R

2019-09-03

Contents

Chapter 1

Lecture

- : 1213 (13:00~16:00)
- :
- : 042-860-4372, haseong@kribb.re.kr (1143)
- site: https://greendaygh.github.io/Rstat2019/

1.1 Goal

- . R
.

1.2 References

- Using R for Introductory Statistics by John Verzani
  - Free version of 1st Edition
    - https://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf
    - http://cbb.sjtu.edu.cn/~mywu/bi217/usingR.pdf
  - Second edition
    - https://www.crcpress.com/Using-R-for-Introductory-Statistics-Second-Edition/Verzani/p/book/9781466590731
- R for Data Science (https://r4ds.had.co.nz, https://github.com/hadley)
- https://resources.rstudio.com/
- (,)

1.3 Evaluation

- 50% / 50% / 80 S, 80 U

1.4 Schedule

- 1 - R basics / introduction of data

- 2 - Univariate data – Summary statistics (, ,)
- 3 - Bivariate data – Correlation / Independence (, ,)
- 4 - Multivariate data – R data structure (, R , R)
- 5 - Populations – Families of distributions
- 6 - Sampling – Distribution and CLT ,
- 7 - Statistical inference
- 8 - Confidence intervals
- 9 - Significance test - parameteric ( )
- 10 - Significance test – non parametric ( )
- 11 - Goodness of fit - parametric ( )
- 12 - Goodness of fit – non parametirc ( )
- 13 - Linear regression – basics & simple LR
- 14 - Multiple linear regression
- 15 - Analysis of variance
- 16 - Logistic / Non-linear regression /
- 9/25 ( )

1.5 References

- R https://www.r-project.org/
- Rstudio https://www.rstudio.com/
- Packages for biologists https://www.bioconductor.org/
- R (, , ,)
- https://cran.r-project.org/doc/manuals/r-release/R-intro.html
- https://cran.r-project.org/doc/manuals/r-release/R-data.html
- https://cran.r-project.org/doc/manuals/r-release/R-admin.html
- R ebooks
- https://bookdown.org/

