

— WORKSHOP —

Applied Classical and Modern Multivariate Statistical Analysis

Module 1: An Introduction to R

Weixing Song, Juan Du

Department of Statistics  
Kansas State University

August 14, 2018

# Outline

- 1 The R Project
- 2 Installation of R
- 3 R Studio
- 4 R Packages and Installation
- 5 Load R Script: RStudio
- 6 R Basics

# Outline

- 1 The R Project
- 2 Installation of R
- 3 R Studio
- 4 R Packages and Installation
- 5 Load R Script: RStudio
- 6 R Basics

# Some Facts

**R** is a platform for statistical computing and visualization.

Versions of **R** exist for Windows, MacOS, Linux and various other Unix systems.

**R** is an implementation of the **S** language, which was principally developed by John Chambers.

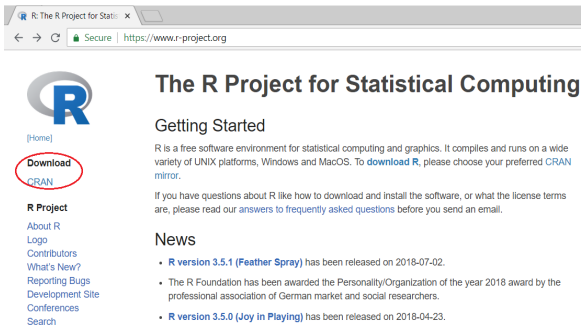
**R** is free and open source.

# Outline

- 1 The R Project
- 2 Installation of R
- 3 R Studio
- 4 R Packages and Installation
- 5 Load R Script: RStudio
- 6 R Basics

# Installation of R (I)

- Go to <https://www.r-project.org/>;
- On the left, click the link CRAN under the Download category;



The screenshot shows a web browser window with the URL <https://www.r-project.org/>. The page title is "The R Project for Statistical Computing". On the left sidebar, the "Download" link is circled in red, and "CRAN" is listed below it. The main content area includes a "Getting Started" section with introductory text and a "News" section with two bullet points.

**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 R](#), please choose your preferred [CRAN mirror](#).

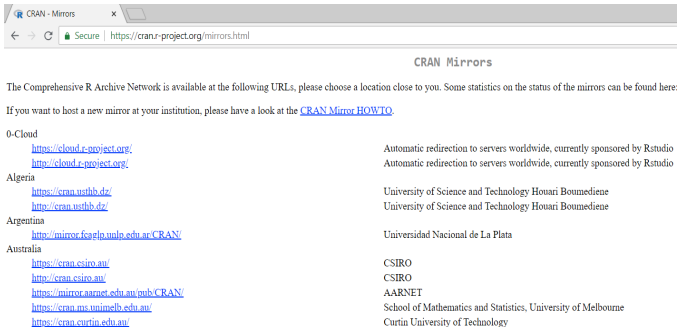
If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

News

- [R version 3.5.1 \(Feather Spray\)](#) has been released on 2018-07-02.
- The R Foundation has been awarded the Personality/Organization of the year 2018 award by the professional association of German market and social researchers.
- [R version 3.5.0 \(Joy in Playing\)](#) has been released on 2018-04-23.

# Installation of R (II)

- Choose an mirror page;



The screenshot shows a web browser window with the title "CRAN - Mirrors" and the address bar containing "https://cran.r-project.org/mirrors.html". The page content is as follows:

**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:

If you want to host a new mirror at your institution, please have a look at the [CRAN Mirror HOWTO](#).

**0-Cloud**

- <https://cloud.r-project.org/>
- <http://cloud.r-project.org/>

Automatic redirection to servers worldwide, currently sponsored by Rstudio

**Algeria**

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

University of Science and Technology Houari Boumediene

**Argentina**

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

Universidad Nacional de La Plata

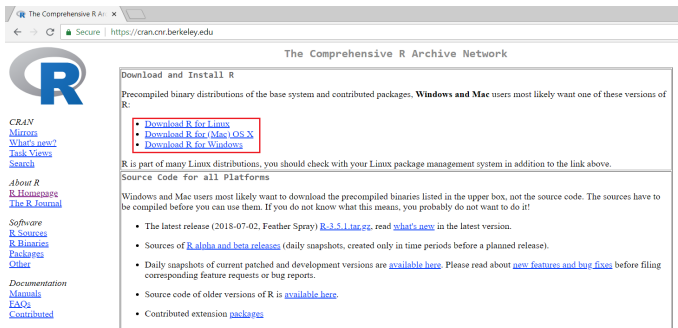
**Australia**

- <https://cran.csiro.au/>
- <http://cran.csiro.au/>
- <https://mirror.narinet.edu.au/pub/CRAN/>
- <https://cran.ms.unimelb.edu.au/>
- <https://cran.curtin.edu.au/>

CSIRO  
CSIRO  
AARNET  
School of Mathematics and Statistics, University of Melbourne  
Curtin University of Technology

# Installation of R (III)

- Download a proper version of **R** you need. We proceed with the Windows version.



