CRAN R

R HISTORY

- “S” language for data analysis developed at Bell Labs 1976
- Licensed by AT&T. Product name: S-plus.
- R: initially written & released as an open source software by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand
- Since 1997: 1000s of code writers and statisticians happy to share their libraries!

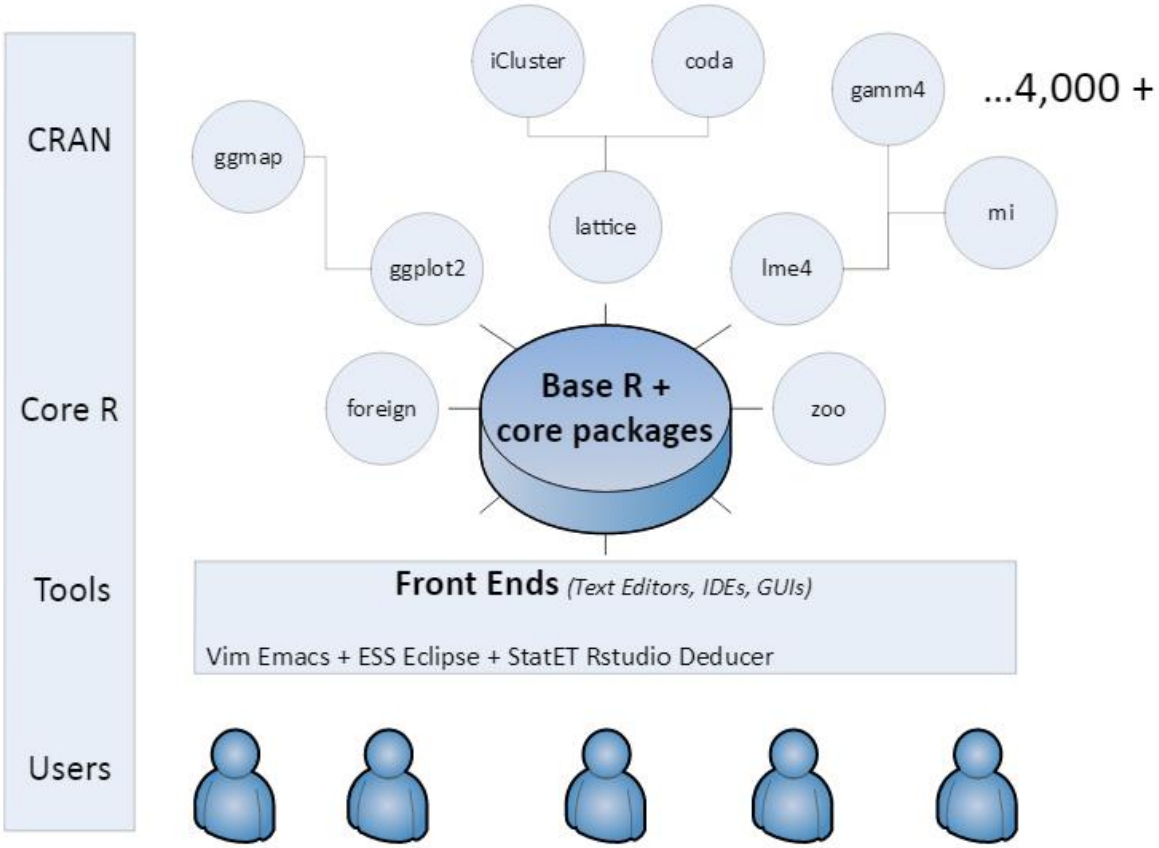
Ross Ihaka Robert Gentleman

Introduction

- CRAN - Comprehensive R Archive Network.
  https://cran.r-project.org/
- R is a system for statistical computation and graphics.
  - Programming language
  - High level graphics
  - Interfaces to other languages
  - Debugging facilities.
- Most widely used data analysis software.
- Create beautiful and unique data visualizations.
- The source code for the R software environment is written primarily in C, Fortran, and R.

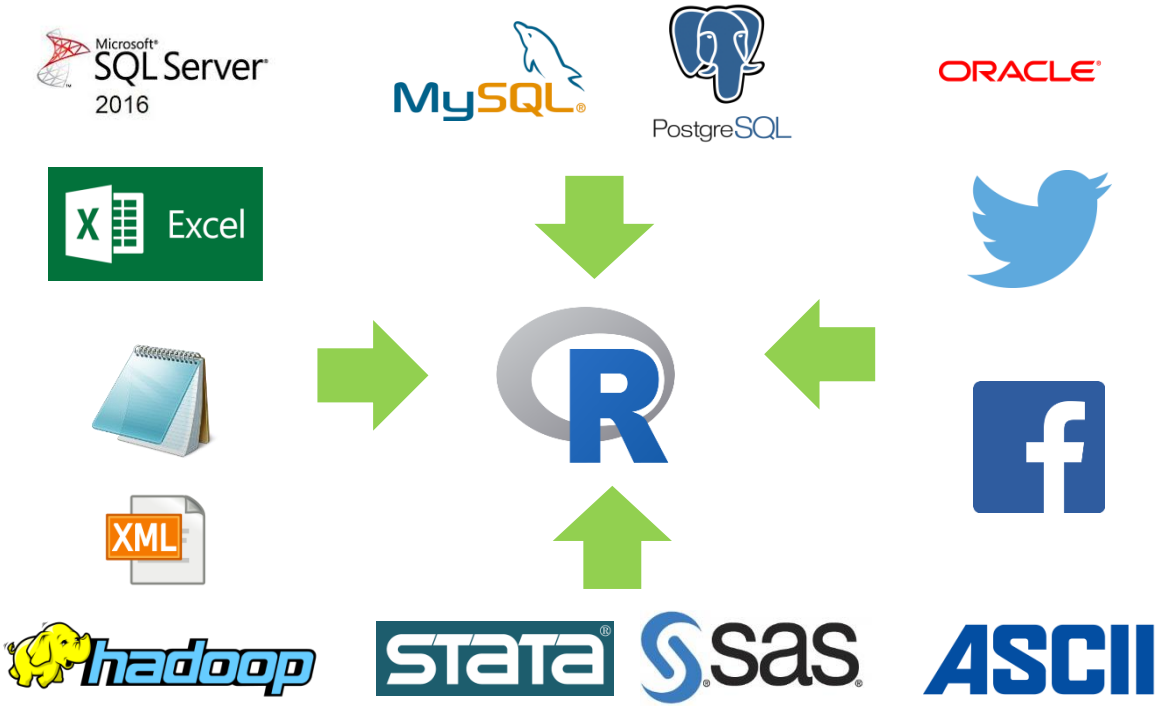
Open Source Project (community)	RAM (Primary Memory) for Processing	Interpreted Language	CASE Sensitive