Chapter 2

Introductin

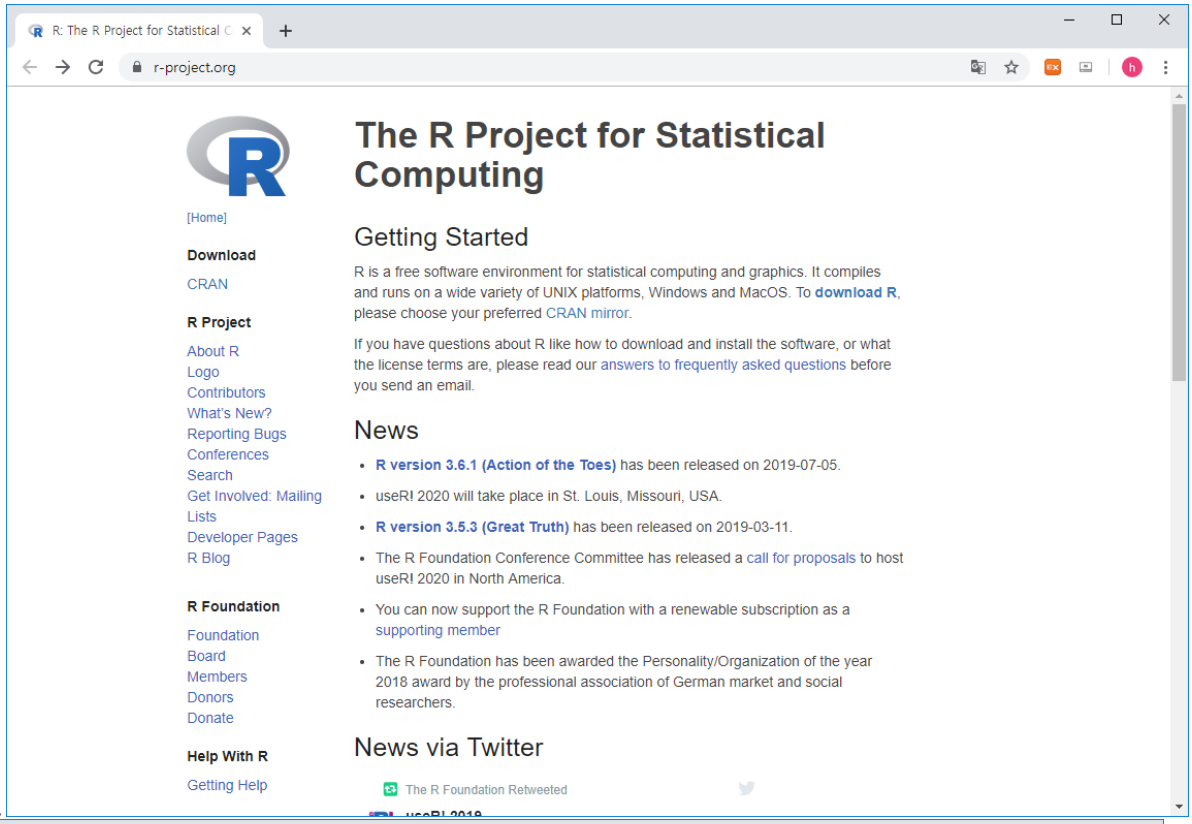
2.1 What is R / Rstudio

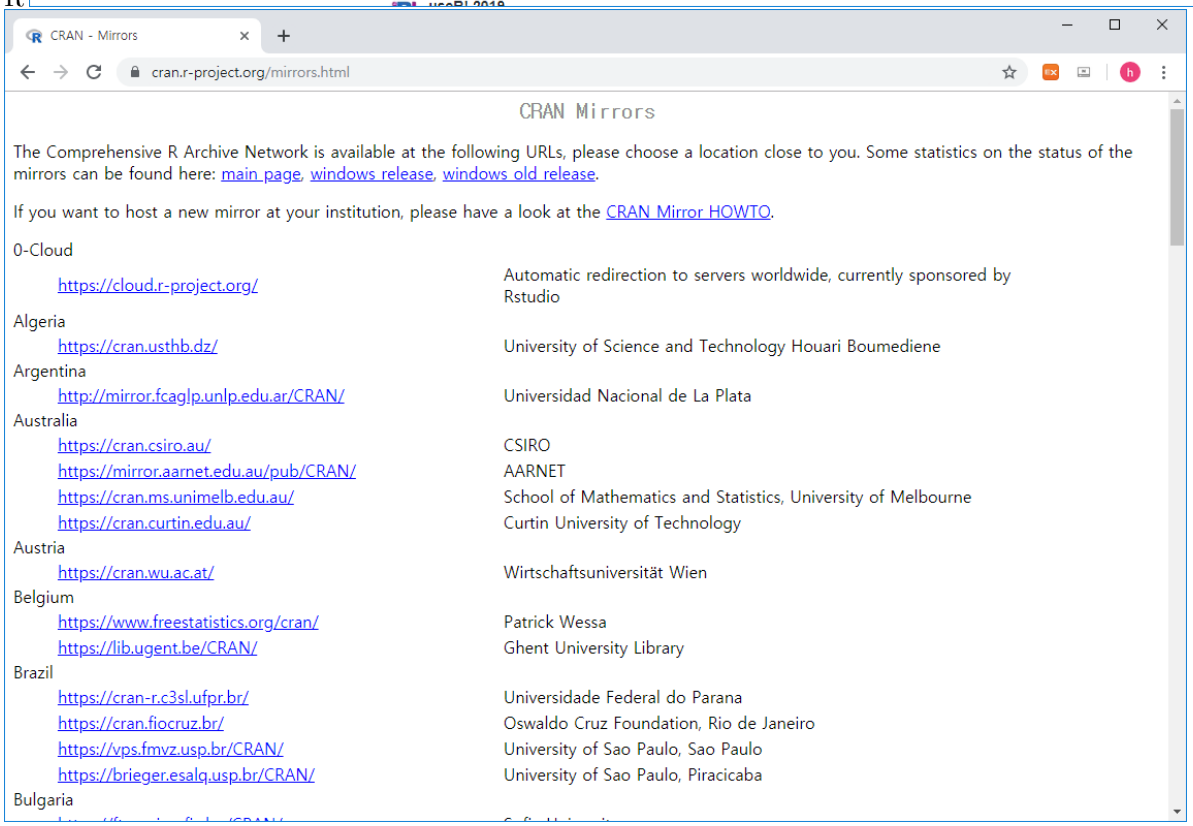
- R is a programming language that runs computations (https://www.r-project.org/)
- RStudio is an integrated development environment (IDE) that provides an interface for the programming (https://www.rstudio.com/)