The screenshot shows a web browser window with the address bar displaying 'https://cran.cnr.berkeley.edu'. The page title is 'The Comprehensive R Archive Network'. On the left side, there is a navigation menu with links for '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', and 'Contributed'. The main content area is titled 'Download and Install R' and contains the following text:

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.

Source Code for all Platforms

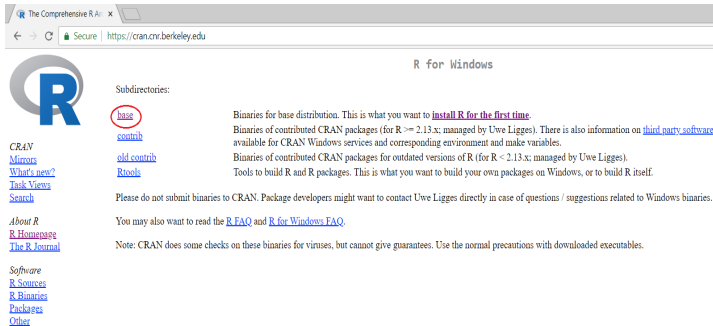
Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2018-07-02, Feather Spray) [R:3.5.1 tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#)



# Installation of R (IV)

- Choose base.



The screenshot shows a web browser window with the address bar displaying `https://cran.cnr.berkeley.edu`. The page title is "R for Windows". On the left side, there is a navigation menu with links for "CRAN", "Mirrors", "What's new?", "Task Views", "Search", "About R", "R Homepage", "The R Journal", "Software", "R Sources", "R Binaries", "Packages", and "Other". The main content area is titled "Subdirectories:" and lists several links: "base" (circled in red), "contrib", "old.contrib", and "Rtools". To the right of these links, there are descriptive paragraphs for each subdirectory. The "base" link is highlighted with a red circle.

**R for Windows**

Subdirectories:

- [base](#) (circled in red)
- [contrib](#)
- [old.contrib](#)
- [Rtools](#)

Binaries for base distribution. This is what you want to [install R for the first time](#).

Binaries of contributed CRAN packages (for R >= 2.13.x; managed by Uwe Ligges). There is also information on [third party software](#) available for CRAN Windows services and corresponding environment and make variables.

Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.x; managed by Uwe Ligges).

Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

Please do not submit binaries to CRAN. Package developers might want to contact Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.

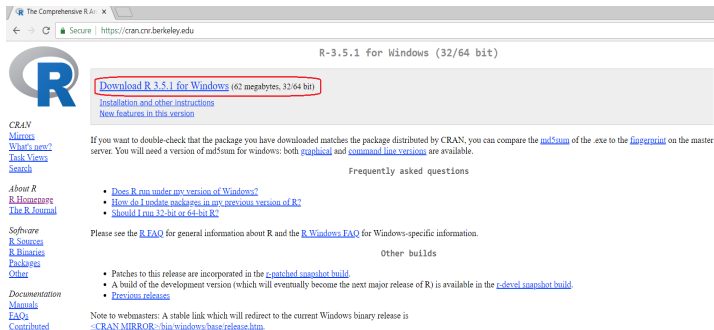
**CRAN**  
[Mirrors](#)  
[What's new?](#)  
[Task Views](#)  
[Search](#)

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

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

# Installation of R (V)

- Click the Download R x.x.x for Windows link to download the R.



The screenshot shows the CRAN website for R-3.5.1 for Windows (32/64 bit). The main heading is "R-3.5.1 for Windows (32/64 bit)". Below it, the link "Download R 3.5.1 for Windows (62 megabytes, 32/64 bit)" is highlighted with a red box. Other links include "Installation and other instructions" and "New features in this version".

**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](#)

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

Frequently asked questions

- Does R run under my version of Windows?
- 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.

