and Liang (1997) and Diggle et al (2002). Either formula is expressed in terms of marginal model or Generalized Estimating Equations (GEE) parameters. This package contains functions which translate pilot mixed effect model parameters (e.g. random intercept and/or slope) into marginal model parameters so that the formulas of Diggle et al or Liu and Liang formula can be applied to produce sample size calculations for two sample longitudinal designs assuming known variance.
- [PIDS](#) generates predicted interval plots, simulates and plots confidence intervals of an effect estimate given observed data and a hypothesis about the distribution of future data.
- [PowerTOST](#) contains functions to calculate power and sample size for various study designs used for bioequivalence studies. See function known.designs() for study designs covered. Moreover the package contains functions for power and sample size based on 'expected' power in case of uncertain (estimated) variability. Added are functions for the power and sample size for the ratio of two means with normally distributed data on the original scale (based on Fieller's confidence ('fiducial') interval).
- [pwr](#) has power analysis functions along the lines of Cohen (1988).
- [PwrGSD](#) is a set of tools to compute power in a group sequential design.
- [qtlDesign](#) provides tools for the design of QTL experiments.
- [sepmem](#) is computes the probability of crossing sequential efficacy and futility boundaries in a clinical trial. It implements the Armitage-McPherson and Rowe Algorithm using the method described in Schoenfeld (2001).

## Design and Analysis

- Package [AGSDes](#) This package provides tools and functions for parameter estimation in adaptive group sequential trials.
- Package [clintutup](#) has functions for both design and analysis of clinical trials. For phase II trials, it has functions to calculate sample size, effect size, and power based on Fisher's exact test, the operating characteristics of a two-stage boundary. Optimal and Minimax 2-stage Phase II designs given by Richard Simon, the exact 1-stage Phase II design and can compute a stopping rule and its operating characteristics for toxicity monitoring based repeated significance testing. For phase III trials, it can calculate sample size for group sequential designs.

# Companies using R

from <http://www.revolutionanalytics.com/companies-using-r>

ANZ, the fourth largest bank in Australia, using R for credit risk analysis

Bank of America uses R for reporting.

The Consumer Financial Protection Bureau uses R for data analysis.

Facebook

Facebook and R:

- Analysis of Facebook Status Updates
- Facebook's Social Network Graph
- How Google and Facebook are using R
- Predicting Colleague Interactions with R

# Pre Requisites

- Installation of R

<http://cran.rstudio.com/bin/windows/base/>



CRAN  
Mirrors  
What's new?  
Task Views  
Search

About R  
R Homepage  
The R Journal

Software  
R Sources  
Binary  
Packages  
Other

Documentation  
Manuals  
FAQs  
Contributed

R-3.1.1 for Windows (32/64 bit)

[Download R 3.1.1 for Windows](#) (54 megabytes, 32/64 bit)

[Installation and other instructions](#)  
[New features in this version](#)

If you want to double-check that the package you have downloaded exactly matches the package distributed by R, you can compare the [md5sum](#) of the .exe to the [true fingerprint](#). You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Frequently asked questions

- How do I install R when using Windows Vista?
- How do I update packages in my previous version of R?
- Should I run 32-bit or 64-bit R?

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Other builds

- Patches to this release are incorporated in the [r-patched.snapshot.build](#)
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel.snapshot.build](#)
- [Previous releases](#)

Note to webmasters: A stable link which will redirect to the current Windows binary release is  
[CRAN MIRROR - bin/windows/base/release.htm](#)

Last change: 2014-07-10, by Duncan Murdoch

- R Studio

- R Packages

# Pre Requisites

- Installation of R
  - Rtools
    - <http://cran.rstudio.com/bin/windows/Rtools/>
- R Studio
- R Packages



[CRAN](#)  
[Mirrors](#)  
[What's new?](#)  
[Data Views](#)  
[Search](#)

[About R](#)  
[R Home page](#)  
[R Journal](#)

[Software](#)  
[R Sources](#)  
[R Binaries](#)  
[Binaries](#)  
[Other](#)  
[Documentation](#)  
[Manuals](#)  
[FAQs](#)  
[Contributed](#)

## Building R for Windows

This document is a collection of resources for building packages for R under Microsoft Windows, or for building R itself (version 1.9.0 or later). The original collection was put together by Prof. Brian Ripley; it is currently being maintained by Duncan Murdoch.

The authoritative source of information for tools to work with the current release of R is the "R Administration and Installation" manual. In particular, please read the "Windows Toolkit" appendix.

### Tools Download

With the change to gcc 4.2.1, some of the tools for 3.2 bit compiles became incompatible with obsolete versions of R. Since then we have been maintaining one actively updated version of the tools, and other "frozen" snapshots of them. We recommend that users use the latest release of Rtools with the latest release of R.

The current version of this file is recorded here: [VERSION.txt](#)

Download	R compatibility	Frozen?
<a href="#">Rtools31.exe</a>	R 3.0.x to 3.1.x	No
<a href="#">Rtools30.exe</a>	R 3.0.1 to R 3.0.x	Yes
<a href="#">Rtools29.exe</a>	R 2.14.x to R 2.15.1	Yes
<a href="#">Rtools214.exe</a>	R 2.13.x or R 2.14.x	Yes
<a href="#">Rtools213.exe</a>	R 2.13.x	Yes
<a href="#">Rtools212.exe</a>	R 2.12.x	Yes
<a href="#">Rtools211.exe</a>	R 2.10.x or R 2.11.x	Yes
<a href="#">Rtools210.exe</a>	R 2.9.x or 2.10.x	Yes
<a href="#">Rtools209.exe</a>	R 2.8.x or 2.9.x	Yes
<a href="#">Rtools208.exe</a>	R 2.7.x or R 2.8.x	Yes
<a href="#">Rtools207.exe</a>	R 2.6.x or R 2.7.x	Yes
<a href="#">Rtools206.exe</a>	R 2.6.x, R 2.5.x or (untested) earlier	Yes

The change history to the Rtools is [below](#).

### Tools for 64 bit Windows builds

Rtools 2.12 and later include both 32 bit and 64 bit tools.

# Pre Requisites

- Installation of R

- RTools

The screenshot shows the RStudio website's download section. At the top, there are navigation links for Products, Resources, Pricing, About Us, and Blog, along with a search icon. Below this, a banner says "Download RStudio". To the right, there's a sidebar with a link to "Do you need support or a commercial license? Check out our commercial offerings". The main content area features a link to "Download RStudio Desktop v0.98.1074 — Release Notes". It notes that RStudio requires R 2.11.1 (or higher). Below this, a "Installers for ALL Platforms" table lists four download links:

Installers	Size	Date	MD5
RStudio 0.98.1074 - Windows XP/Vista/7/8	49 MB	2014-10-14	74d7bc76ec04287fac79cdd8dfaa8dd
RStudio 0.98.1074 - Mac OS X 10.6+ (64-bit)	38.4 MB	2014-10-14	f01c43fa29af679400c0faeae7ee33fb
RStudio 0.98.1074 - Debian 6+/Ubuntu 10.04+ (32-bit)	54.3 MB	2014-10-14	759d86599b22b28202a5aa0025e77278
RStudio 0.98.1074 - Debian 6+/Ubuntu 10.04+ (64-bit)	66.1 MB	2014-10-14	077e51c714a7df28ffcc5b9732404ab8

- R Studio

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

- R Packages

# Pre Requisites

- Installation of R
  - RTools

- R Studio

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

- R Packages
- about eight packages supplied with the R distribution and many more are available through the CRAN family of Internet sites covering a very wide range of modern statistics.

The screenshot shows the R Documentation homepage. At the top, there's a search bar with placeholder text "Start searching the documentation". Below it is a "TASK VIEWS" sidebar with a list of categories like Bayesian, ChemPhys, ClinicalTrials, Cluster, DifferentialEquations, Distributions, Econometrics, Environmetrics, ExperimentalDesign, Finance, Genetics, gR, Graphics, HighPerformanceComputing, MachineLearning, MedicalImaging, MetaAnalysis, Multivariate, NaturalLanguageProcessing, NumericalMathematics, OfficialStatistics, Optimization, Pharmacokinetics, Phylogenetics, Psychometrics, ReproducibleResearch, Robust, SocialSciences, Spatial, and SpatioTemporal. The main content area has a header "R Documentation" and a sub-header "Search the R documentation of 7128 R packages and 145117 R functions:". It features logos for CRAN, Bioconductor, and GitHub. Below the header are search fields for "All Fields", "Package Name", "Function Name", "Title", "Description", and "Author(s)". A green "Start search" button is at the bottom of these fields. To the right, there are sections for "Top Ranked CRAN Packages" (listing Rcpp, ggplot2, plyr, stringr, digest, RColorBrewer, reshape2, colorspace, labeling, scales) and "New Packages" (listing CP, mmnormpow, mblock, NB, BSGW, stabs, CEC, mvprpb, ENIRG, mdstdt). Navigation links for Discussion, About, and documentation package are at the top right.

# CRAN

107 sites in 49 regions



[CRAN](#)  
[Mirrors](#)  
[What's new?](#)  
[Task Views](#)  
[Search](#)

[About R](#)  
[R Homepage](#)  
[The R Journal](#)

[Software](#)  
[R Sources](#)  
[R Binaries](#)  
[Packages](#)  
[Other](#)

[Documentation](#)  
[Manuals](#)  
[FAQs](#)  
[Contributed](#)

## CRAN Mirrors

The Comprehensive R Archive Network is available at the following URLs, please choose a location close to you. Some statistics on the status of the mirrors can be found here: [main page](#), [windows release](#), [windows old release](#).

### 0-Cloud

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

### Algeria

<http://cran.usthb.dz/>

### Argentina

<http://mirror.fcaglp.unlp.edu.ar/CRAN/>

### Australia

<http://cran.csiro.au/>

<http://cran.ms.unimelb.edu.au/>

### Austria

<http://cran.at.r-project.org/>

### Belgium

<http://www.freestatistics.org/cran/>

### Brazil

<http://nbcgib.uesc.br/mirrors/cran/>

<http://cran-r.c3sl.ufpr.br/>

<http://cran.fiocruz.br/>

<http://www.vps.fmvz.usp.br/CRAN/>

<http://brieger.esalq.usp.br/CRAN/>

### Canada

<http://cran.stat.sfu.ca/>

<http://mirror.its.dal.ca/cran/>

<http://cran.ustat.utoronto.ca/>

<http://cran.skazkaforyou.com/>

<http://cran.parentingamerica.com/>

### Chile

<http://dirichlet.mat.puc.cl/>

### China

<http://ftp.ctex.org/mirrors/CRAN/>

<http://mirror.bjtu.edu.cn/cran/>

<http://mirrors.opencas.cn/cran/>

<http://crandoc.csail.mit.edu/~crandoc/CRAN/>

Rstudio, automatic redirection to servers worldwide



University of Science and Technology Houari Boumediene

Universidad Nacional de La Plata

CSIRO

University of Melbourne

Wirtschaftsuniversitaet Wien

K.U.Leuven Association

Center for Comp. Biol. at Universidade Estadual de Santa Cruz

Universidade Federal do Parana

Oswaldo Cruz Foundation, Rio de Janeiro

University of Sao Paulo, Sao Paulo

University of Sao Paulo, Piracicaba

Simon Fraser University, Burnaby

Dalhousie University, Halifax

University of Toronto

iWeb, Montreal

iWeb, Montreal

Pontificia Universidad Catolica de Chile, Santiago

CTEX.ORG

Beijing Jiaotong University, Beijing

Chinese Academy of Sciences, Beijing

TUMA Team, Tsinghua University

# Non CRAN Repositories

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

The screenshot shows the R Documentation homepage. At the top, there's a search bar with placeholder text "Start searching the documentation". Below it is a sidebar titled "TASK VIEWS" containing a long list of R package categories, each preceded by a blue triangle icon. The main content area has a header "R Documentation" and a sub-header "Search the R documentation of 7393 R packages and 150600 R functions:". A descriptive text below the sub-header encourages users to click the search bar or use the advanced search form. The main search form consists of six input fields: "All Fields", "Package Name", "Function Name", "Title", "Description", and "Author(s)". A large green "Start search" button is positioned at the bottom of the form. To the right of the search area, there's a vertical sidebar with advertisements. One ad from DataCamp promotes "Learn Data Science with R" for \$25/month, featuring a laptop icon with a gold medal. Another section at the bottom lists package sources: "Aggregating packages from: CRAN, Bioconductor, GitHub".