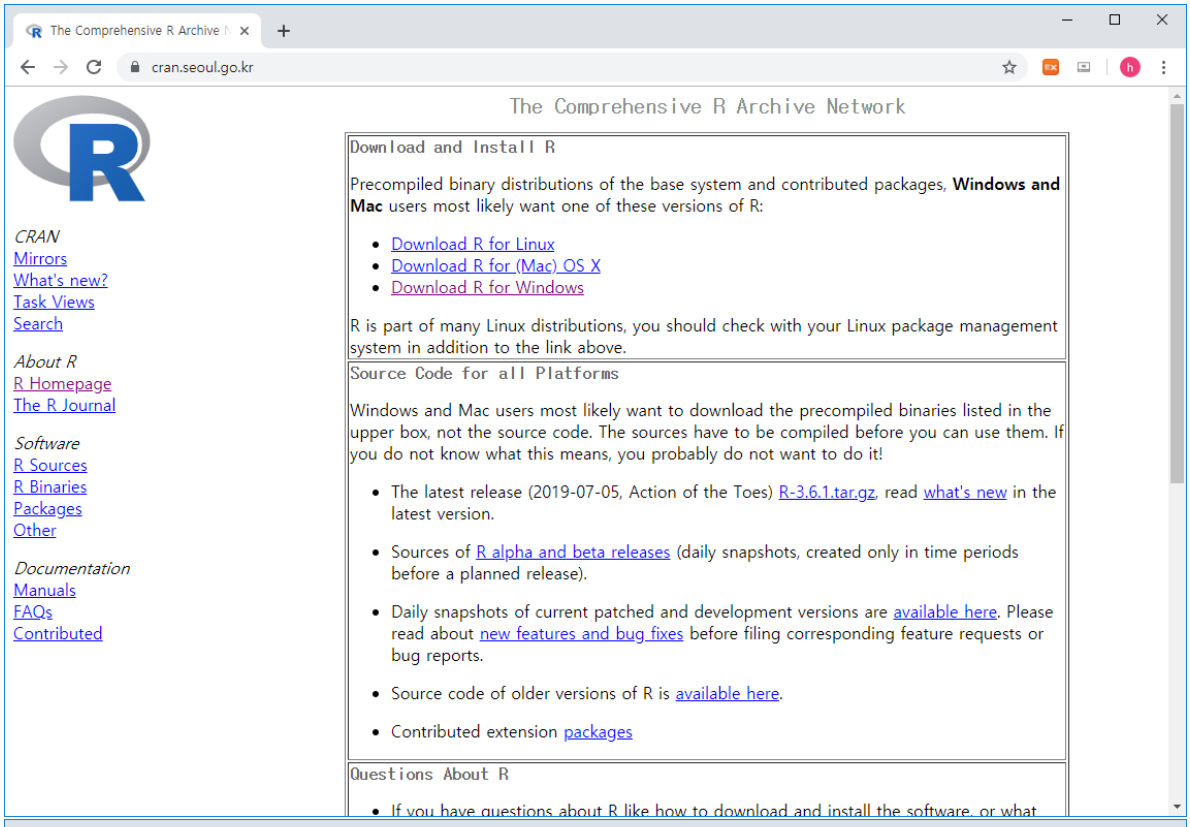
2.2 R / Rstudio installation

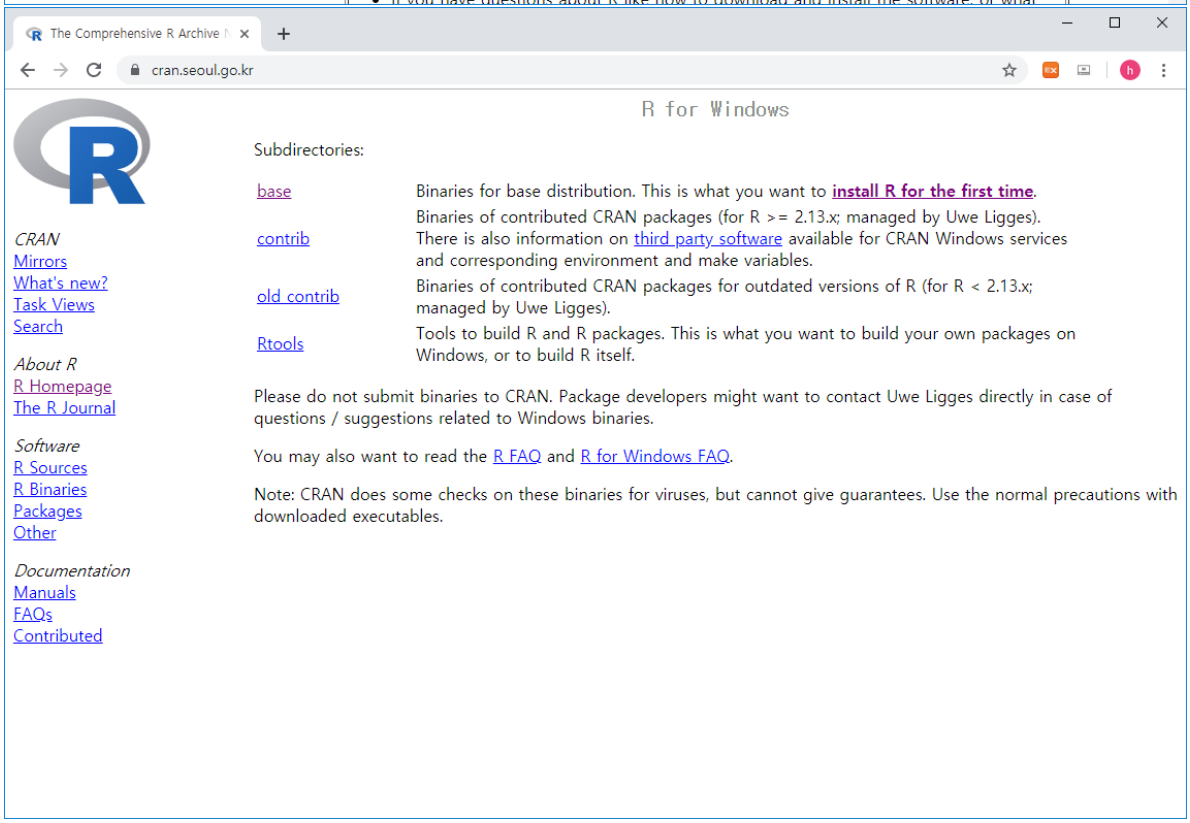
- Install R first and then install RStudio second