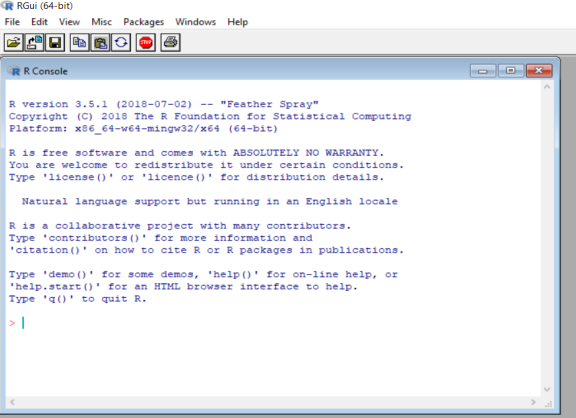
Other builds

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

# R interface

After installation, running **R**, you are expected to see the following



The screenshot shows the R GUI (64-bit) window. The title bar reads "RGui (64-bit)". The menu bar includes "File", "Edit", "View", "Misc", "Packages", "Windows", and "Help". The toolbar contains icons for file operations and execution. The main window is titled "R Console" and displays the following text:

```
R version 3.5.1 (2018-07-02) -- "Feather Spray"  
Copyright (C) 2018 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.  
  
> |
```

# Outline

- 1 The R Project
- 2 Installation of R
- 3 R Studio**
- 4 R Packages and Installation
- 5 Load R Script: RStudio
- 6 R Basics

# What is R-Studio?

- RStudio is a platform including many integrated tools to help you be more productive with **R**.
- RStudio is also free!

# Installation of RStudio (I)

- Go to <https://www.rstudio.com/>, look for the Download RStudio link, or
- Go to <https://www.rstudio.com/products/rstudio/download/>
- Choose the **free** download link

Download RStudio - RStudio

Secure | <https://www.rstudio.com/products/rstudio/download/>

RStudio

Products Resources Pricing About Us Blogs

### Choose Your Version of RStudio

RStudio is a set of integrated tools designed to help you be more productive with R. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools for plotting, viewing history, debugging and managing your workspace. [Learn More](#) about RStudio features.

RStudio Desktop Open Source License	RStudio Desktop Commercial License	RStudio Server Open Source License	RStudio Server Pro Commercial License	RStudio Server Pro + RStudio Connect Commercial License
FREE	\$995 per year	FREE	\$9,995 per year	\$29,995 per year
<a href="#">DOWNLOAD</a> <small>Learn More</small>	<a href="#">BUY</a> <small>Learn More</small>	<a href="#">DOWNLOAD</a> <small>Learn More</small>	<a href="#">DOWNLOAD</a> <small>Learn More</small>	<a href="#">TALK</a> <small>Learn More</small>

# Installation of RStudio (II)

- Choose the right version you want to use. We use the Windows version.

Download RStudio - RStudio

Products Resources Pricing About Us Blogs

**RStudio Desktop 1.1.456** — Release Notes

RStudio requires R 3.0.1+. If you don't already have R, download it here.

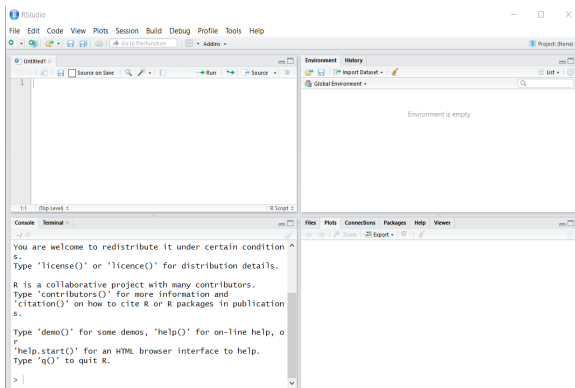
Linux users may need to import RStudio's public code-signing key prior to installation, depending on the operating system's security policy.

**Installers for Supported Platforms**

Installers	Size	Date	MDS
RStudio 1.1.456 - Windows Vista/7/8/10	85.8 MB	2018-07-19	24ca3Fe0dad0187aab040fbfb9dc2b5ad
RStudio 1.1.456 - Mac OS X 10.6+ (64-bit)	74.5 MB	2018-07-19	4fc4f4f700450142bf96dc1a5b1dc556
RStudio 1.1.456 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-07-19	3493f9d5839e3a3d697f46b7bb1ce961
RStudio 1.1.456 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	97.4 MB	2018-07-19	863ae806120358fa0146e4d14cd75be4
RStudio 1.1.456 - Ubuntu 16.04+/Debian 9+ (64-bit)	64.9 MB	2018-07-19	d96e63548c2add890bac633bdb883f32
RStudio 1.1.456 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.1 MB	2018-07-19	1df56c7cd80e2634f8a9fd011ca1fb3d
RStudio 1.1.456 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	90.6 MB	2018-07-19	5e77094a88fdbd555db6d35700752462