Start searching the documentation

TASK VIEWS

- Bayesian
- ChemPhys
- ClinicalTrials
- Cluster
- DifferentialEquations
- Distributions
- Econometrics
- Environmetrics
- ExperimentalDesign
- Finance
- Genetics
- gR
- Graphics
- HighPerformanceComputing
- MachineLearning
- MedicalImaging
- MetaAnalysis
- Multivariate
- NaturalLanguageProcessing
- NumericalMathematics
- OfficialStatistics
- Optimization
- Pharmacokinetics
- Phylogenetics
- Psychometrics
- ReproducibleResearch
- Robust
- SocialSciences
- Spatial
- SpatioTemporal

R Documentation

Discussion About

Search the R documentation of 7393 R packages and 150600 R functions:

Rdocumentation is a tool that helps you easily find and browse the documentation of all current and some past packages on CRAN. Click on the search bar at the top left for instant search or fill out the forms below for advanced search!

All Fields

Package Name

Function Name

Title

Description

Author(s)

**Start search**

DataCamp

Learn Data Science with R

\$25/month

Data Manipulation, Data Visualization, R Programming, Big Data, and much more.

Discover All Courses

Aggregating packages from:

CRAN

Bioconductor  
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

GitHub

# github

<https://github.com/trending?l=R>

The screenshot shows the GitHub trending repositories page for the language R. The URL in the address bar is <https://github.com/trending?l=R>. The page title is "Trending repositories". A sidebar on the right lists various programming languages: All languages, Unknown languages, C, C++, HTML, Java, JavaScript, Python, and R, which is highlighted with a blue background. A "Languages" dropdown menu is also present. A "ProTip!" box suggests looking for most forked R repositories. The main content area displays four trending R repositories:

- rdpeng/ProgrammingAssignment2**  
Repository for Programming Assignment 2 for R Programming on Coursera  
R • Built by [Profile icons]
- qinwf/awesome-R**  
A curated list of awesome R frameworks, packages and software.  
R • 7 stars today • Built by [Profile icons]
- berndbischl/mlr**  
mlr: Machine Learning in R  
R • Built by [Profile icons]
- rstudio/shinyapps**  
[Profile icons]

At the top of the page, there is a navigation bar with links for Explore, Gist, Blog, Help, and decisionstats. Below the navigation bar, there are tabs for All, Showcases, Trending (which is selected), and Stars.

# bioconductor

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

The screenshot shows the Bioconductor website with a dark teal header bar. The header includes the Bioconductor logo (a stylized DNA helix icon) and the text "Bioconductor OPEN SOURCE SOFTWARE FOR BIOINFORMATICS". Below the header is a navigation menu with links for Home, Install, Help, Developers, and About. A search bar is located in the top right corner. The main content area features several sections: "BioC2015" (with a bioinformatics conference announcement), "About Bioconductor" (with a brief description of the project's purpose and history), "News" (with a list of recent updates), and four large cards for "Install", "Learn", "Use", and "Develop". Each card contains a summary and a bulleted list of resources.

**BioC2015**

Join us for morning talks from distinguished speakers and community members, afternoon workshops to hone your skills, and poster sessions and social activities to get to know members of the Bioconductor community at our [Annual Conference](#), July 20 (Developer Day), 21 and 22 in Seattle, WA.

**About Bioconductor**

Bioconductor provides tools for the analysis and comprehension of high-throughput genomic data. Bioconductor uses the R statistical programming language, and is open source and open development. It has two releases each year, [1024 software packages](#), and an active user community. Bioconductor is also available as an [AMI](#) (Amazon Machine Image) and a series of [Docker](#) images.

**News**

- Bioconductor [3.1](#) is available.
- Orchestrating high-throughput genomic analysis with [Bioconductor \(abstract\)](#) and other [recent literature](#).
- Read our latest [newsletter](#) and [course](#)

**Install »**

Get started with *Bioconductor*

- [Install \*Bioconductor\*](#)
- [Explore packages](#)
- [Get support](#)
- [Latest newsletter](#)
- [Follow us on twitter](#)
- [Install R](#)

**Learn »**

Master *Bioconductor* tools

- [Courses](#)
- [Support site](#)
- [Package vignettes](#)
- [Literature citations](#)
- [Common work flows](#)
- [FAQ](#)
- [Community resources](#)
- [Videos](#)

**Use »**

Create bioinformatic solutions with *Bioconductor*

- [Software, Annotation, and Experiment packages](#)
- [Amazon Machine Image](#)
- [Latest release announcement](#)
- [Support site](#)

**Develop »**

Contribute to *Bioconductor*

- [Use BioC 'devel'](#)
- ['Devel' Software, Annotation and Experiment packages](#)
- [Package guidelines](#)
- [New package submission](#)
- [Developer resources](#)
- [Build reports](#)

# Pre Requisites

- R Packages

`install.packages()` INSTALLS

`update.packages()` UPDATES

`library()` LOADS

- Packages are **installed** once, updated periodically, but **loaded** every time

# Interfaces to R

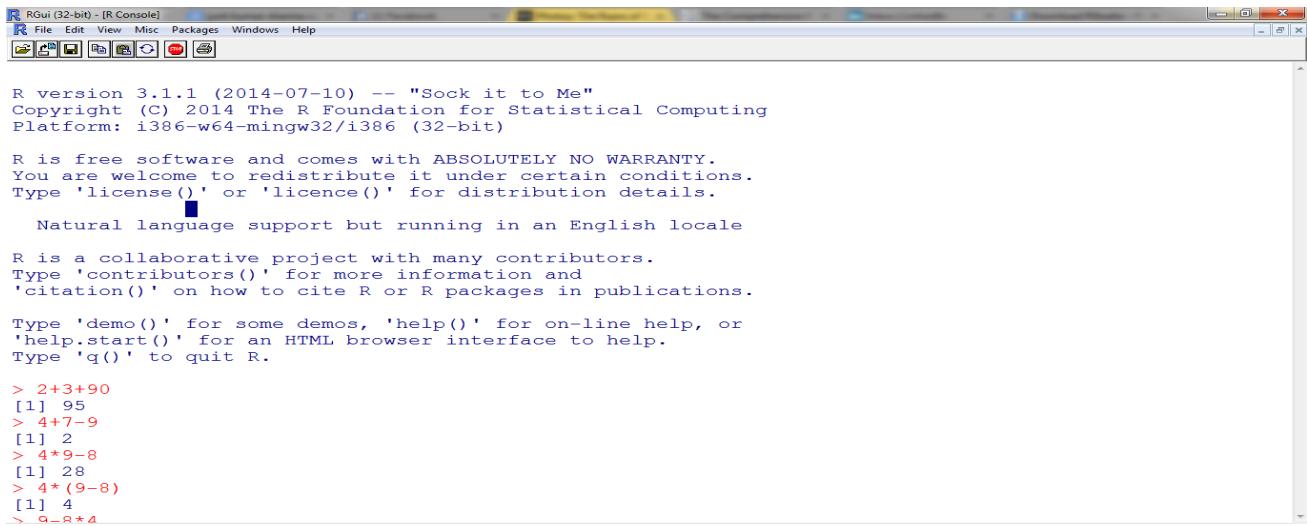
- Console

*Default*

*Customization*

- IDE

- GUI



R version 3.1.1 (2014-07-10) -- "Sock it to Me"  
Copyright (C) 2014 The R Foundation for Statistical Computing  
Platform: i386-w64-mingw32/i386 (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.

Natural language support but running in an English locale

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.

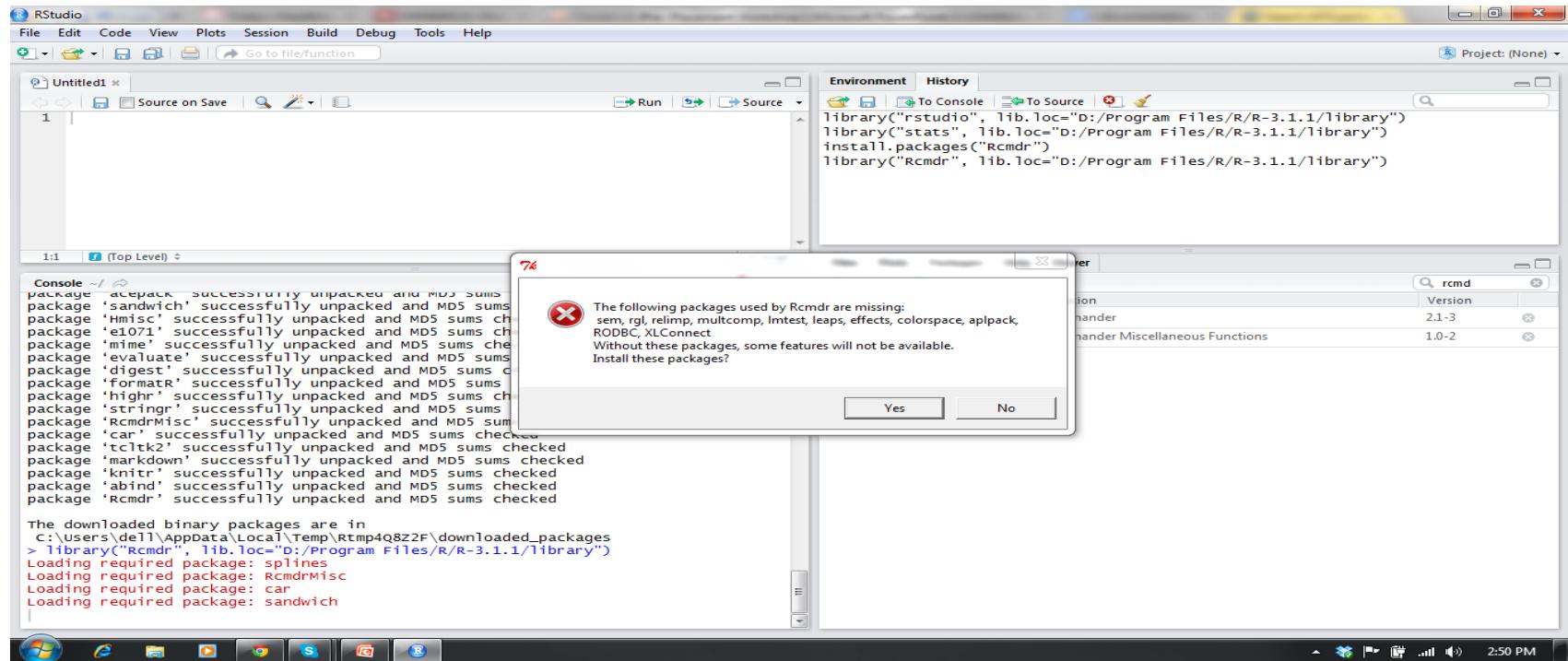
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

```
> 2+3+90
[1] 95
> 4+7-9
[1] 2
> 4*9-8
[1] 28
> 4*(9-8)
[1] 4
> a-a*4
```