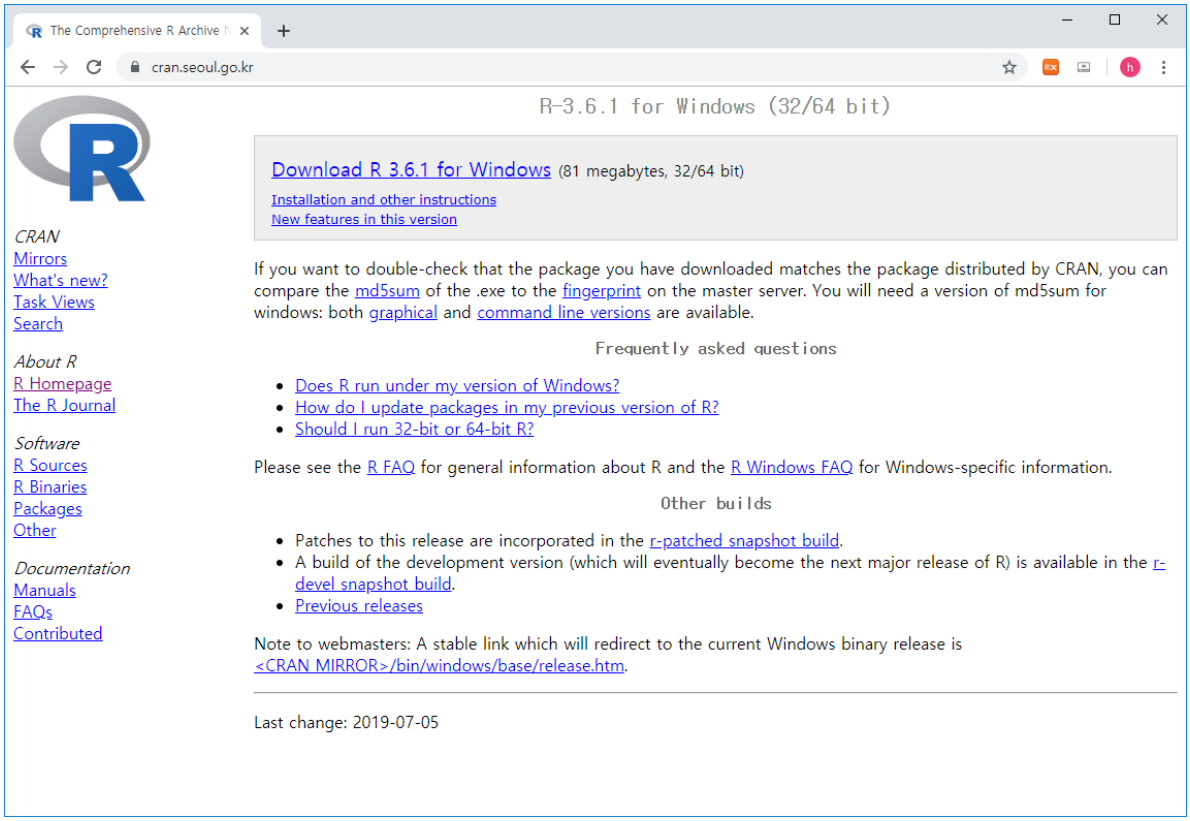
- R


R: The R Project for Statistical C
r-project.org
The R Project for Statistical Computing
[Home]
Download
CRAN
R Project
About R
Logo
Contributors
What's New?
Reporting Bugs
Conferences
Search
Get Involved: Mailing Lists
Developer Pages
R Blog
R Foundation
Foundation
Board
Members
Donors
Donate
Help With R
Getting Help
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 3.6.1 (Action of the Toes) has been released on 2019-07-05.
useR! 2020 will take place in St. Louis, Missouri, USA.
R version 3.5.3 (Great Truth) has been released on 2019-03-11.
The R Foundation Conference Committee has released a call for proposals to host useR! 2020 in North America.
You can now support the R Foundation with a renewable subscription as a supporting member
The R Foundation has been awarded the Personality/Organization of the year 2018 award by the professional association of German market and social researchers.
News via Twitter
The R Foundation Retweeted