CRAN R



- R consists of a core and packages.
- Base R and most R packages are available for download from the Comprehensive R Archive Network (CRAN)
- Packages contain functions that are not available in the core.
- Packages need to be installed in R from repositories.
- Currently, the CRAN package repository features 8513 available packages.

R DATA SOURCE CONNECTIONS



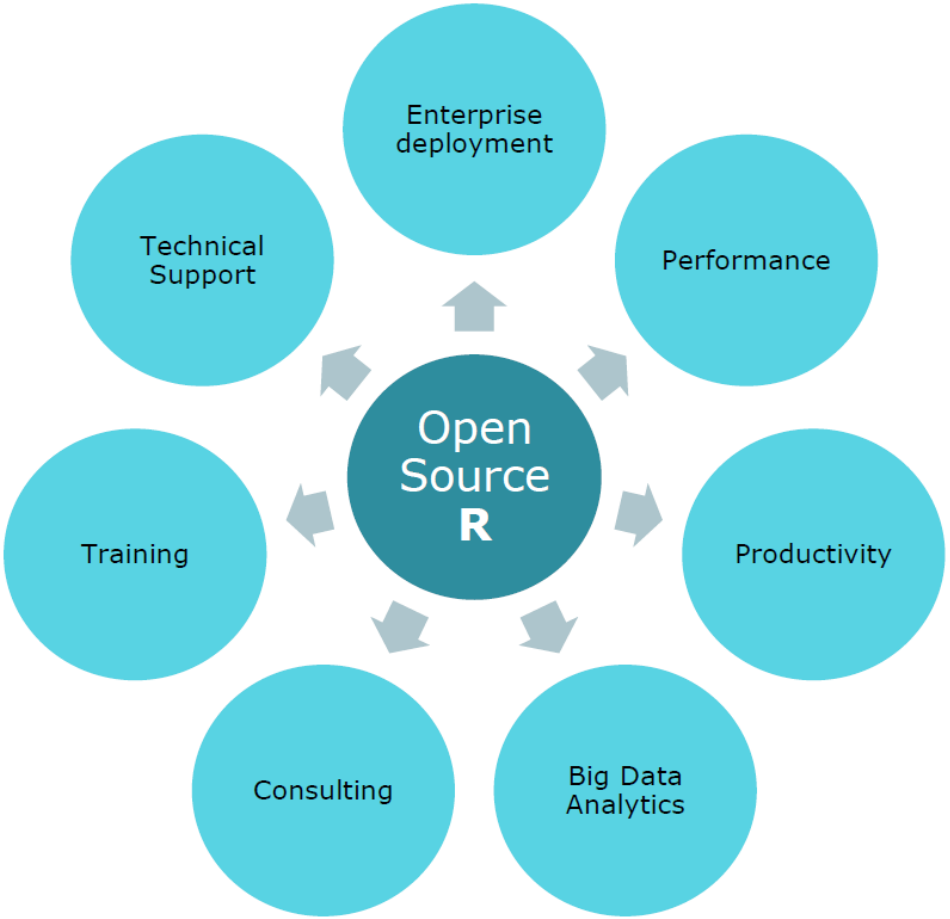
R COMMERCIAL

- Other major commercial software systems supporting connections to or integration with R include:
  - Microsoft R (Revolution Analytics acquired by Microsoft)
  - R shiny
  - SAS
  - SQL Server 2016
  - SPSS
  - STATISTICA
  - Tableau
  - Esri ArcGis
  - Dundas
  - MATLAB
  - Spotfire etc.