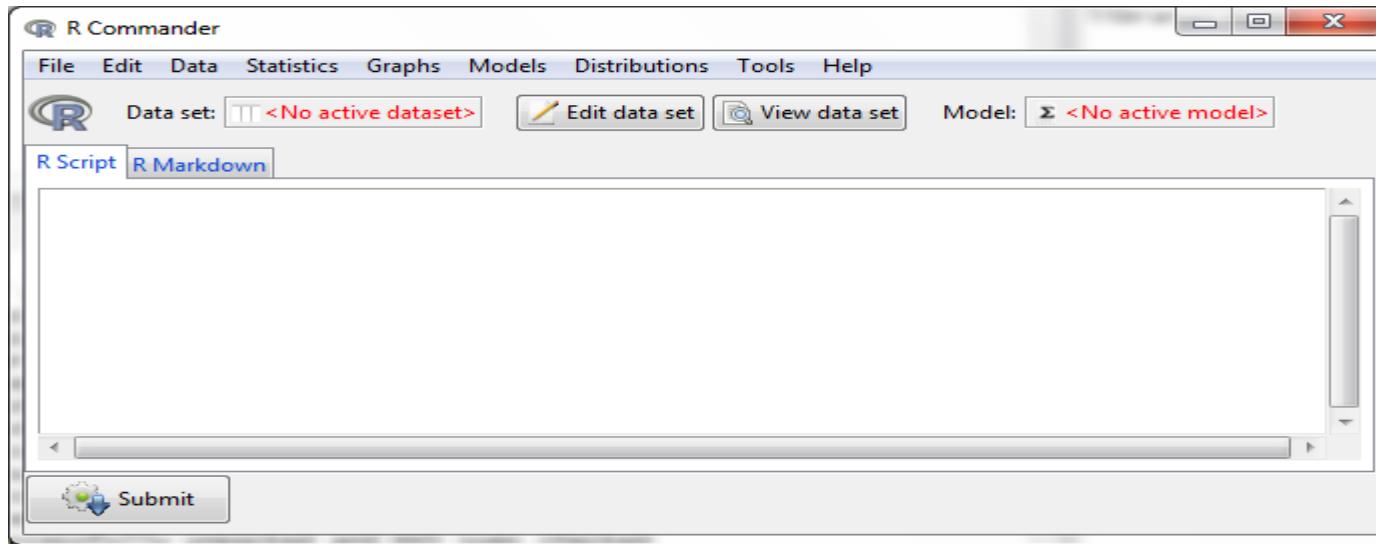
# Graphical Interfaces to R

- R Commander
- Rattle
- Deducer

# Installation of R Commander

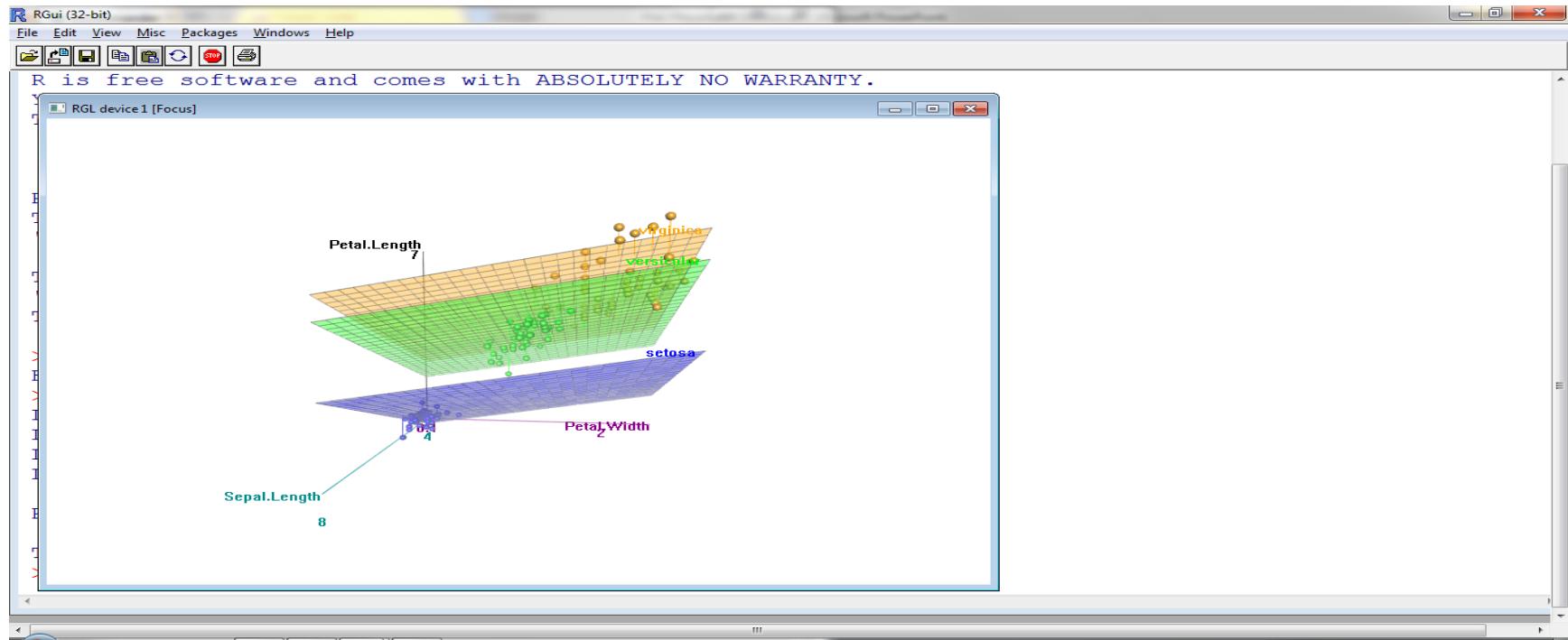


# Overview of R Commander

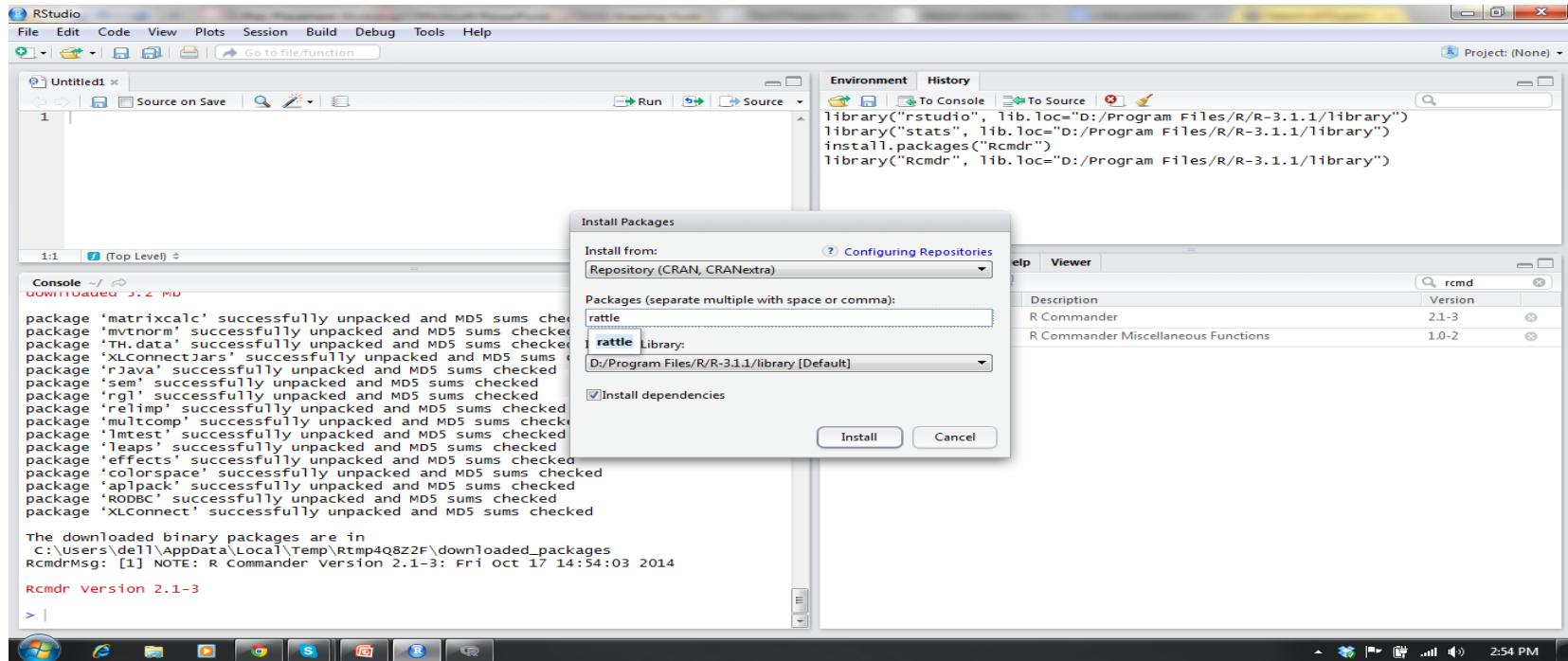


# Demo

## R Commander – 3D Graphs



# Installation of Rattle



# Installation of Rattle

The screenshot shows the RStudio interface with the following details:

- Top Bar:** RStudio, File, Edit, Code, View, Plots, Session, Build, Debug, Tools, Help.
- Project Bar:** Project: (None).
- Console Tab:** Untitled1 x, Source on Save, Run, Source.
- Console Output:**

```
library("rstudio", lib.loc="D:/Program Files/R/R-3.1.1/library")
library("stats", lib.loc="D:/Program Files/R/R-3.1.1/library")
install.packages("Rcmdr")
library("Rcmdr", lib.loc="D:/Program Files/R/R-3.1.1/library")
install.packages("rattle")
library("rattle", lib.loc="D:/Program Files/R/R-3.1.1/library")

1:1 [ Top Level ] R Script

Console ~ 
package 'rattle' successfully unpacked and MD5 sums checked
package 'leaps' successfully unpacked and MD5 sums checked
package 'effects' successfully unpacked and MD5 sums checked
package 'colorspace' successfully unpacked and MD5 sums checked
package 'alppack' successfully unpacked and MD5 sums checked
package 'RODBC' successfully unpacked and MD5 sums checked
package 'XLConnect' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:\Users\dell\AppData\Local\Temp\Rtmp4Q8Z2F\downloaded_packages
RcmdrMsg: [1] NOTE: R Commander Version 2.1-3: Fri Oct 17 14:54:03 2014
Rcmdr version 2.1-3
> install.packages("rattle")
trying URL 'http://cran.rstudio.com/bin/windows/contrib/3.1/rattle_3.3.0.zip'
Content type 'application/zip' length 3211375 bytes (3.1 Mb)
opened URL
downloaded 3.1 Mb
package 'rattle' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:\Users\dell\AppData\Local\Temp\Rtmp4Q8Z2F\downloaded_packages
> library("rattle", lib.loc="D:/Program Files/R/R-3.1.1/library")
```
- Environment Tab:** Shows the loaded packages: rattle, Rcmdr, stats, and rstudio.
- Packages Tab:** Shows the installed packages table.
- Bottom Status Bar:** Screenshot Added, A screenshot was added to your Dropbox.

# Installation of Rattle

The screenshot shows the RStudio interface with the following components:

- Top Bar:** RStudio, File, Edit, Code, View, Plots, Session, Build, Debug, Tools, Help.
- Console:** Shows the command-line output of the R session. It includes the download of the rattle package, its unpacking, and the resulting message: "The downloaded binary packages are in C:/Users/dell/AppData/Local/Temp/Rtmp4Q8Z2F/downloaded\_packages". It also shows the Rattle welcome message and a prompt asking if the user would like to install RGtk2.
- Code Editor:** Untitled1.R, showing the R code used to install the package: `library("rattle")` and `install.packages("rattle")`.
- Environment Tab:** Shows the loaded packages: rstudio, stats, Rcmdr, Rcmdr, rattle, and rattle.
- Packages Tab:** Shows a list of installed packages with their details:

Name	Description	Version
multcomp	Simultaneous Inference in General Parametric Models	1.3-7
mvtnorm	Multivariate Normal and t Distributions	1.0-0
nlme	Linear and Nonlinear Mixed Effects Models	3.1-117
nnet	Feed-Forward Neural Networks and Multinomial Log-Linear Models	7.3-8
parallel	Support for Parallel computation in R	3.1.1
<input checked="" type="checkbox"/> rattle	Graphical user interface for data mining in R	3.3.0
Rcmdr	R Commander	2.1-3
RcmdrMisc	R Commander Miscellaneous Functions	1.0-2
RColorBrewer	ColorBrewer palettes	1.0-5
relimp	Relative Contribution of Effects in a Regression Model	1.0-3
rgl	3D visualization device system (OpenGL)	0.94.1143
rJava	Low-level R to Java interface	0.9-6
RODBC	ODBC Database Access	1.3-10
rpart	Recursive Partitioning and Regression Trees	4.1-8
rstudio	Tools and Utilities for RStudio	0.98.1074
sandwich	Robust Covariance Matrix Estimators	2.3-2

# Installation of Rattle

The screenshot shows the RStudio interface with the following details:

- Console:** Displays the R command to install Rattle and its dependencies, followed by the download progress bar.
- Output:** Shows the successful download of the Rattle package and its dependencies.
- Progress Bar:** Shows "90% downloaded" with a progress bar indicating the download status.
- Message Box:** A modal window displays the URL for the download: "URL: ... //cran.rstudio.com/bin/windows/contrib/3.1/RGtk2\_2.20.31.zip".
- Environment View:** Shows the installed packages and their versions.
- Table:** A table of installed packages with their descriptions and versions.