# RStudio Interfact

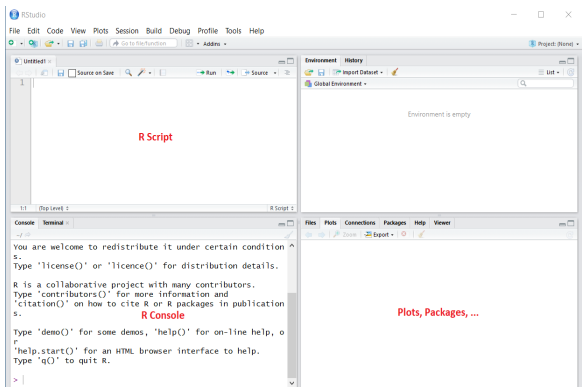
After installation, run RStudio, you are expect to see





# RStudio Interfact

After installation, run RStudio, you are expect to see



# Outline

- 1 The R Project
- 2 Installation of R
- 3 R Studio
- 4 R Packages and Installation**
- 5 Load R Script: RStudio
- 6 R Basics

# R Packages

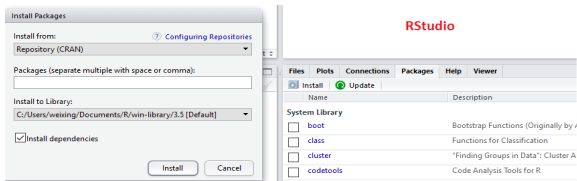
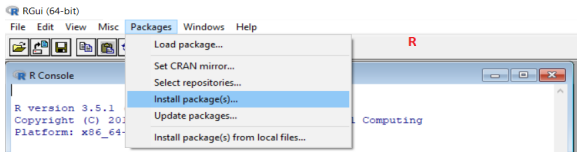
R packages are collections of functions and data sets developed by the community.

The power of R greatly increased by R packages by improving existing base R functionalities, or by adding new ones.

By January 27, 2017, CRAN, the global repository of open-source packages that extend the capabilities of R, has 10,000 R-packages available for download. Note that this number doesn't even include all the R packages outside of CRAN. There are also another 1294 packages for genomic analysis in the BioConductor repository, hundreds of R packages published only on GitHub, commercial R packages from vendors such as Microsoft and Oracle, and an unknowable number of private, unpublished packages.

# How to install R packages?

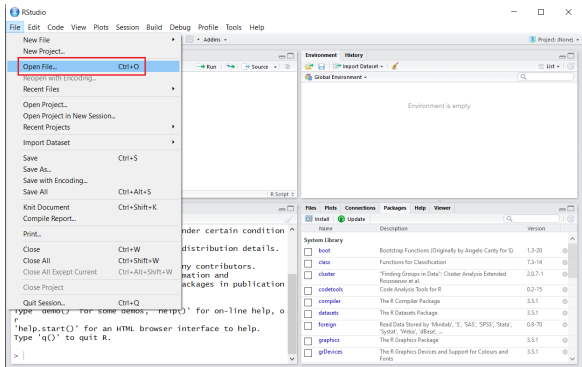
- In R or RStudio Console, issue the following commands
  - > `install.packages("package.name")`
  - > `library("package.name")`
- Using menu tools in R or Rstudio



# Outline

- 1 The R Project
- 2 Installation of R
- 3 R Studio
- 4 R Packages and Installation
- 5 Load R Script: RStudio**
- 6 R Basics

To load an R script file, simply choose **Open File ...** from File menu.



# Required Packages

In today's workshop, we need the following packages:

```
aplpack, biotools, boot, car, cluster, e1071, ggfortify, ggplot2,  
graphics, ISLR, kernlab, klaR, lattice, latticeExtra, MASS, mclust,  
MVA, mvtnorm, plot3D, plot3Drgl, plotly, pracma, psych, rattle,  
RcolorBrewer, rgl, rlang, rpart, rrcov, scatterplot3d, tidyverse, tree
```

Use the following command in RStudio console to load these packages

```
> install.packages(c("ggplot2", "MASS"))
```

# Outline

- 1 The R Project
- 2 Installation of R
- 3 R Studio
- 4 R Packages and Installation
- 5 Load R Script: RStudio
- 6 R Basics**



## Working Directory

- Check the current working directory

```
> getwd()
```

- Set the working directory

```
> setwd("path.to.folder")
```

## Load Data Set into R

- Load a TXT file into R

```
mydata=read.table("path.to.file",header = FALSE)
```

- Load an EXCEL file into R

```
mydata=read.csv("path.to.file",header = FALSE)
```

- Load SPSS, STAT, Minitab data into R

```
> library(foreign)
> mydata=read.spss("path.to.file")    SPSS
> mydata=read.dta("path.to.file")    Stata
> mydata=read.mtp("path.to.file")    Minitab
```

- Load SAS into R

```
> library(sas7bdat)
> mydata=read.sas7bdat("path.to.file")    SAS
```

### Note:

- R is case sensitive!
- When encounter questions, try `?`, `help()` functions, and if still no clue, GOOGLE!