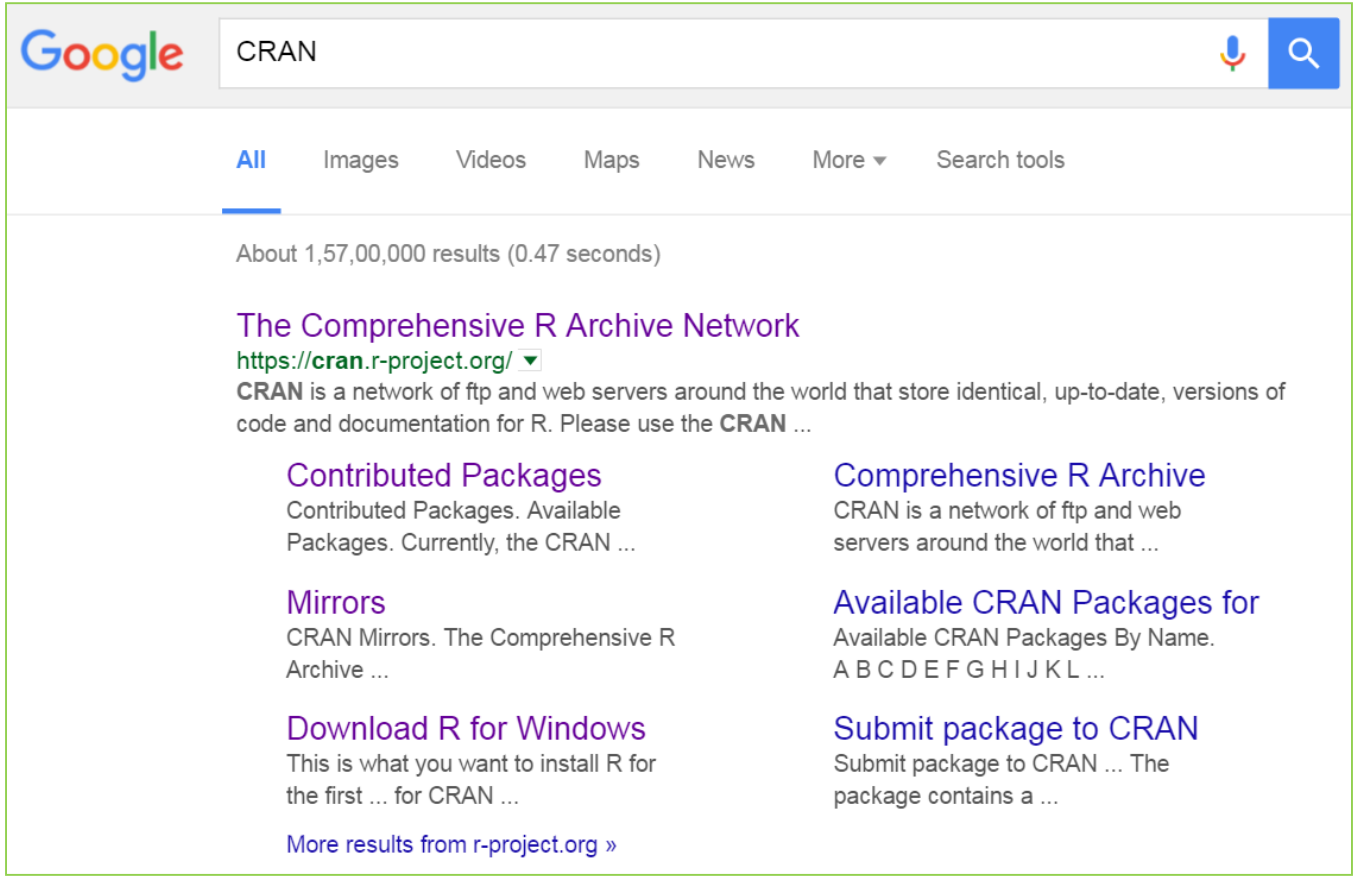
R ADVANTAGES & DISADVANTAGES

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

DISTINCTIVE FEATURES OF R



R DOWNLOAD


Google
CRAN
All
Images
Videos
Maps
News
More
Search tools
About 1,57,00,000 results (0.47 seconds)
The Comprehensive R Archive Network
https://cran.r-project.org/
CRAN is a network of ftp and web servers around the world that store identical, up-to-date, versions of code and documentation for R. Please use the CRAN ...
Contributed Packages
Contributed Packages. Available Packages. Currently, the CRAN ...
Comprehensive R Archive
CRAN is a network of ftp and web servers around the world that ...
Mirrors
CRAN Mirrors. The Comprehensive R Archive ...
Available CRAN Packages for
Available CRAN Packages By Name.
A B C D E F G H I J K L ...
Download R for Windows
This is what you want to install R for the first ... for CRAN ...
Submit package to CRAN
Submit package to CRAN ... The package contains a ...
More results from r-project.org »

R DOWNLOAD

The Comprehensive R Archive Network

Download and Install R

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

- Download R for Linux
- Download R for (Mac) OS X
- Download R for Windows

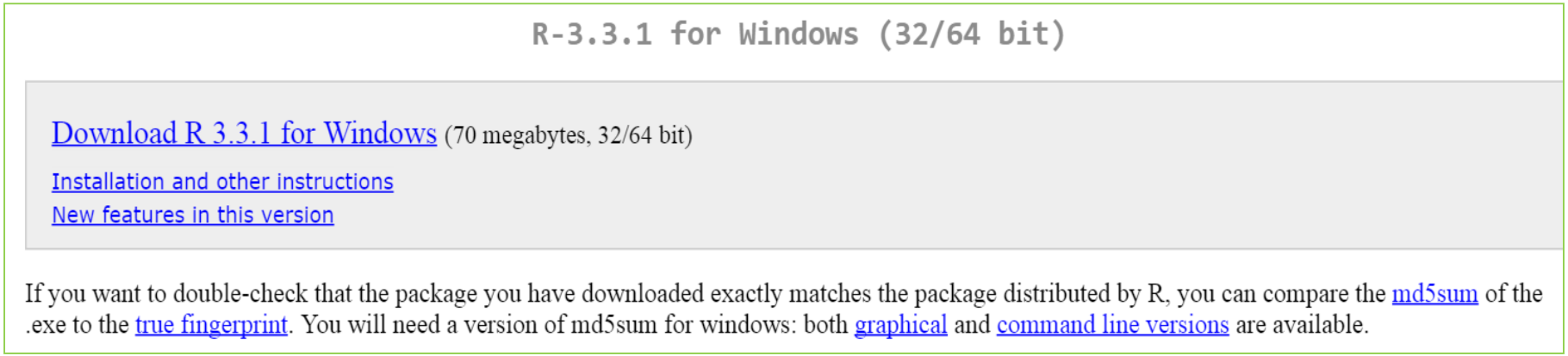
R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