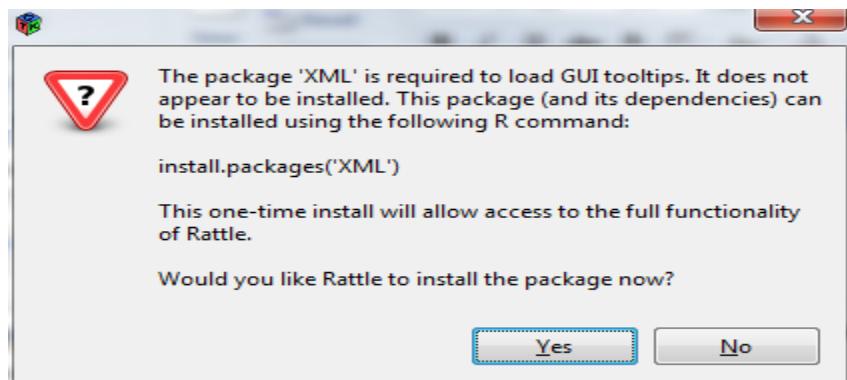
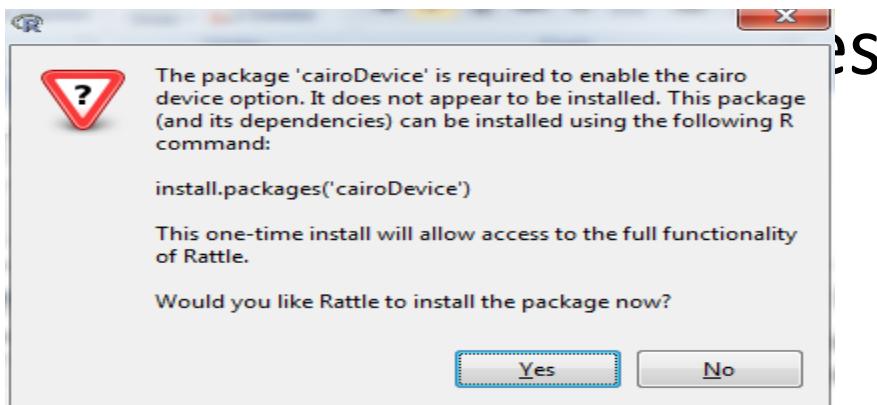
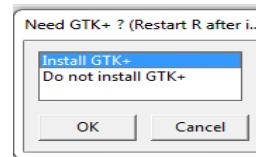
```
library("rstudio", lib.loc="D:/Program Files/R/R-3.1.1/library")
library("stats", lib.loc="D:/Program Files/R/R-3.1.1/library")
install.packages("Rcmdr")
library("Rcmdr", lib.loc="D:/Program Files/R/R-3.1.1/library")
install.packages("rattle")
```

package 'rattle' successfully downloaded  
The downloaded binary packages were...  
C:/Users/dell/AppData/Local/Temp/Rtmp4Q8Z2F/downloaded\_packages  
> library('rattle', lib.loc="D:/Program Files/R/R-3.1.1/library")  
Rattle: A free graphical interface for data mining with R.  
Version 3.3.0 copyright (c) 2006-2014 Togaware Pty Ltd.  
Type 'rattle()' to shake, rattle, and roll your data.  
> rattle()  
The package 'RGtk2' is required to display the Rattle GUI. It does not appear to be installed. This package (and its dependencies) can be installed using the following R command:  
install.packages('RGtk2')  
This one-time install will allow access to the full functionality of Rattle.  
Would you like Rattle to install the package now?  
(yes/no) yes  
trying URL 'http://cran.rstudio.com/bin/windows/contrib/3.1/RGtk2\_2.20.31.zip'  
Content type 'application/zip' length 13884133 bytes (13.2 Mb)  
opened URL

Package	Description	Version
nlme	Linear and Nonlinear Mixed Effects Models	3.1-117
nnet	Feed-Forward Neural Networks and Multinomial Log-Linear Models	7.3-8
parallel	Support for Parallel computation in R	3.1.1
<b>rattle</b>	Graphical user interface for data mining in R	3.3.0
Rcmdr	R Commander	2.1-3
RcmdrMisc	R Commander Miscellaneous Functions	1.0-2
RColorBrewer	ColorBrewer palettes	1.0-5
relimp	Relative Contribution of Effects in a Regression Model	1.0-3
rgl	3D visualization device system (OpenGL)	0.94.1143
rJava	Low-level R to Java interface	0.9-6
RODBC	ODBC Database Access	1.3-10
rpart	Recursive Partitioning and Regression Trees	4.1-8
rstudio	Tools and Utilities for RStudio	0.98.1074
sandwich	Robust Covariance Matrix Estimators	2.3-2

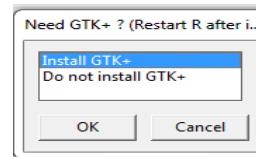
# Installation of Rattle

- GTK+ Installation Necessary



# Installation of Rattle

- GTK+ Installation Necessary

A screenshot of the Rattle application window showing four separate dependency installation dialogs. Each dialog has a red warning icon with a question mark and a message about a required package, an R command to install it, and a "Would you like Rattle to install the package now?" question with "Yes" and "No" buttons.

The package 'cairoDevice' is required to enable the cairo device option. It does not appear to be installed. This package (and its dependencies) can be installed using the following R command:

```
install.packages('cairoDevice')
```

This one-time install will allow access to the full functionality of Rattle.

Would you like Rattle to install the package now?

The package 'XML' is required to load GUI tooltips. It does not appear to be installed. This package (and its dependencies) can be installed using the following R command:

```
install.packages('XML')
```

This one-time install will allow access to the full functionality of Rattle.

Would you like Rattle to install the package now?

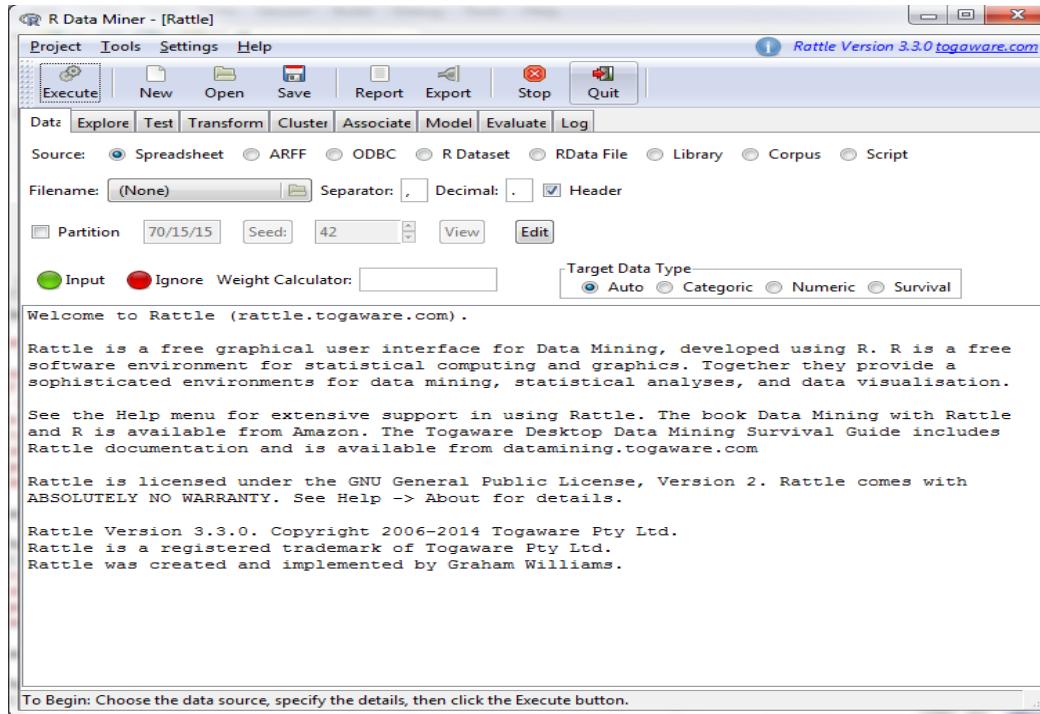
The package 'rpart.plot' is required to plot nice looking decision trees. It does not appear to be installed. This package (and its dependencies) can be installed using the following R command:

```
install.packages('rpart.plot')
```

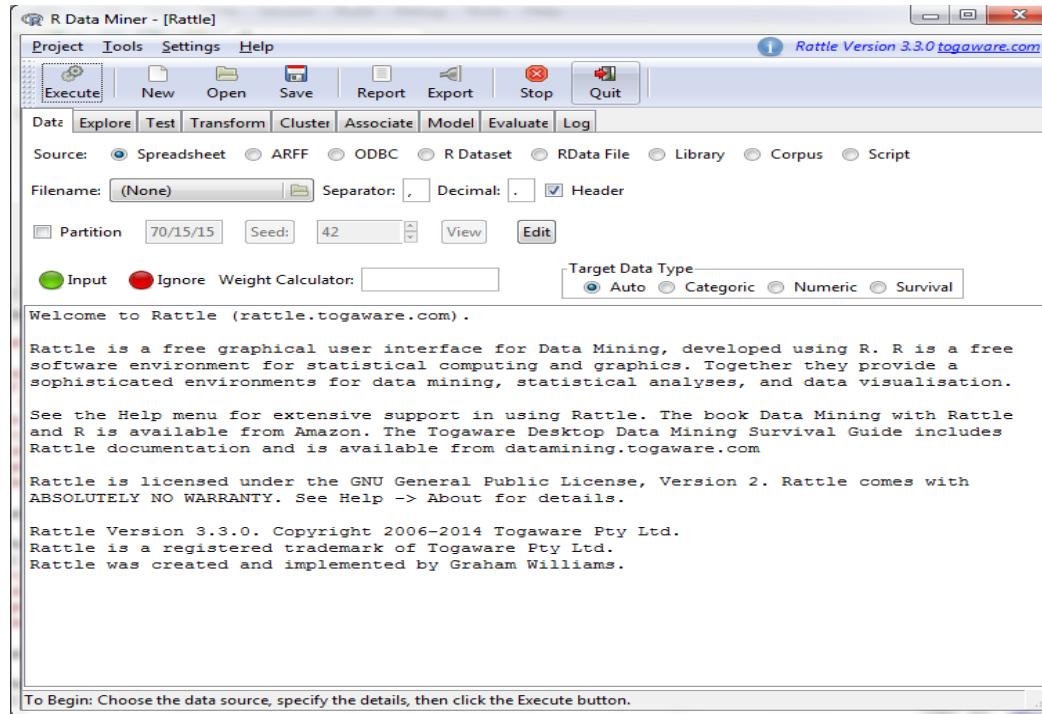
This one-time install will allow access to the full functionality of Rattle.

Would you like Rattle to install the package now?