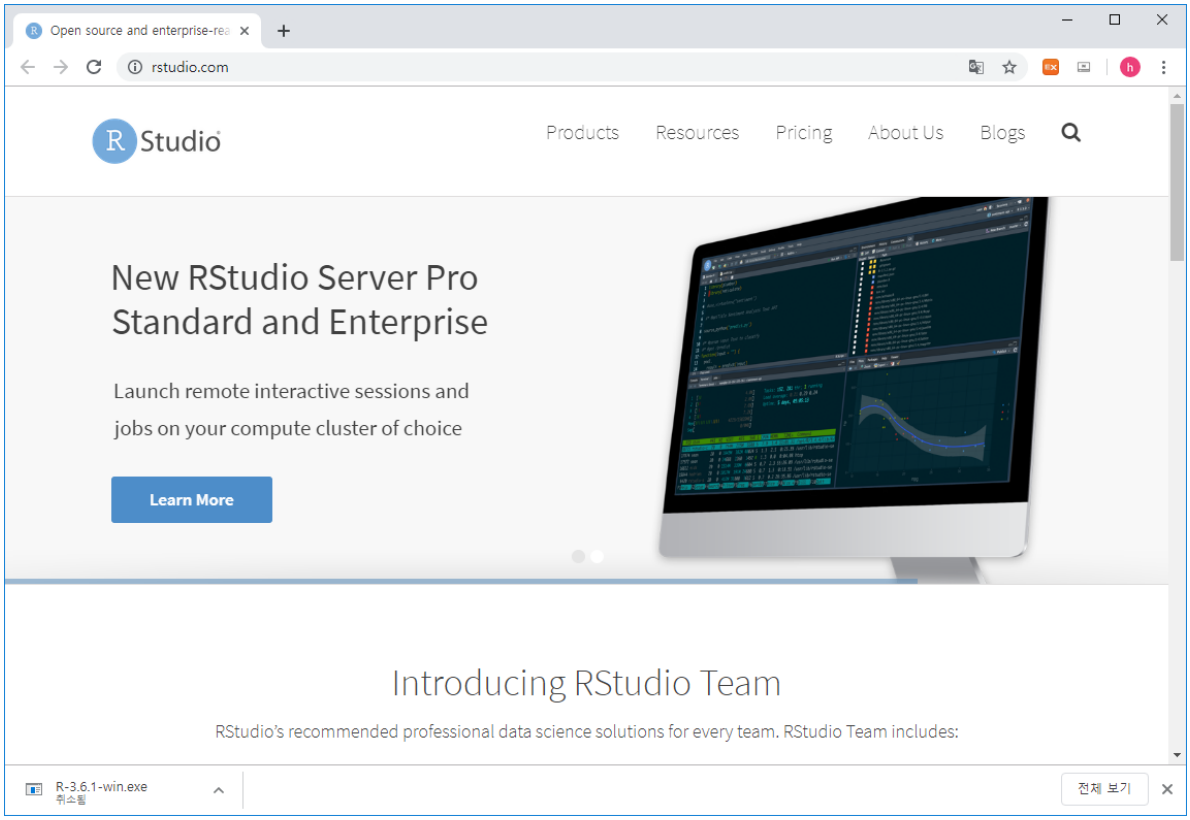

CRAN - Mirrors
cran.r-project.org/mirrors.html
CRAN Mirrors
The Comprehensive R Archive Network is available at the following URLs, please choose a location close to you. Some statistics on the status of the mirrors can be found here: main page, windows release, windows old release.
If you want to host a new mirror at your institution, please have a look at the CRAN Mirror HOWTO.
0-Cloud
https://cloud.r-project.org/ Automatic redirection to servers worldwide, currently sponsored by Rstudio
Algeria
https://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/ CSIRO
https://mirror.aarnet.edu.au/pub/CRAN/ AARNET
https://cran.ms.unimelb.edu.au/ School of Mathematics and Statistics, University of Melbourne
https://cran.curtin.edu.au/ Curtin University of Technology
Austria
https://cran.wu.ac.at/ Wirtschaftsuniversität Wien
Belgium
https://www.freestatistics.org/cran/ Patrick Wessa
https://lib.ugent.be/CRAN/ Ghent University Library
Brazil
https://cran-r.c3sl.ufpr.br/ Universidade Federal do Parana
https://cran.fiocruz.br/ Oswaldo Cruz Foundation, Rio de Janeiro
https://vps.fmvz.usp.br/CRAN/ University of Sao Paulo, Sao Paulo
https://brieger.esalq.usp.br/CRAN/ University of Sao Paulo, Piracicaba
Bulgaria

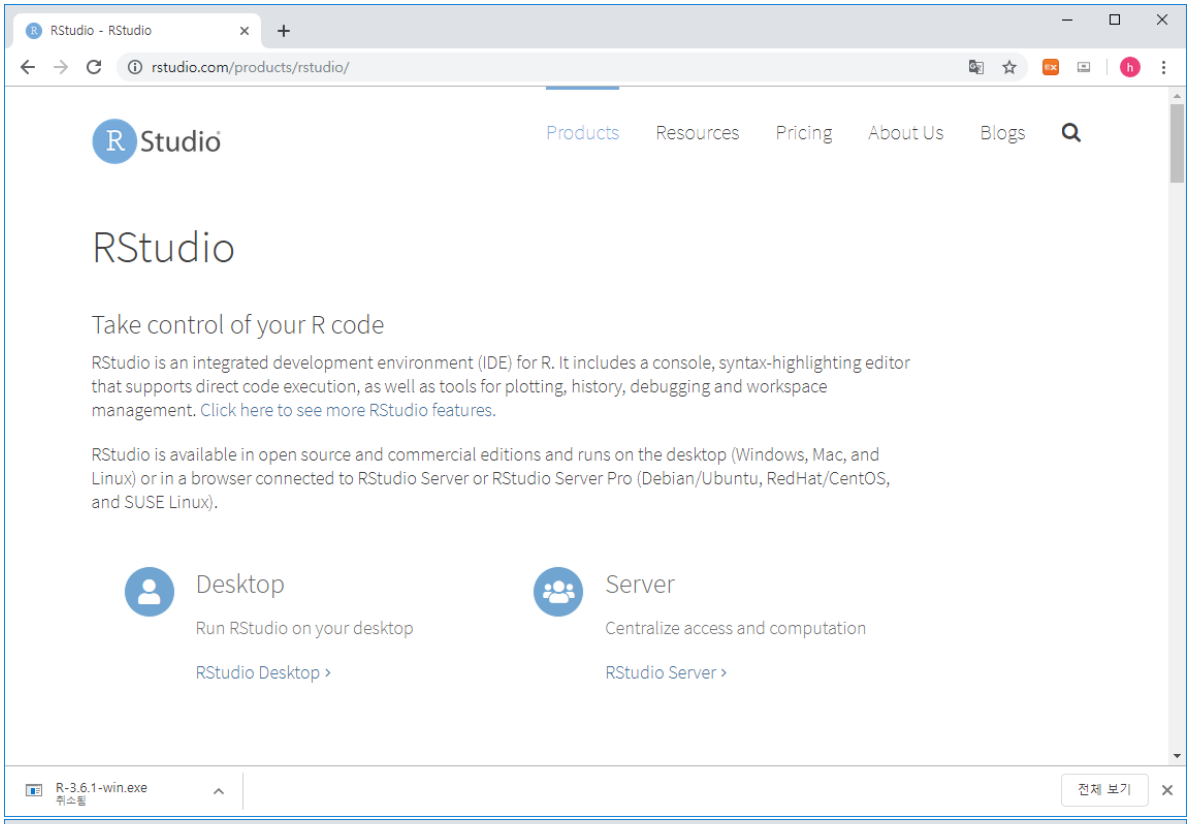

The Comprehensive R Archive N
cran.seoul.go.kr
The Comprehensive R Archive Network
CRAN
Mirrors
What's new?
Task Views
Search
About R
R Homepage
The R Journal
Software
R Sources
R Binaries
Packages
Other
Documentation
Manuals
FAQs
Contributed
Download and Install R
Precompiled binary distributions of 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 (2019-07-05, Action of the Toes) R-3.6.1.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
Questions About R