R for Windows

Subdirectories:

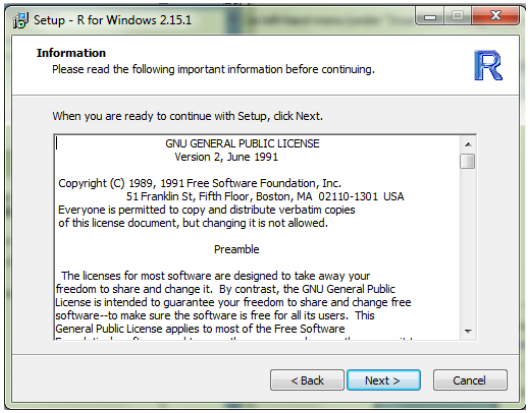
base	Binaries for base distribution (managed by Duncan Murdoch). This is what you want to **install R for the first time**.
contrib	Binaries of contributed CRAN packages (for R >= 2.11.x; managed by Uwe Ligges). There is also information on third party software available for CRAN Windows services and corresponding environment and make variables.
old contrib	Binaries of contributed CRAN packages for outdated versions of R (for R < 2.11.x; managed by Uwe Ligges).
Rtools	Tools to build R and R packages (managed by Duncan Murdoch). This is what you want to build your own packages on Windows, or to build R itself.

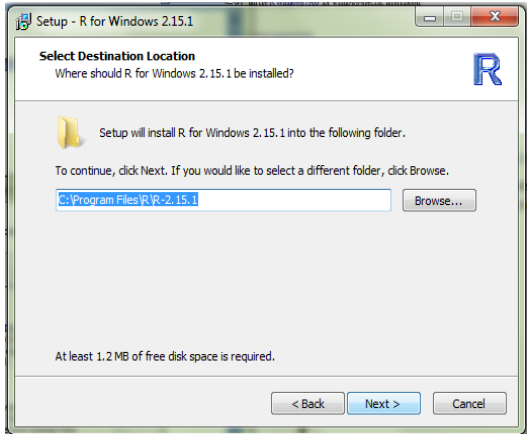
R DOWNLOAD

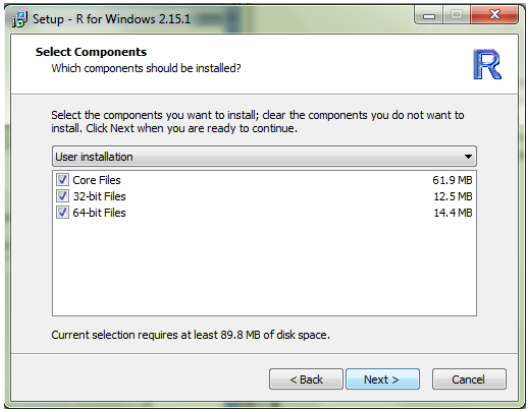


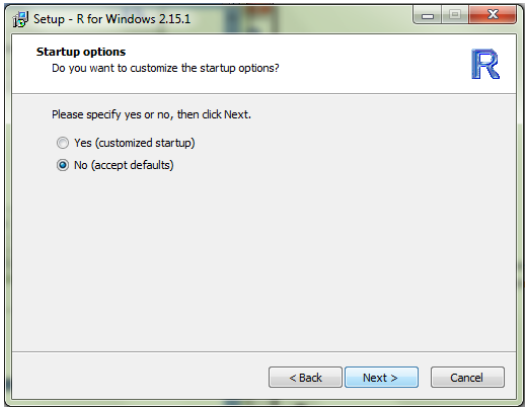
- Download the software and install it. (Simple installation just hit next button till installation complete).