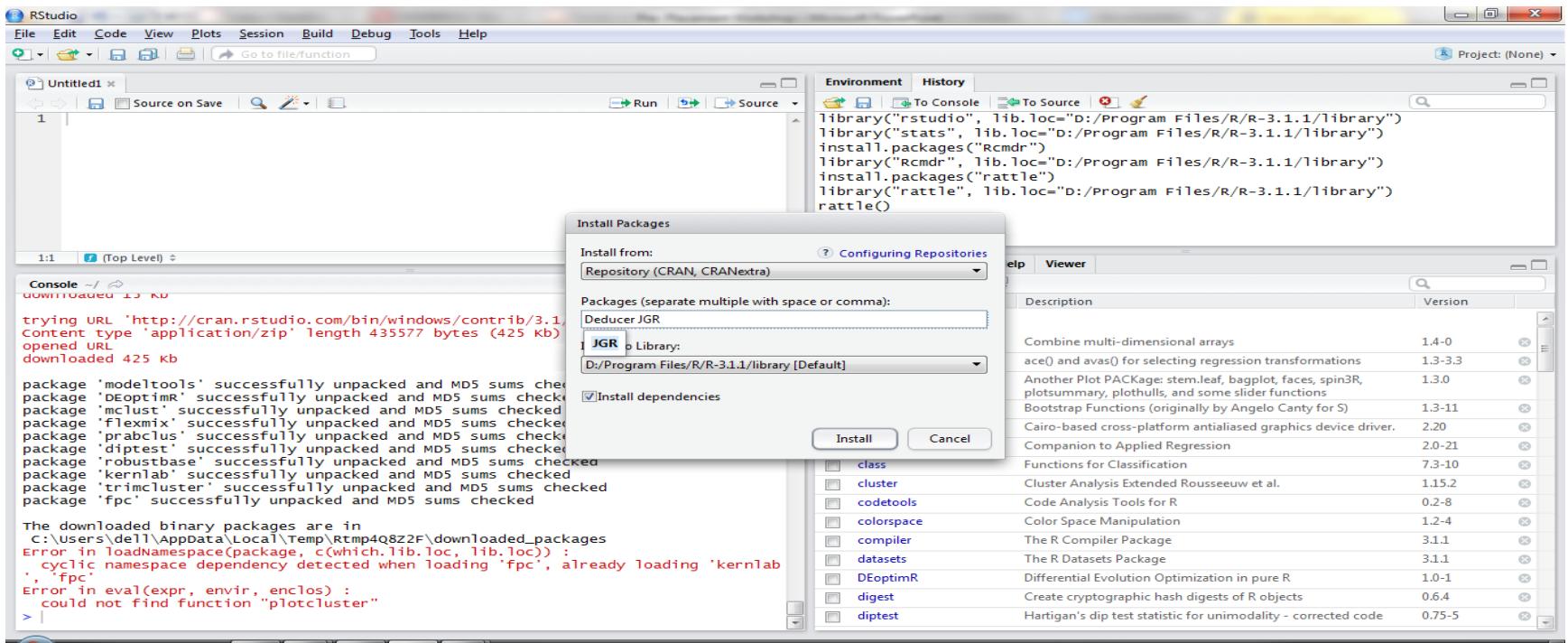
# Overview of Rattle



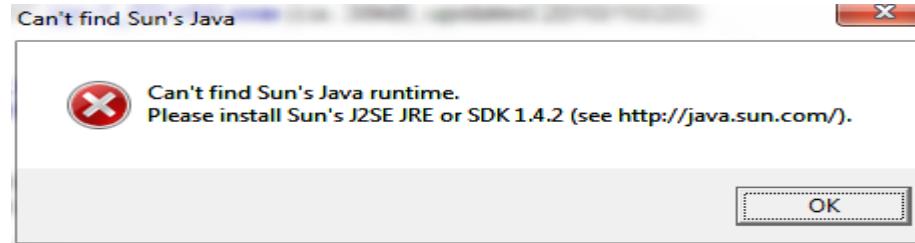
# Demo Rattle



# Installation Deducer (with JGR)



# Installation Deducer (with JGR)



# Installation Deducer (with JGR)

- [Java ME](#)
- [Java SE Support](#)
- [Java SE Advanced & Suite](#)
- [Java Embedded](#)
- [Java DB](#)
- [Web Tier](#)
- [Java Card](#)
- [Java TV](#)
- [New to Java](#)
- [Community](#)
- [Java Magazine](#)

## Java SE Runtime Environment 7 Downloads

Do you want to run Java™ programs, or do you want to develop Java programs? If you want to run Java programs, but not develop them, download the Java Runtime Environment, or JRE™.

If you want to develop applications for Java, download the Java Development Kit, or JDK™. The JDK includes the JRE, so you do not have to download both separately.

[7u71 JRE MD5 Checksum](#)  
[7u72 JRE MD5 Checksum](#)

**What is the difference between a Java CPU (7u71) and PSU (7u72) release?**

Java SE Critical Patch Updates (CPU) contain fixes to security vulnerabilities and critical bug fixes. Oracle strongly recommends that all Java SE users upgrade to the latest CPU releases as they are made available. Most user should choose this release.

Java SE Patch Set Updates (PSU) contain all of the security fixes in the CPUs released up to that version, as well as additional non-critical fixes. Java PSU releases should only be used if you are being impacted by one of the additional bugs fixed in that version.

[Visit Java CPU and PSU Releases Explained for details.](#)

### Java SE Runtime Environment 7u71

You must accept the Oracle Binary Code License Agreement for Java SE to download this software.

Thank you for accepting the Oracle Binary Code License Agreement for Java SE; you may now download this software.

Product / File Description	File Size	Download
Linux x86	31.58 MB	<a href="#">jre-7u71-linux-i586.rpm</a>
Linux x86	46.22 MB	<a href="#">jre-7u71-linux-i586.tar.gz</a>
Linux x64	32.1 MB	<a href="#">jre-7u71-linux-x64.rpm</a>
Linux x64	44.84 MB	<a href="#">jre-7u71-linux-x64.tar.gz</a>
Mac OS X x64	48.57 MB	<a href="#">jre-7u71-macosx-x64.dmg</a>
Mac OS X x64	44.52 MB	<a href="#">jre-7u71-macosx-x64.tar.gz</a>
Solaris x86	52.16 MB	<a href="#">jre-7u71-solaris-i586.tar.gz</a>
Solaris x64	16.14 MB	<a href="#">jre-7u71-solaris-x64.tar.gz</a>
Solaris SPARC	54.95 MB	<a href="#">jre-7u71-solaris-sparc.tar.gz</a>
Solaris SPARC 64-bit	18.1 MB	<a href="#">jre-7u71-solaris-sparcv9.tar.gz</a>
Windows x86 Online	0.89 MB	<a href="#">jre-7u71-windows-i586-ifw.exe</a>
Windows x86 Offline	28.09 MB	<a href="#">jre-7u71-windows-i586.exe</a>
Windows x86	40 MB	<a href="#">jre-7u71-windows-i586.tar.gz</a>
Windows x64	29.59 MB	<a href="#">jre-7u71-windows-x64.exe</a>
Windows x64	41.71 MB	<a href="#">jre-7u71-windows-x64.tar.gz</a>

[Java EE and Glassfish](#)

[Java ME](#)

[Java Card](#)

[NetBeans IDE](#)

[Java Mission Control](#)

**Java Resources**

[Java APIs](#)

[Technical Articles](#)

[Demos and Videos](#)

[Forums](#)

[Java Magazine](#)

[Java.net](#)

[Developer Training](#)

[Tutorials](#)

[Java.com](#)

**NEW!**

**Java magazine** Get it now for FREE!

Subscribe Today

Webcast  
Introducing Java 8



# Installation Deducer (with JGR)

- [Java ME](#)
- [Java SE Support](#)
- [Java SE Advanced & Suite](#)
- [Java Embedded](#)
- [Java DB](#)
- [Web Tier](#)
- [Java Card](#)
- [Java TV](#)
- [New to Java](#)
- [Community](#)
- [Java Magazine](#)

## Java SE Runtime Environment 7 Downloads

Do you want to run Java™ programs, or do you want to develop Java programs? If you want to run Java programs, but not develop them, download the Java Runtime Environment, or JRE™.

If you want to develop applications for Java, download the Java Development Kit, or JDK™. The JDK includes the JRE, so you do not have to download both separately.

[7u71 JRE MD5 Checksum](#)  
[7u72 JRE MD5 Checksum](#)

### What is the difference between a Java CPU (7u71) and PSU (7u72) release?

Java SE Critical Patch Updates (CPU) contain fixes to security vulnerabilities and critical bug fixes. Oracle strongly recommends that all Java SE users upgrade to the latest CPU releases as they are made available. Most user should choose this release.

Java SE Patch Set Updates (PSU) contain all of the security fixes in the CPUs released up to that version, as well as additional non-critical fixes. Java PSU releases should only be used if you are being impacted by one of the additional bugs fixed in that version.

[Visit Java CPU and PSU Releases Explained for details.](#)

### Java SE Runtime Environment 7u71

You must accept the Oracle Binary Code License Agreement for Java SE to download this software.

Thank you for accepting the Oracle Binary Code License Agreement for Java SE; you may now download this software.

Product / File Description	File Size	Download
Linux x86	31.58 MB	<a href="#">jre-7u71-linux-i586.rpm</a>
Linux x86	46.22 MB	<a href="#">jre-7u71-linux-i586.tar.gz</a>
Linux x64	32.1 MB	<a href="#">jre-7u71-linux-x64.rpm</a>
Linux x64	44.84 MB	<a href="#">jre-7u71-linux-x64.tar.gz</a>
Mac OS X x64	48.57 MB	<a href="#">jre-7u71-macosx-x64.dmg</a>
Mac OS X x64	44.52 MB	<a href="#">jre-7u71-macosx-i586.tar.gz</a>
Solaris x86	52.16 MB	<a href="#">jre-7u71-solaris-i586.tar.gz</a>
Solaris x64	16.14 MB	<a href="#">jre-7u71-solaris-x64.tar.gz</a>
Solaris SPARC	54.95 MB	<a href="#">jre-7u71-solaris-sparc.tar.gz</a>
Solaris SPARC 64-bit	18.1 MB	<a href="#">jre-7u71-solaris-sparcv9.tar.gz</a>
Windows x86 Online	0.89 MB	<a href="#">jre-7u71-windows-i586-iftw.exe</a>
Windows x86 Offline	28.09 MB	<a href="#">jre-7u71-windows-i586.exe</a>
Windows x86	40 MB	<a href="#">jre-7u71-windows-i586.tar.gz</a>
Windows x64	29.59 MB	<a href="#">jre-7u71-windows-x64.exe</a>
Windows x64	41.71 MB	<a href="#">jre-7u71-windows-x64.tar.gz</a>

[Java EE and Glassfish](#)

[Java ME](#)

[Java Card](#)

[NetBeans IDE](#)

[Java Mission Control](#)

### Java Resources

[Java APIs](#)

[Technical Articles](#)

[Demos and Videos](#)

[Forums](#)

[Java Magazine](#)

[Java.net](#)

[Developer Training](#)

[Tutorials](#)

[Java.com](#)

**NEW!**

**Java magazine** Get it now for FREE!

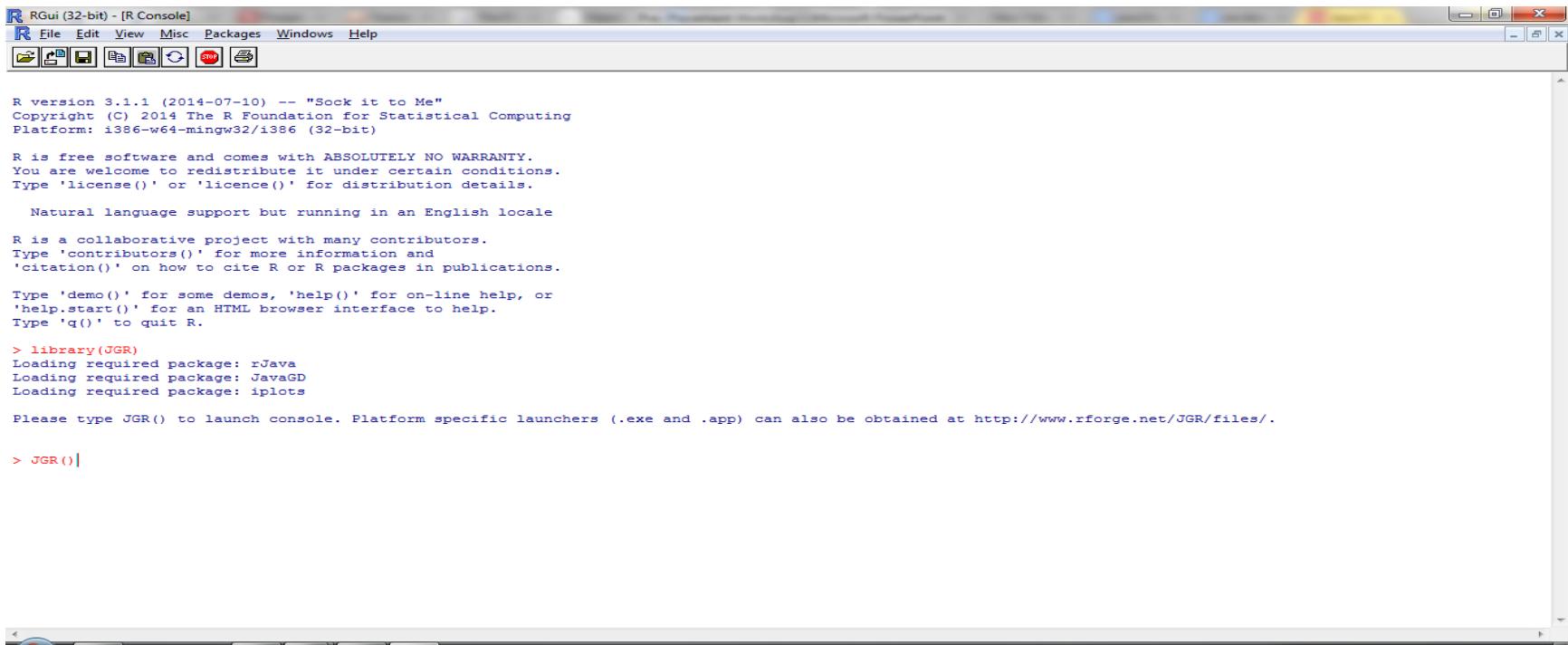
Subscribe Today

Webcast  
Introducing Java 8

Watch Now >



# Installation Deducer (with JGR)



R Gui (32-bit) - [R Console]

R File Edit View Misc Packages Windows Help

R version 3.1.1 (2014-07-10) -- "Sock it to Me"  
Copyright (C) 2014 The R Foundation for Statistical Computing  
Platform: i386-w64-mingw32/i386 (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.