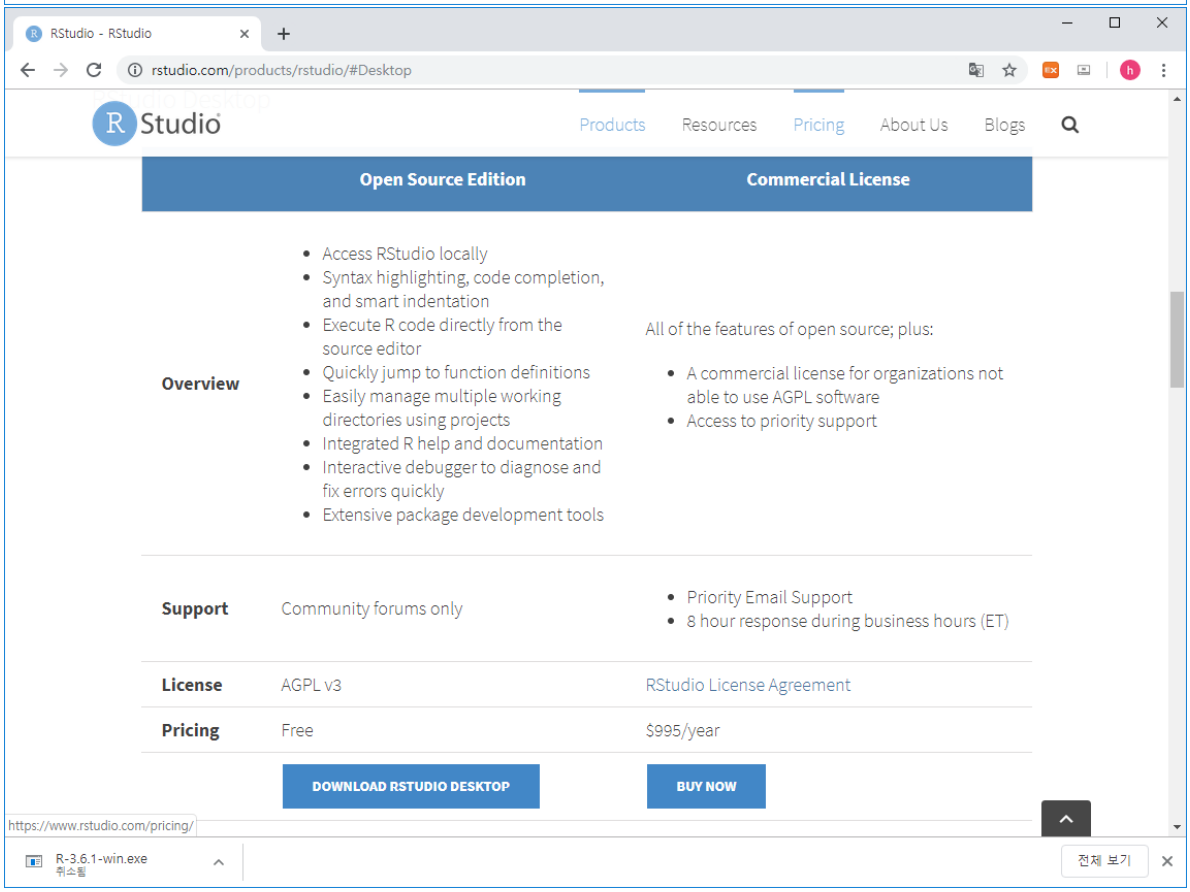

The Comprehensive R Archive N
cran.seoul.go.kr
R for Windows
CRAN
Mirrors
What's new?
Task Views
Search
About R
R Homepage
The R Journal
Software
R Sources
R Binaries
Packages
Other
Documentation
Manuals
FAQs
Contributed
Subdirectories:
base
Binaries for base distribution. This is what you want to install R for the first time.
contrib
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.
old contrib
Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.x; managed by Uwe Ligges).
Rtools
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.