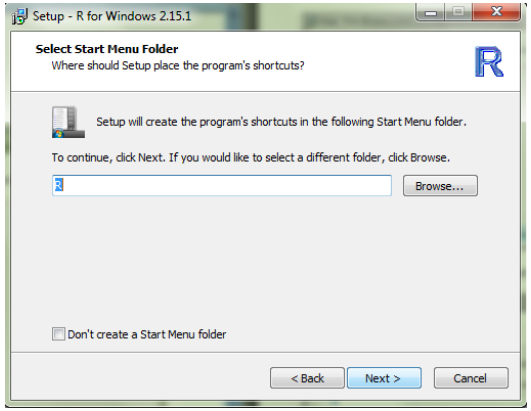
R INSTALLATION

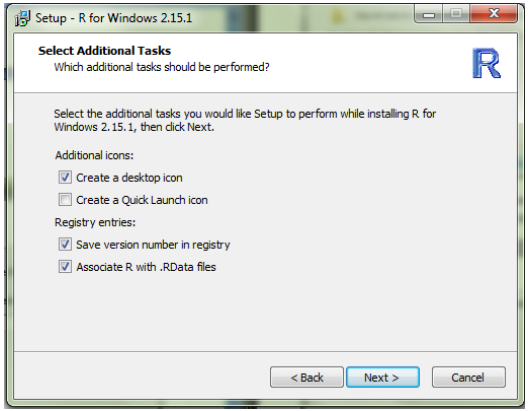




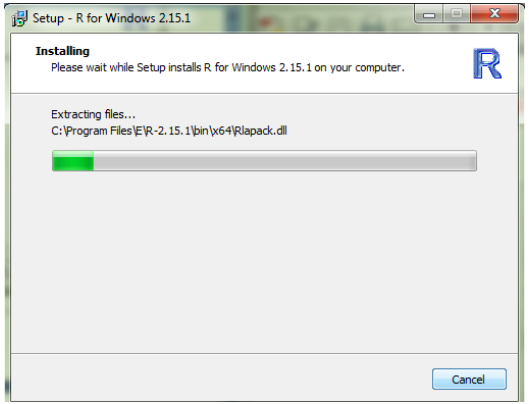






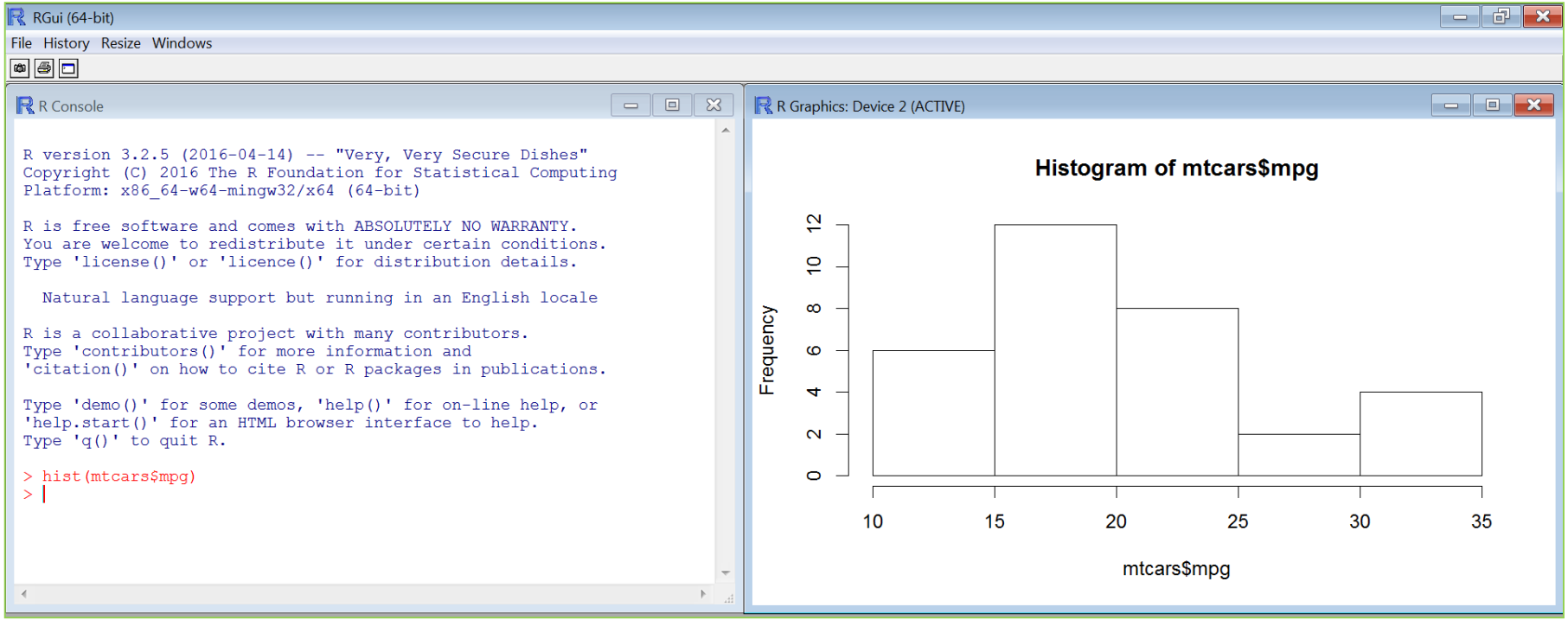


R INSTALLATION



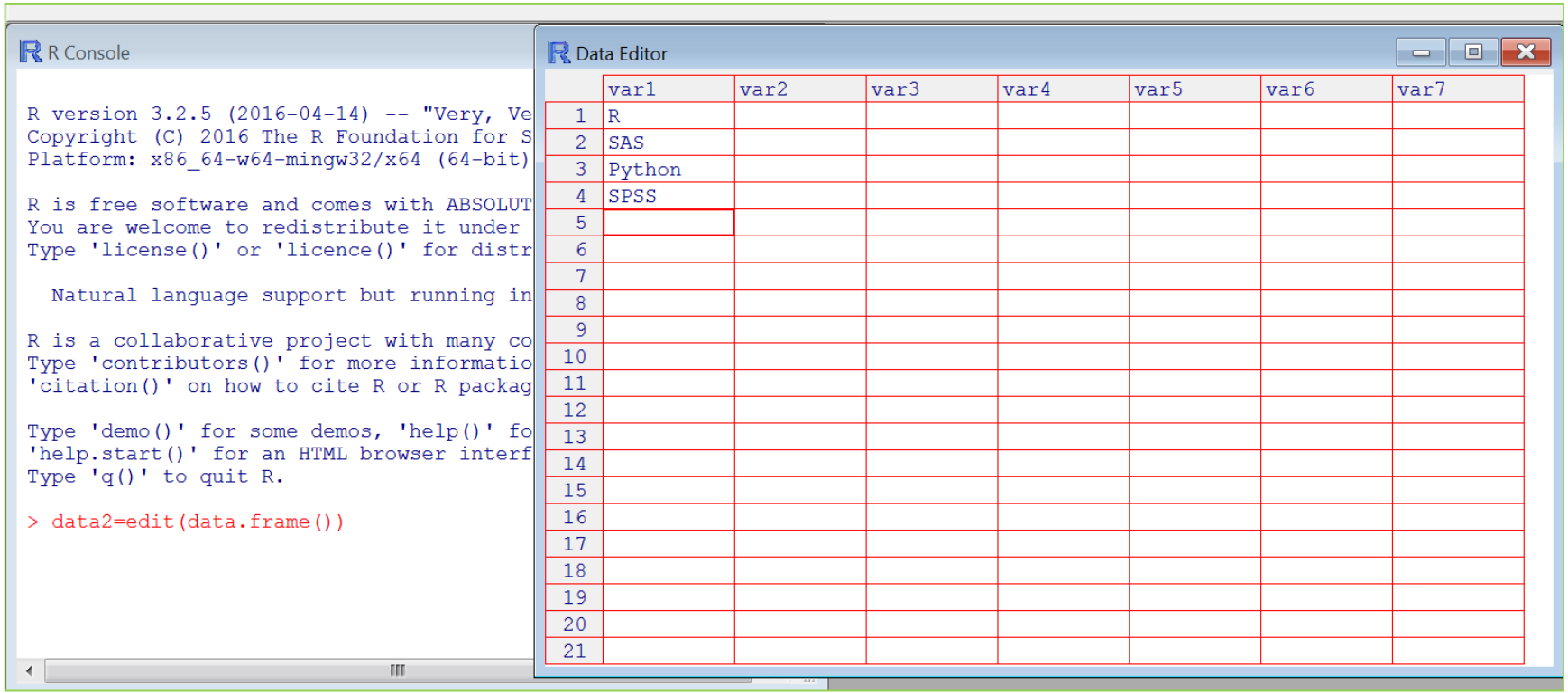
- Click on R icon in desktop/Startup Menu.

R WORKSPACE



- R consist of:
  - Editor/Console window → Coding, viewing data.
  - Output Window → Results

R DATA EDITOR



R DATASETS

R Console

```
R version 3.2.5 (2016-04-14) -- "Very, Very Secure Dishes"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-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.

> require(utils)
> data()
```

R data sets

```
state.abb (state)       US State Facts and Figures
state.area (state)      US State Facts and Figures
state.center (state)    US State Facts and Figures
state.division (state)
                        US State Facts and Figures
state.name (state)      US State Facts and Figures
state.region (state)    US State Facts and Figures
state.x77 (state)       US State Facts and Figures
sunspot.month           Monthly Sunspot Data, from 1749 to "Present"
sunspot.year            Yearly Sunspot Data, 1700-1988
sunspots                Monthly Sunspot Numbers, 1749-1983
swiss                   Swiss Fertility and Socioeconomic Indicators