Natural language support but running in an English locale

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.

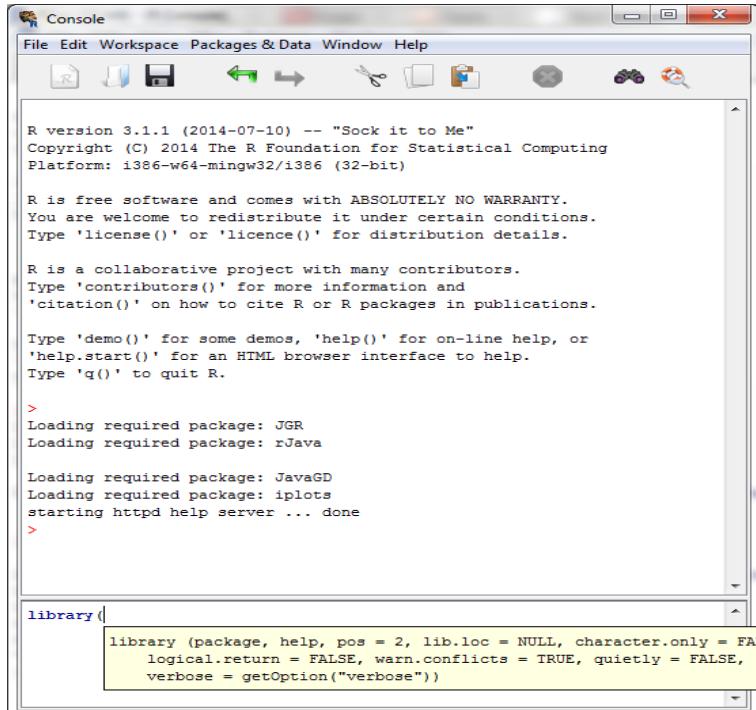
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

> library(JGR)  
Loading required package: rJava  
Loading required package: JavaGD  
Loading required package: ipplots

Please type JGR() to launch console. Platform specific launchers (.exe and .app) can also be obtained at <http://www.rforge.net/JGR/files/>.

> JGR()

# Installation Deducer (with JGR)



The screenshot shows the R Console window with the following output:

```
R version 3.1.1 (2014-07-10) -- "Sock it to Me"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (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.

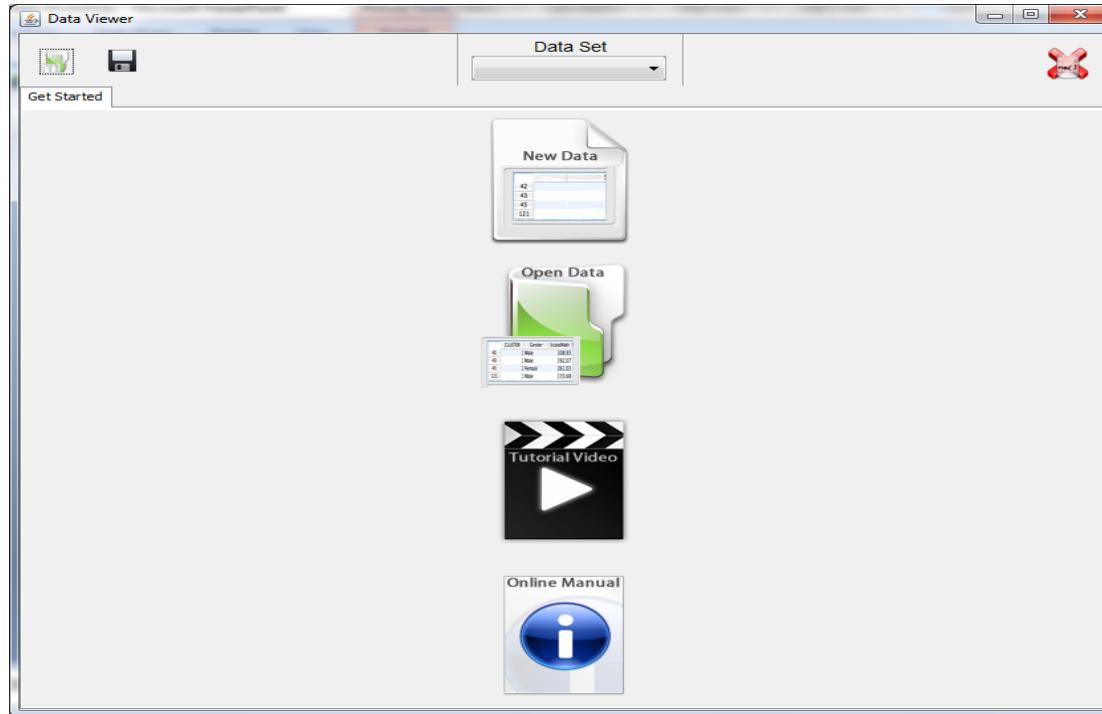
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
Loading required package: JGR
Loading required package: rJava

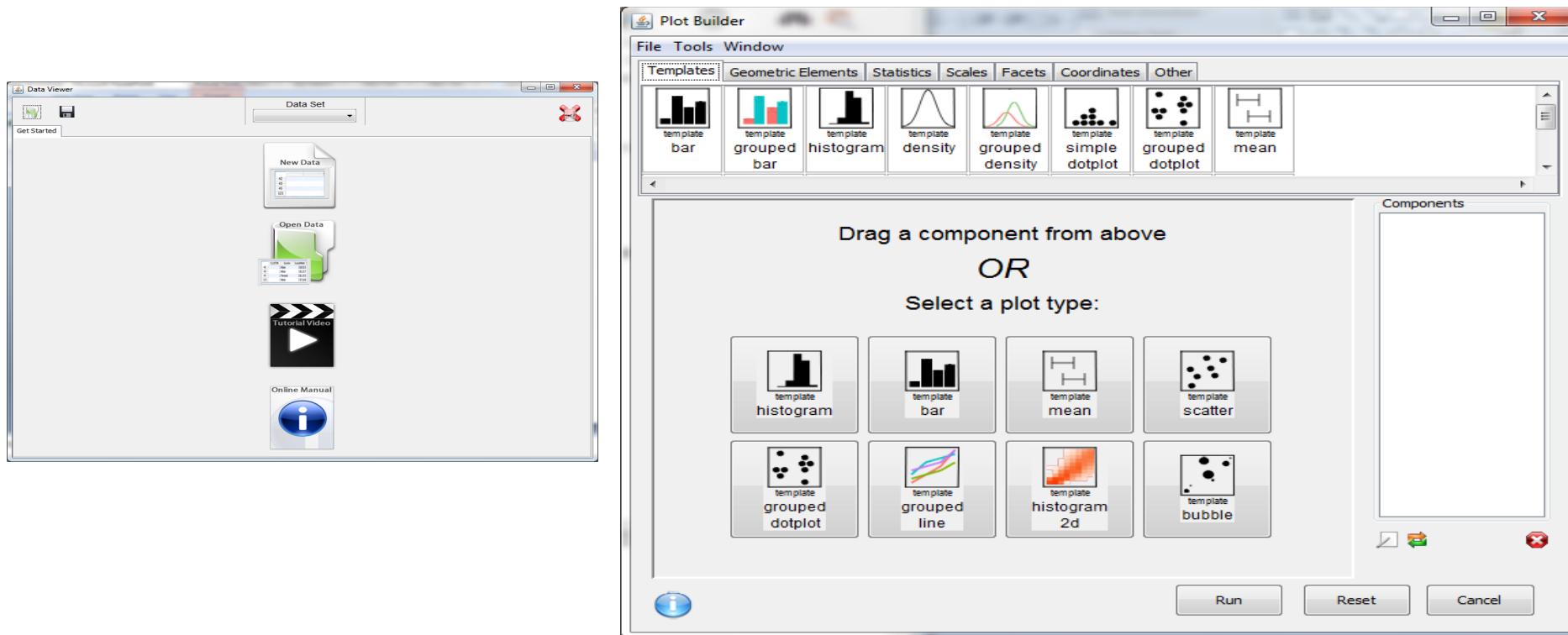
Loading required package: JavaGD
Loading required package: ipplots
starting httpd help server ... done
>

library()
library(package, help, pos = 2, lib.loc = NULL, character.only = FALSE,
logical.return = FALSE, warn.conflicts = TRUE, quietly = FALSE,
verbose = getOption("verbose"))
```

# Installation Deducer (with JGR)



# Overview of Deducer (with JGR)



# Demo Deducer

- `data()`
  - `data(mtcars)`

# RStudio

RStudio Desktop enables you with following advantages of native R console

- Syntax highlighting, code completion, and smart indentation
- Execute R code directly from the source editor
- Quickly jump to function definitions
- Easily manage multiple working directories using projects
- Integrated R help and documentation
- Interactive debugger to diagnose and fix errors quickly
- Extensive package development tools

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

# RStudio

RStudio Server enables you to provide a browser based interface (the RStudio IDE) to a version of R running on a remote Linux server. Deploying R and RStudio on a server has a number of benefits, including:

- The ability to access your R workspace from any computer in any location;
- Easy sharing of code, data, and other files with colleagues;
- Allowing multiple users to share access to the more powerful compute resources (memory, processors, etc.) available on a well equipped server; and
- Centralized installation and configuration of R, R packages, TeX, and other supporting libraries.