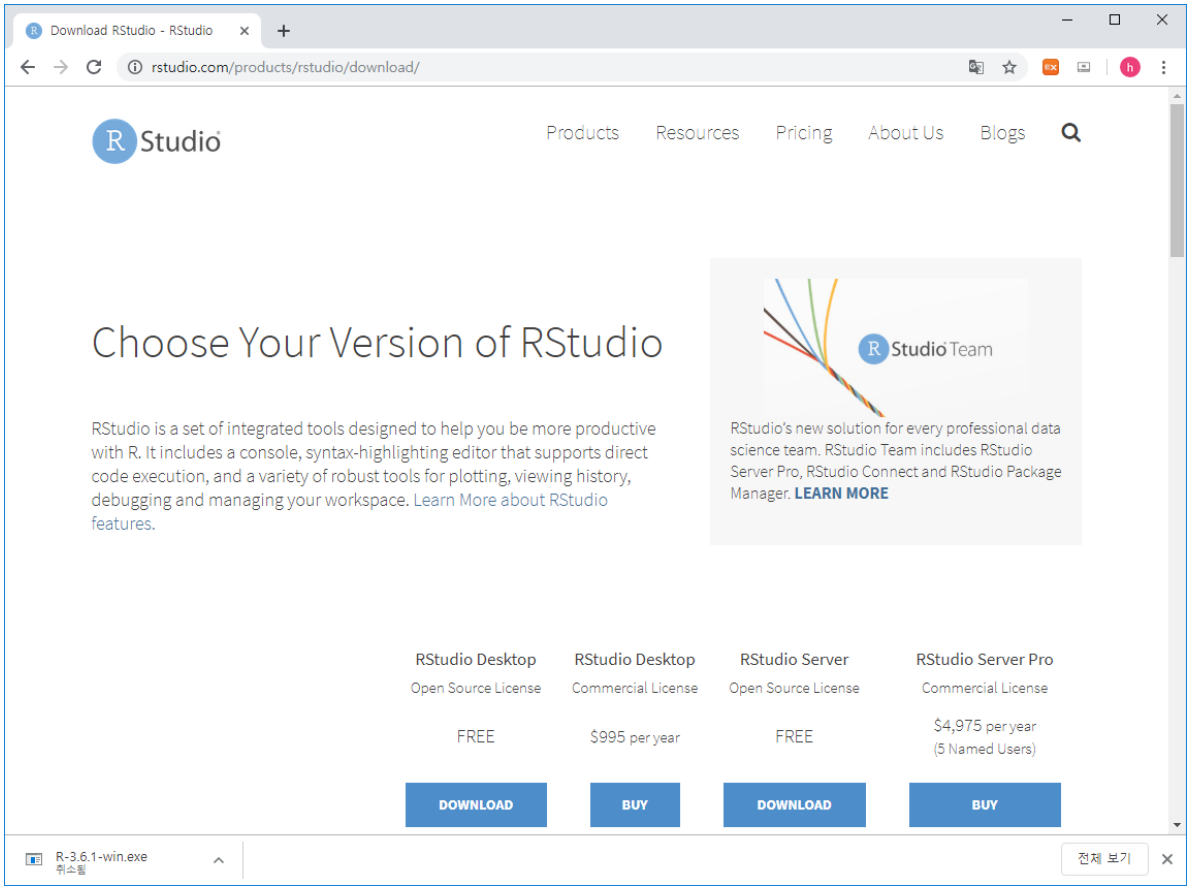

The Comprehensive R Archive
cran.seoul.go.kr
R-3.6.1 for Windows (32/64 bit)
Download R 3.6.1 for Windows (81 megabytes, 32/64 bit)
Installation and other instructions
New features in this version
If you want to double-check that the package you have downloaded matches the package distributed by CRAN, 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 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: 2019-07-05
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

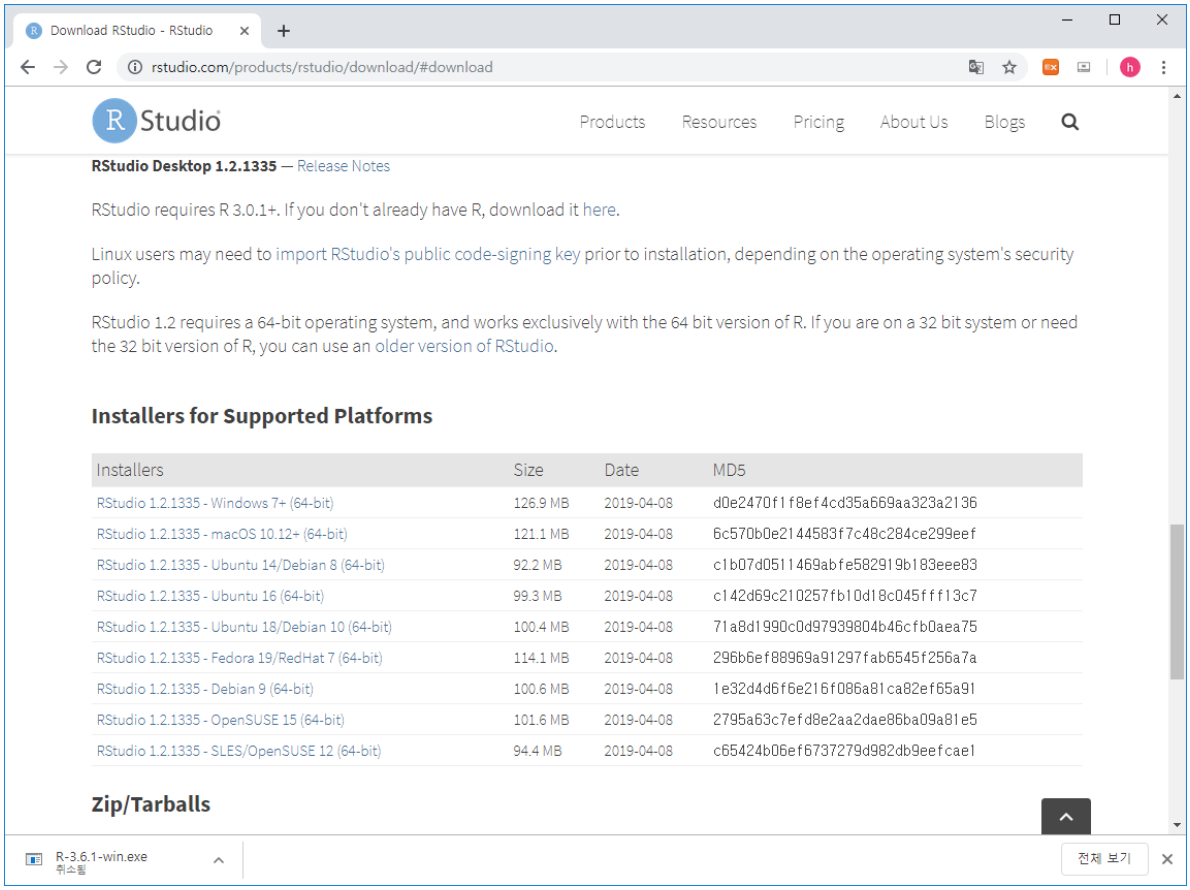
- Rstudio


Open source and enterprise-rea
rstudio.com
RStudio
Products
Resources
Pricing
About Us
Blogs
New RStudio Server Pro
Standard and Enterprise
Launch remote interactive sessions and
jobs on your compute cluster of choice
Learn More
Introducing RStudio Team
RStudio's recommended professional data science solutions for every team. RStudio Team includes:
R-3.6.1-win.exe
취소됨
전체 보기