                        (1888) Data
treering                Yearly Treering Data, -6000-1979
trees                   Girth, Height and Volume for Black Cherry Trees
uspop                   Populations Recorded by the US Census
volcano                 Topographic Information on Auckland's Maunga
                        Whau Volcano
warpbreaks              The Number of Breaks in Yarn during Weaving
women                   Average Heights and Weights for American Women
```

- R consist of default datasets for practice purpose.
- You can access the data from a dataset by typing the name directly
  - mtcars

To view all the default datasets use below code:
require(utils)
data()

R – SURVIVAL

- To get a description of the version of R and its attached packages used in the current session, we can use the session Info function
  sessionInfo()

- Get the default working library
  getwd()

- Setting a new work library
  setwd("D:\\LMDM")

- Listing objects created in current session
  ls()

- Get the R code types in current session
  history()

- Get help
  help.search("histogram")
  ??logistic

- Remove Objects from R session
  rm()

- Code Comment
  #

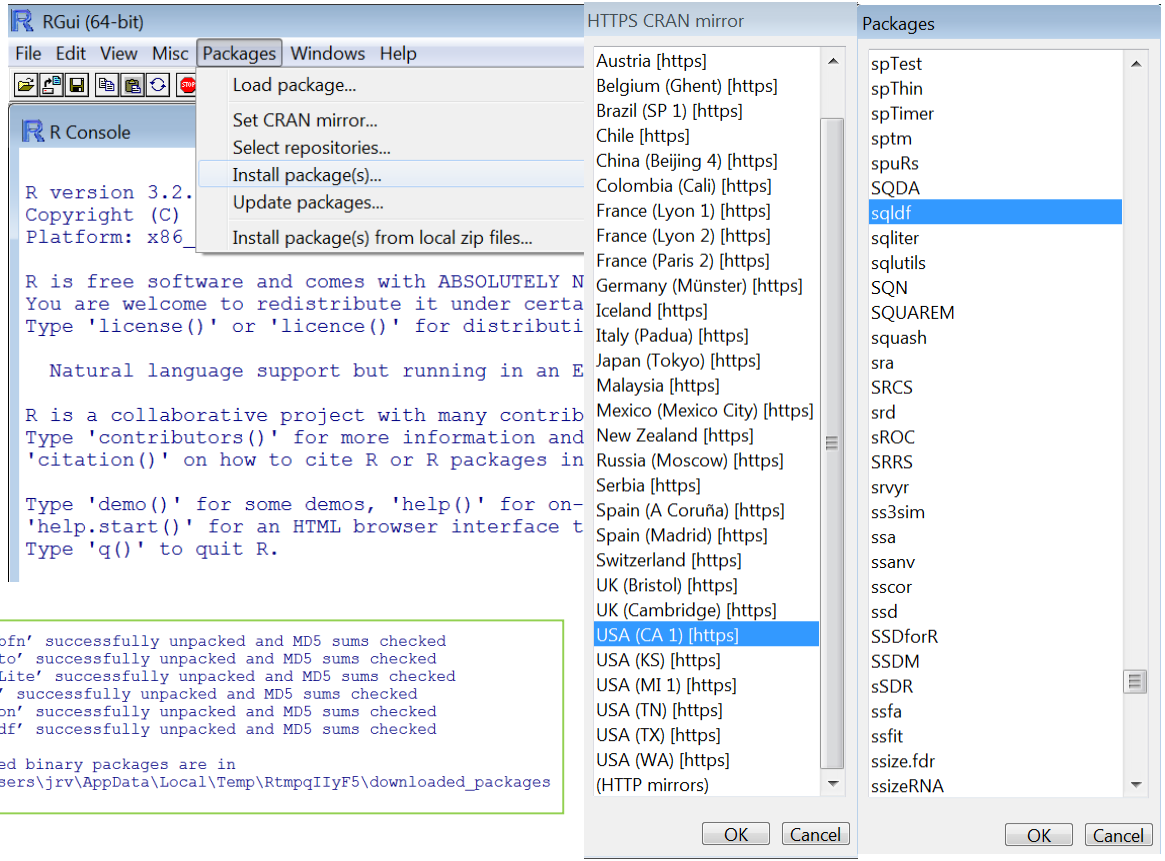
SQL	SAS	R
Table	Dataset	Object
Row	Observation	Observation
Column	Variable	Variable