Go to file/function

```

new2.R x packages.R x chapter1.Rmd x Untitled1* x
Source on Save Run Source
1 library(ggplot2)
2 data(diamonds)
3 barplot(diamonds$price)
4 plot(diamonds$price)
5 plot(diamonds$price,diamonds$carat)
6 pie(table(diamonds$cut))
7 boxplot(diamonds$price)
8 boxplot(diamonds$price-diamonds$cut)
9 boxplot(diamonds$price-diamonds$color)
10 plot(diamonds$cut,diamonds$color)
11 hist(diamonds$price)
12
12:1 (Top Level) R Script

```

## RStudio - Interface

Console

&gt;

kmeans

```

function (x, centers, iter.max = 10, nstart = 1, algorithm = c("Hartigan-Wong",
  "Lloyd", "Forgy", "MacQueen"), trace = FALSE)
{
  do_one <- function(nmeth) {
    switch(nmeth, {
      isteps.Qtran <- 50 * m
      iTran <- c(as.integer(isteps.Qtran), integer(max(0,
        k - 1)))
      Z <- .Fortran(C_kmns, x, m, p, centers = centers,
        as.integer(k), c1 = integer(m), c2 = integer(m),
        nc = integer(k), double(k), double(k), ncp = integer(k),
        D = double(m), iTran = iTran, live = integer(k),
        iter = iter.max, wss = double(k), ifault = as.integer(trace))
      switch(Z$ifault, stop("empty cluster: try a better set of initial centers",
        call. = FALSE), Z$iter <- max(Z$iter, iter.max +
          1L), stop("number of cluster centres must lie between 1 and nrow(x)",
        call. = FALSE), warning(gettextf("Quick-TRANSFER stage steps exceeded maximum (= %d)",
          isteps.Qtran), call. = FALSE))
    }, {
      Z <- .C(C_kmeans_Lloyd, x, m, p, centers = centers,
        k, c1 = integer(m), iter = iter.max, nc = integer(k),
        wss = double(k))
    })
  }
}

```

Environment History

Import Dataset Clear

Global Environment

Data

 diamonds	53940 obs. of 10 variables
 iris3	50 obs. of 12 variables

Values

a	NULL (empty)
i	90L

Files Plots Packages Help Viewer

R: Search Results



The search string was "kmeans"

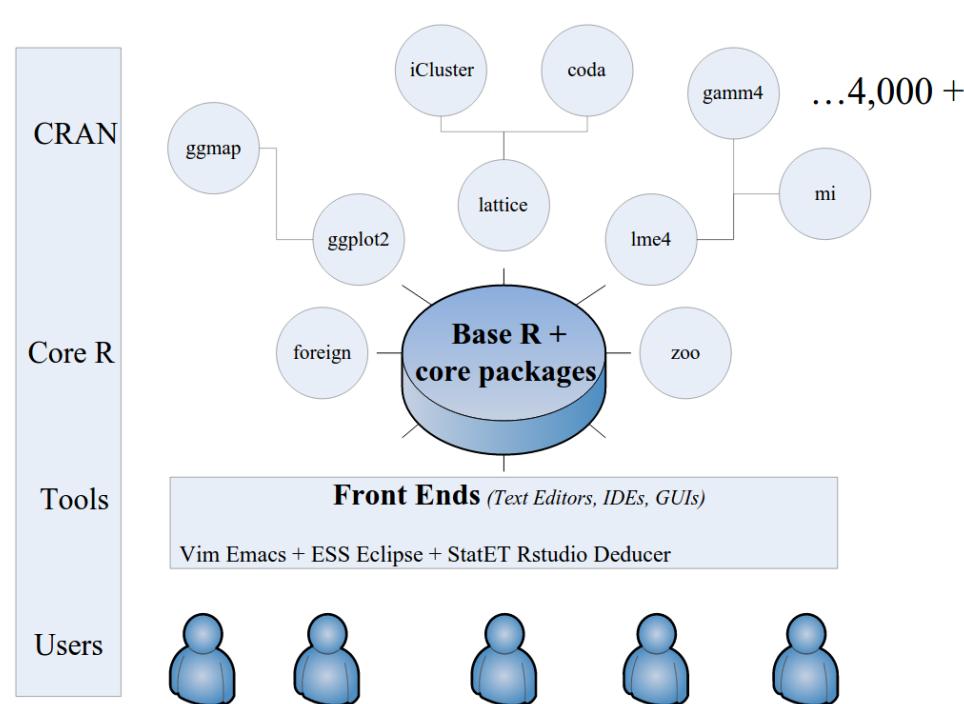
### Vignettes:

[broom::kmeans](#) kmeans with dplyr+broom[HTML](#) [source](#) [R code](#)

### Help pages:

[amap::Kmeans](#) K-Means Clustering[broom::augment.kmeans](#) Tidying methods for kmeans objects[e1071::cmeans](#) Fuzzy C-Means Clustering

# R Landscape



# R Documentation

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

## 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 platforms. The version of the manuals for each platform are part of the respective R installations. The manuals change with R, hence we provide three versions for the patched release version (R-patched) and finally a version for the forthcoming R version that is still in development.

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

Manual	R-release
<b>An Introduction to R</b> 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.	<a href="#">HTML</a>   <a href="#">PDF</a>   <a href="#">EPUB</a>
<b>R Data Import/Export</b> describes the import and export facilities available either in R itself or via packages which are available from CRAN.	<a href="#">HTML</a>   <a href="#">PDF</a>   <a href="#">EPUB</a>
<b>R Installation and Administration</b>	<a href="#">HTML</a>   <a href="#">PDF</a>   <a href="#">EPUB</a>
<b>Writing R Extensions</b> covers how to create your own packages, write R help files, and the foreign language (C, C++, Fortran, ...) interfaces.	<a href="#">HTML</a>   <a href="#">PDF</a>   <a href="#">EPUB</a>
A draft of <b>The R language definition</b> documents the language <i>per se</i> . 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.	<a href="#">HTML</a>   <a href="#">PDF</a>   <a href="#">EPUB</a>
<b>R Internals</b> : a guide to the internal structures of R and coding standards for the core team working on R itself.	<a href="#">HTML</a>   <a href="#">PDF</a>   <a href="#">EPUB</a>
<b>The R Reference Index</b> : contains all help files of the R standard and recommended packages in printable form. (9MB, approx. 3500 pages)	<a href="#">PDF</a>

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

The LaTeX or Texinfo sources of the latest version of these documents are contained in every R source distribution (in the `doc/manuals` directory). They can be found in the respective [archives of the R sources](#). The HTML versions of the manuals are also part of most R installation distributions.

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

R

# Documentation

## Vignettes

### ggplot2: An Implementation of the Grammar of Graphics

An implementation of the grammar of graphics in R. It combines the advantages of both base and lattice graphics by step from multiple data sources. It also implements a sophisticated multidimensional conditioning system with information, documentation and examples.

Version:	1.0.1
Depends:	R (≥ 2.14), stats, methods
Imports:	<a href="#">plyr</a> (≥ 1.7.1), <a href="#">digest</a> , <a href="#">grid</a> , <a href="#">gtable</a> (≥ 0.1.1), <a href="#">reshape2</a> , <a href="#">scales</a> (≥ 0.2.3), <a href="#">proto</a> , <a href="#">MASS</a>
Suggests:	<a href="#">quantreg</a> , <a href="#">Hmisc</a> , <a href="#">mapproj</a> , <a href="#">maps</a> , <a href="#">hexbin</a> , <a href="#">maptools</a> , <a href="#">multcomp</a> , <a href="#">nlme</a> , <a href="#">testthat</a> , <a href="#">knitr</a> , <a href="#">mgcv</a>
Enhances:	<a href="#">sp</a>
Published:	2015-03-17
Author:	Hadley Wickham [aut, cre], Winston Chang [aut]
Maintainer:	Hadley Wickham <h.wickham at gmail.com>
BugReports:	<a href="https://github.com/hadley/ggplot2/issues">https://github.com/hadley/ggplot2/issues</a>
License:	<a href="#">GPL-2</a>
URL:	<a href="http://ggplot2.org">http://ggplot2.org</a> , <a href="https://github.com/hadley/ggplot2">https://github.com/hadley/ggplot2</a>
NeedsCompilation:	no
Citation:	<a href="#">ggplot2 citation info</a>
Materials:	<a href="#">README</a> <a href="#">NEWS</a>
In views:	<a href="#">Graphics</a> , <a href="#">Phylogenetics</a>
CRAN checks:	<a href="#">ggplot2 results</a>

#### Downloads :

Reference manual:	<a href="#">ggplot2.pdf</a>
Vignettes:	<a href="#">Contributing to ggplot2 development</a> <a href="#">ggplot2 release process</a>
Package source:	<a href="#">ggplot2_1.0.1.tar.gz</a>
Windows binaries:	r-devel: <a href="#">ggplot2_1.0.1.zip</a> , r-release: <a href="#">ggplot2_1.0.1.zip</a> , r-oldrel: <a href="#">ggplot2_1.0.1.zip</a>
OS X Snow Leopard binaries:	r-release: not available, r-oldrel: <a href="#">ggplot2_1.0.1.tgz</a>
OS X Mavericks binaries:	r-release: <a href="#">ggplot2_1.0.1.tgz</a>
Old sources:	<a href="#">ggplot2 archive</a>

#### Reverse dependencies :

Reverse depends: [alphahull](#), [AmpliconDuo](#), [aoristic](#), [apsimr](#), [bcrm](#), [bde](#), [benchmark](#), [biomod2](#), [bootnet](#), [brms](#)

# CRAN Views

<http://cran.r-project.org/web/views/>

<a href="#">Bayesian</a>	Bayesian Inference
<a href="#">ChemPhys</a>	Chemometrics and Computational Physics
<a href="#">ClinicalTrials</a>	Clinical Trial Design, Monitoring, and Analysis
<a href="#">Cluster</a>	Cluster Analysis & Finite Mixture Models
<a href="#">DifferentialEquations</a>	Differential Equations
<a href="#">Distributions</a>	Probability Distributions
<a href="#">Econometrics</a>	Econometrics
<a href="#">Environmetrics</a>	Analysis of Ecological and Environmental Data
<a href="#">ExperimentalDesign</a>	Design of Experiments (DoE) & Analysis of Experimental Data
<a href="#">Finance</a>	Empirical Finance
<a href="#">Genetics</a>	Statistical Genetics
<a href="#">Graphics</a>	Graphic Displays & Dynamic Graphics & Graphic Devices & Visualization
<a href="#">HighPerformanceComputing</a>	High-Performance and Parallel Computing with R
<a href="#">MachineLearning</a>	Machine Learning & Statistical Learning
<a href="#">MedicalImaging</a>	Medical Image Analysis
<a href="#">MetaAnalysis</a>	Meta-Analysis
<a href="#">Multivariate</a>	Multivariate Statistics
<a href="#">NaturalLanguageProcessing</a>	Natural Language Processing
<a href="#">NumericalMathematics</a>	Numerical Mathematics
<a href="#">OfficialStatistics</a>	Official Statistics & Survey Methodology
<a href="#">Optimization</a>	Optimization and Mathematical Programming
<a href="#">Pharmacokinetics</a>	Analysis of Pharmacokinetic Data
<a href="#">Phylogenetics</a>	Phylogenetics, Especially Comparative Methods
<a href="#">Psychometrics</a>	Psychometric Models and Methods
<a href="#">ReproducibleResearch</a>	Reproducible Research
<a href="#">Robust</a>	Robust Statistical Methods
<a href="#">SocialSciences</a>	Statistics for the Social Sciences
<a href="#">Spatial</a>	Analysis of Spatial Data
<a href="#">SpatioTemporal</a>	Handling and Analyzing Spatio-Temporal Data
<a href="#">Survival</a>	Survival Analysis
<a href="#">TimeSeries</a>	Time Series Analysis
<a href="#">WebTechnologies</a>	Web Technologies and Services
<a href="#">gR</a>	gRaphical Models in R

# R Community

- email groups <http://www.r-project.org/mail.html>

**R-announce**

**R-help**

**R-package-devel**

**R-devel**

**R-packages**

## Special Interest Groups

- Stack Overflow [r]
- Twitter #rstats
- Blogs at <http://www.r-bloggers.com/> (573 blogs)

# Stack Overflow

<http://stackoverflow.com/questions/tagged/r>

The screenshot shows the Stack Overflow homepage with the search bar set to 'r'. The main content area displays 'Tagged Questions' for the 'r' tag. There are three visible posts:

- R Count number of rows in one column of a data frame?**  
I just want to know how to get r to list the number of occupied rows of a specific column of a data frame. My guess was nrow(dataframe\$column) though that didn't work.  
asked 2 mins ago by RyanMe321  
1 answer, 3 views
- Create interactive webmap with markers in R using Shiny, Leaflet and rCharts**  
I am trying to create an interactive webmap in R to display storms using Shiny, Leaflet and rCharts (the structure is loosely based on the <http://ramnath.github.io/bikeshare> app). The idea is that ...  
asked 5 mins ago by Louise  
0 votes, 0 answers, 2 views
- R - gsub a specific character of a specific position**  
I would like to delete the last character of a variable. I was wondering if it is possible to select the position with gsub and delete the character of this particular position. In this example, I ...  
asked 15 mins ago by giacomoV  
0 votes, 0 answers, 7 views

On the right side, there is a sidebar with the following sections:

- 90,861 questions tagged**  
[about »](#)
- Featured on Meta**
  - April 2015 Community Moderator Election Results
- Hot Meta Posts**
  - Failed edit to a question says: "Your answer couldn't be submitted"
  - The Font Awesome child tags are too specific - are they even necessary?
  - Flagging questions with details only in comments
- Favorite Tags** [edit](#)  
[Add a favorite tag](#)
- Looking for a job?**
  - Chief Software Architect - Java - \$100K
  - Crossover
  - Bengaluru, India / remote
  - eclipse netbeans

# Twitter

<https://twitter.com/search?q=rstats&src=spqr>

...