R PACKAGES – INSTALL, LOAD & UPDATE

- To use packages in R, we must first install them.

Use the **install.packages** function, which typically downloads the package from CRAN and installs it for use

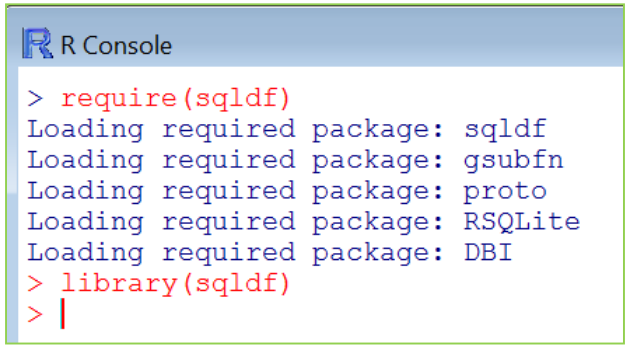
#install.packages("sqldf")
#install.packages("brewer")
#install.packages("reshape2")



R PACKAGES – INSTALL, LOAD & UPDATE

- Loading Packages:
  - If we know we will need a particular package for our current R session, we must load it into the R environment using the library or require functions

    library(sqldf)
    require(sqldf)



Packages should be updated frequently. Outdated packages won't work

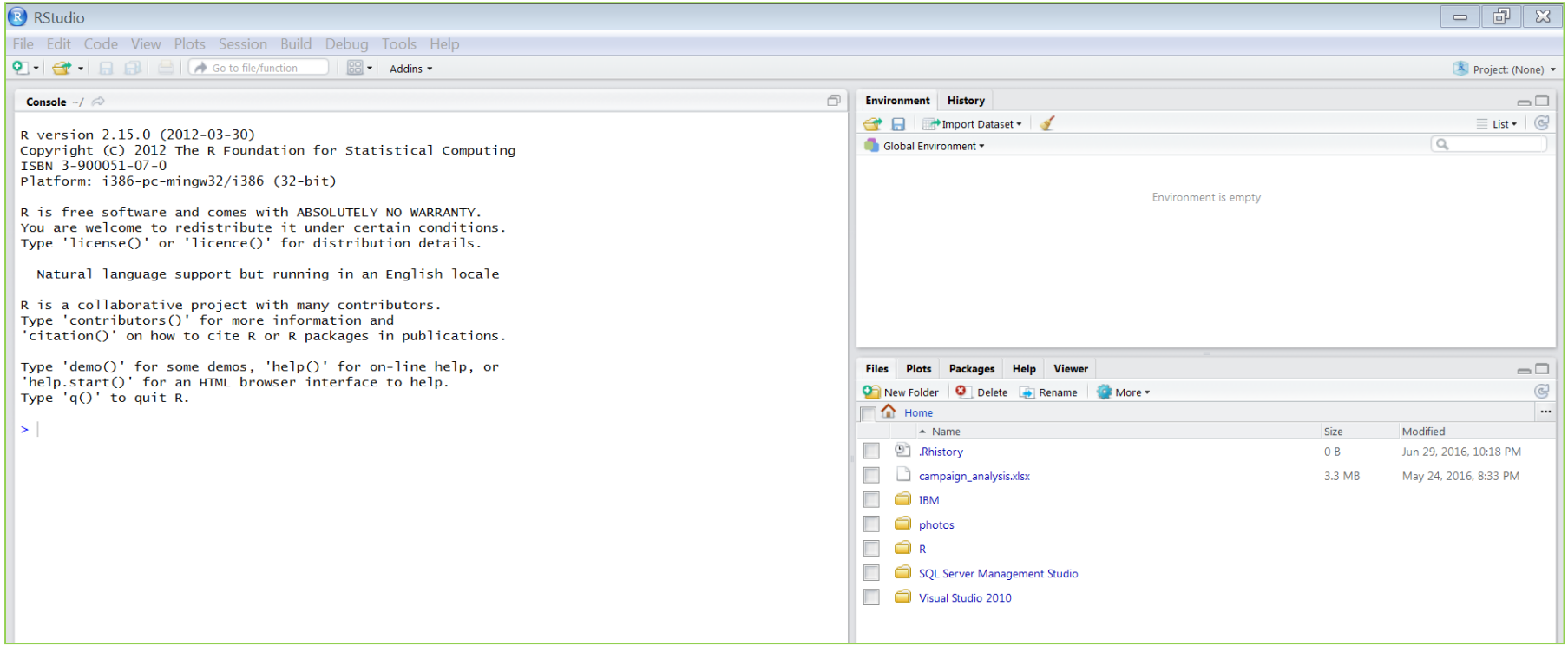
- Updating Packages:
  - Go to Packages in menu tab → Update Packages → Select mirror → select package to update.
  - You can also use below code,

    require(sqldf) || {install.packages("sqldf"); require(sqldf)}

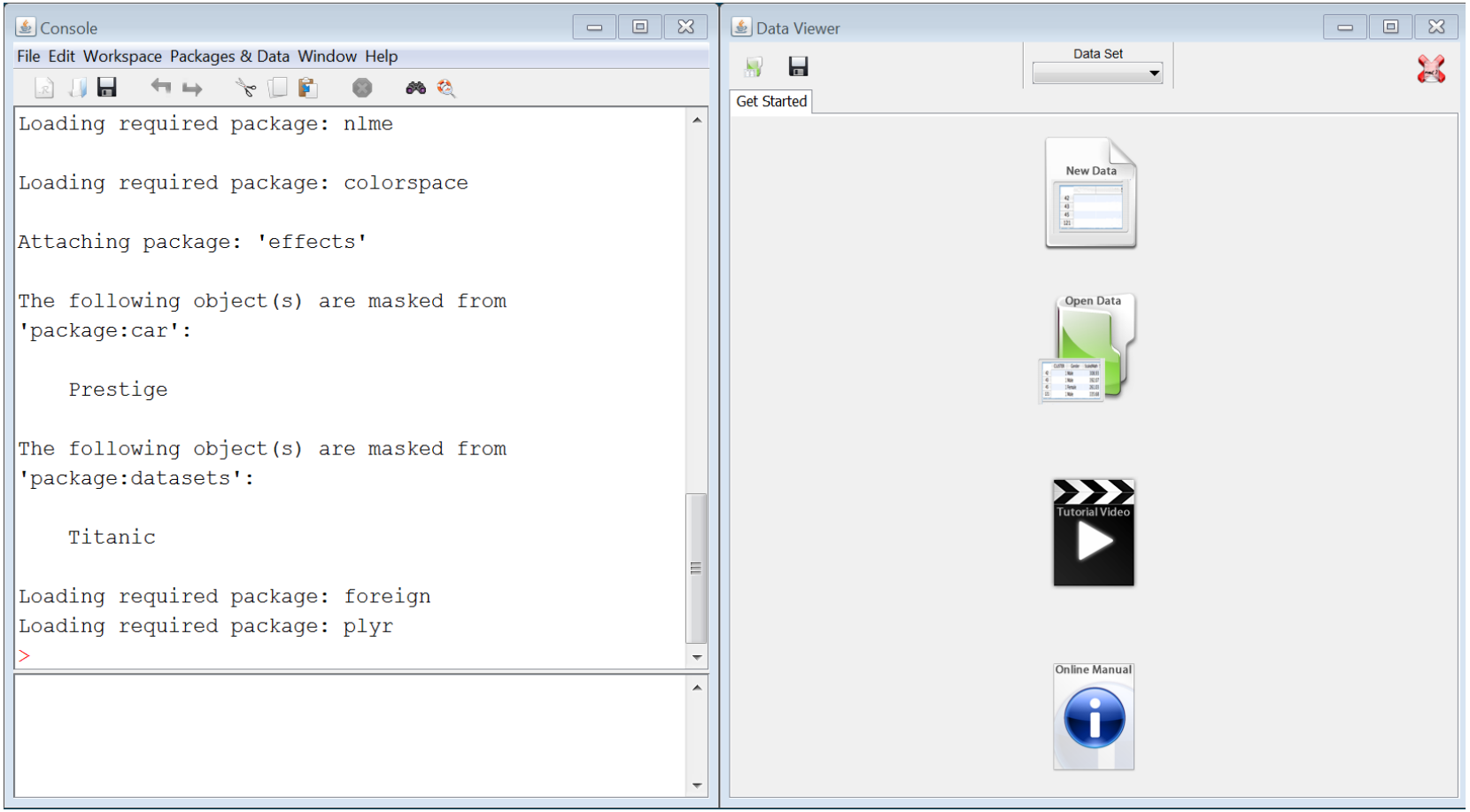
R GUI

- Rstudio
- Deducer
- Rcammandor
- RWinEdt
- Tinn-R
- JGR (Java Gui for R)
- Emacs + ESS

R STUDIO


RStudio
File Edit Code View Plots Session Build Debug Tools Help
Go to file/function
Addins
Project: (None)
Console ~/
R version 2.15.0 (2012-03-30)
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-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.
>
Environment History
Import Dataset
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.Rhistory 0 B Jun 29, 2016, 10:18 PM
campaign_analysis.xlsx 3.3 MB May 24, 2016, 8:33 PM
IBM
photos
R
SQL Server Management Studio
Visual Studio 2010

R DEDUCER


Console
File Edit Workspace Packages & Data Window Help
Loading required package: nlme
Loading required package: colorspace
Attaching package: 'effects'
The following object(s) are masked from 'package:car':
Prestige
The following object(s) are masked from 'package:datasets':
Titanic
Loading required package: foreign
Loading required package: plyr
>
Data Viewer
Data Set
Get Started
New Data
Open Data
Tutorial Video
Online Manual