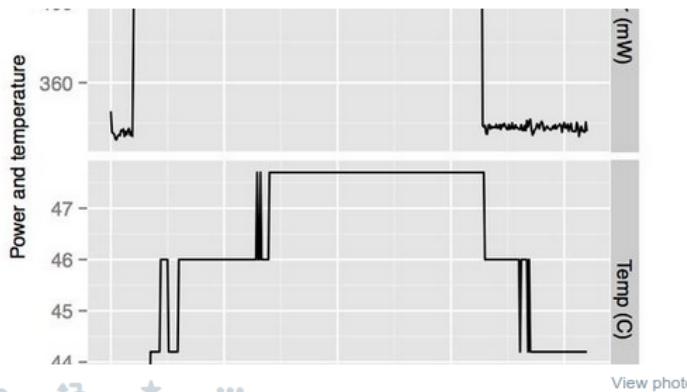
Results for #rstats

Top / All



Mark Benson @markbenson · 5m

Power and heat are related. Here's an R plot I did that proves it on the Kindle Fire. [#rstats vanilladraft.com/stmes/](http://vanilladraft.com/stmes/)



Stéphane Fréchette @sfrechette · 8m

How to get your very own RStudio Server and Shiny Server with DigitalOcean [r-bloggers.com/how-to-get-you...](http://r-bloggers.com/how-to-get-you...) #datascience #feedly #rstats #shiny



Ankit kansal @sinisterinankit · 9m

Interesting post on configuring parallel computing on #r #rstudio #rstats #dataprocessing #data



Learn R @R\_Programming

How to do parallel computing with R? [rstatistics.net/parallel-compu...](http://rstatistics.net/parallel-compu...)  
#rstats #datascience

# Help within R

? "keyword"

?? "keyword"

Example-

```
> ?kmeans
```

```
> ??kmeans
```

# Introductory R

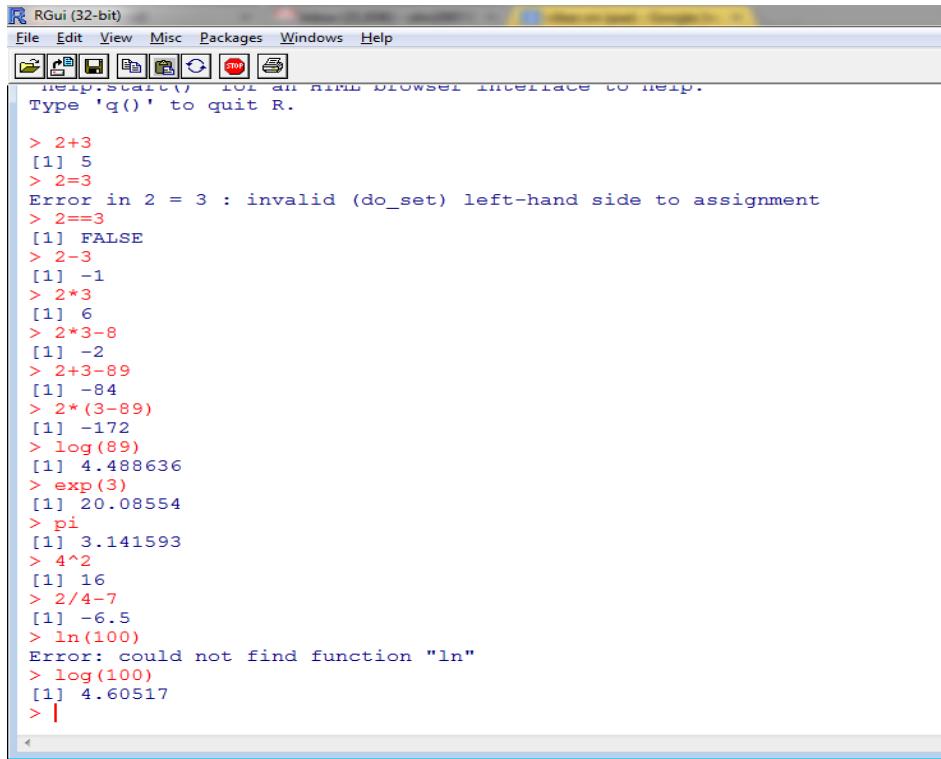
```
> Sys.Date()  
[1] "2015-05-10"  
> Sys.time()  
[1] "2015-05-10 18:28:32 IST"
```

# R as a Calculator

## Basic Math on R Console

- +
  - -
  - Log
  - Exp
  - \*
  - /
  - ()
- mean
  - sum
  - sd
  - log
  - median
  - exp

# Demo- Basic Math on R Console



RGui (32-bit)

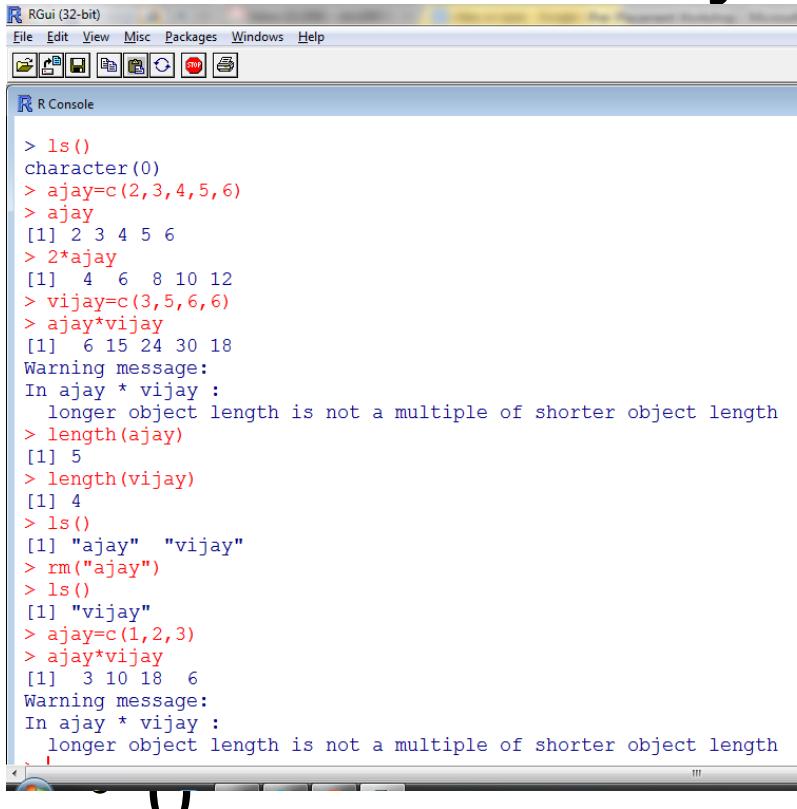
File Edit View Misc Packages Windows Help

help.start() for an HTML browser interface to help.  
Type 'q()' to quit R.

```
> 2+3
[1] 5
> 2=3
Error in 2 = 3 : invalid (do_set) left-hand side to assignment
> 2==3
[1] FALSE
> 2-3
[1] -1
> 2*3
[1] 6
> 2*3-8
[1] -2
> 2+3-89
[1] -84
> 2*(3-89)
[1] -172
> log(89)
[1] 4.488636
> exp(3)
[1] 20.08554
> pi
[1] 3.141593
> 4^2
[1] 16
> 2/4-7
[1] -6.5
> ln(100)
Error: could not find function "ln"
> log(100)
[1] 4.60517
> |
```

Hint- Ctrl +L clears screen

# Demo- Basic Objects on R Console



```
R Gui (32-bit)
File Edit View Misc Packages Windows Help
R Console

> ls()
character(0)
> ajay=c(2,3,4,5,6)
> ajay
[1] 2 3 4 5 6
> 2*ajay
[1] 4 6 8 10 12
> vijay=c(3,5,6,6)
> ajay*vijay
[1] 6 15 24 30 18
Warning message:
In ajay * vijay :
  longer object length is not a multiple of shorter object length
> length(ajay)
[1] 5
> length(vijay)
[1] 4
> ls()
[1] "ajay" "vijay"
> rm("ajay")
> ls()
[1] "vijay"
> ajay=c(1,2,3)
> ajay*vijay
[1] 3 10 18 6
Warning message:
In ajay * vijay :
  longer object length is not a multiple of shorter object length
I
```

## Functions-

ls() – what objects are here

rm("foo") removes object named foo

## Assignment

Using = or -> assigns object names to values

Hint- Up arrow ↑ gives you last typed command

# Functions and Loops

- Loops

```
for (number in 1:5){ print (number) }
```

```
> for (number in 1:5){ print (number) }
[1] 1
[1] 2
[1] 3
[1] 4
[1] 5
> for (i in 1:5){ print (i) }
[1] 1
[1] 2
[1] 3
[1] 4
[1] 5
> for (i in 1:5){ rnorm(i,10,10) }
> for (i in 1:5){ print(rnorm(i,10,10)) }
[1] 1.090406
[1] 8.611727 16.670168
[1] 10.84623 13.13938 11.56230
[1] 6.068250 -18.723389 33.174107 -1.320091
[1] 13.939702 -9.037375 13.755986 9.459680 9.625309
> |
```

# Functions and Loops

- Function

```
functionajay=function(a)(a^2+2*a+1)
```

```
> -----  
> functionajay=function(a) (a^2+2*a+1)  
[1] 4  
> functionajay(2)  
[1] 9  
> for (i in 1:5){ print(rnorm(i) )  
Error: unexpected '}' in "for (i in 1:5){ print(rnorm(i) )"  
>  
> for (i in 1:5){ print(functionajay(i)) }  
[1] 4  
[1] 9  
[1] 16  
[1] 25  
[1] 36  
> |
```

Hint: Always match brackets

Each ( deserves a )

Each { deserves a }

Each [ deserves a ]

# Other sources to learn R

swirlstats

<http://swirlstats.com/>

datacamp

<https://www.datacamp.com/>

codeschool

<http://tryr.codeschool.com/>

coursera

<https://www.coursera.org/course/compdata>

<https://www.coursera.org/course/rprog>



# Good coding practices

- Use # for comment
- Use git for version control
- Use Rstudio for multiple lines of code

# Functions in R

- custom functions
- source code for a function

```
Console ~/ 
> kmeans
function (x, centers, iter.max = 10, nstart = 1, algorithm = c("Hartigan-Wong",
  "Lloyd", "Forgy", "MacQueen"), trace = FALSE)
{
  do_one <- function(nmeth) {
    switch(nmeth, {
      isteps.Qtran <- 50 * m
      iTran <- c(as.integer(isteps.Qtran), integer(max(0,
        k - 1)))
      Z <- .Fortran(C_kmeans, x, m, p, centers = centers,
        as.integer(k), c1 = integer(m), c2 = integer(m),
        nc = integer(k), double(k), double(k), ncp = integer(k),
        D = double(m), iTran = iTran, live = integer(k),
        iter = iter.max, wss = double(k), ifault = as.integer(trace))
      switch(Z$ifault, stop("empty cluster: try a better set of initial centers",
        call. = FALSE), Z$iter <- max(Z$iter, iter.max +
          1L), stop("number of cluster centres must lie between 1 and nrow(x)",
        call. = FALSE), warning(gettextf("Quick-TRANSFER stage steps exceeded maximum (= %d",
        isteps.Qtran), call. = FALSE))
    }, {
      Z <- .C(C_kmeans_Lloyd, x, m, p, centers = centers,
        k, c1 = integer(m), iter = iter.max, nc = integer(k),
        wss = double(k))
    }, {
      Z <- .C(C_kmeans_MacQueen, x, m, p, centers = as.double(centers),
        k, c1 = integer(m), iter = iter.max, nc = integer(k),
        wss = double(k))
    })
    if (m23 <- any(nmeth == c(2L, 3L))) {
      if (any(Z$nc == 0))
        warning("empty cluster: try a better set of initial centers",
          call. = FALSE)
    }
  }
}
```

# HOMEWORK TIME !



# Functions Used in this Lesson

function(x)

for

library

install.packages

update.packages

ls

rm

print

# Citations and References

> citation()

To cite R in publications use:

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.