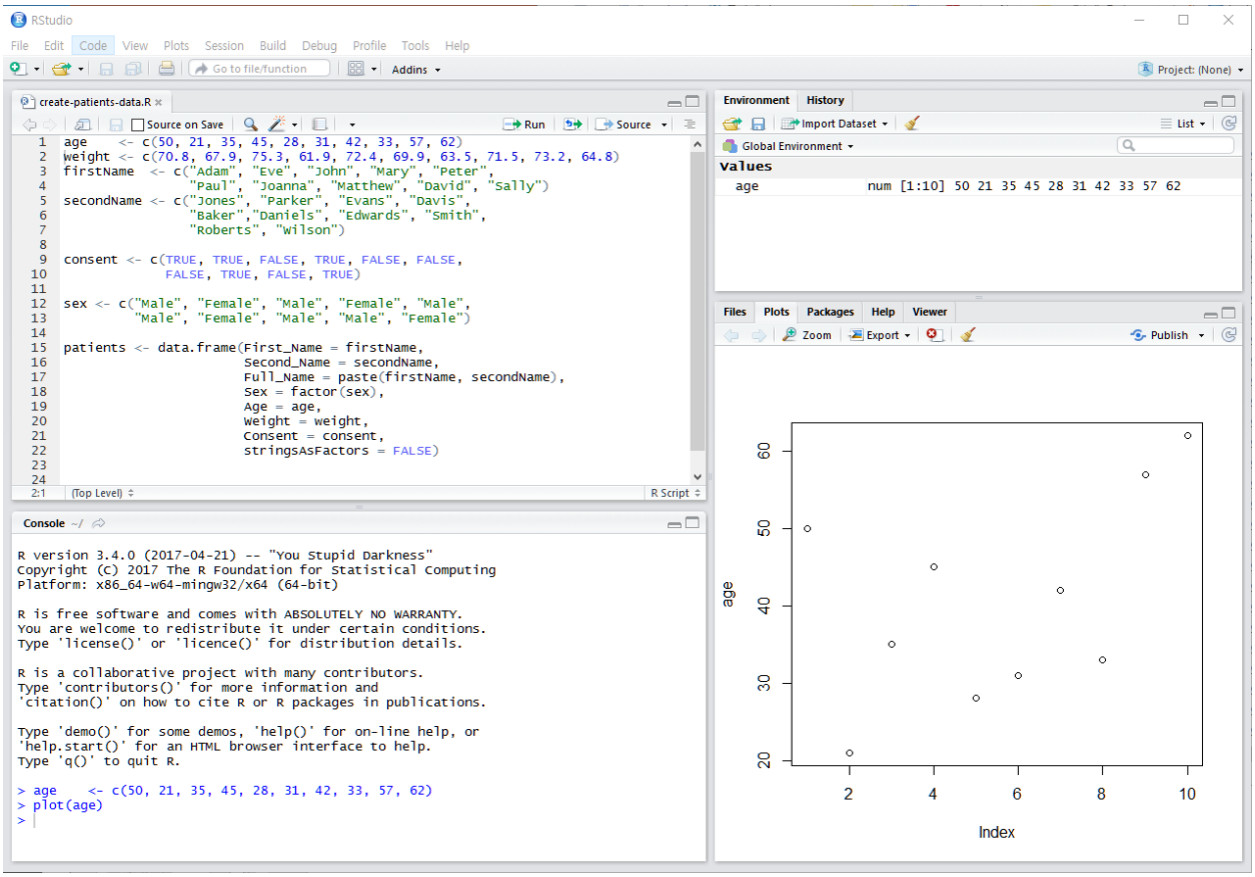

RStudio - RStudio
rstudio.com/products/rstudio/
RStudio
Products
Resources
Pricing
About Us
Blogs
RStudio
Take control of your R code
RStudio is an integrated development environment (IDE) for R. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management. Click here to see more RStudio features.
RStudio is available in open source and commercial editions and runs on the desktop (Windows, Mac, and Linux) or in a browser connected to RStudio Server or RStudio Server Pro (Debian/Ubuntu, RedHat/CentOS, and SUSE Linux).
Desktop
Run RStudio on your desktop
RStudio Desktop ›
Server
Centralize access and computation
RStudio Server ›
R-3.6.1-win.exe
취소됨
전체 보기


RStudio - RStudio
rstudio.com/products/rstudio/#Desktop
RStudio
Products
Resources
Pricing
About Us
Blogs
Open Source Edition
Commercial License
Overview
Access RStudio locally
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
All of the features of open source; plus:
A commercial license for organizations not able to use AGPL software
Access to priority support
Support
Community forums only
Priority Email Support
8 hour response during business hours (ET)
License
AGPL v3
RStudio License Agreement
Pricing
Free
$995/year
DOWNLOAD RSTUDIO DESKTOP
BUY NOW
https://www.rstudio.com/pricing/
R-3.6.1-win.exe
취소됨
전체 보기


Download RStudio - RStudio
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 Team
RStudio's new solution for every professional data science team. RStudio Team includes RStudio Server Pro, RStudio Connect and RStudio Package Manager. LEARN MORE
RStudio Desktop
Open Source License
FREE
DOWNLOAD
RStudio Desktop
Commercial License
$995 per year
BUY
RStudio Server
Open Source License
FREE
DOWNLOAD
RStudio Server Pro
Commercial License
$4,975 per year
(5 Named Users)
BUY
R-3.6.1-win.exe
취소됨
전체 보기


Download RStudio - RStudio
rstudio.com/products/rstudio/download/#download
RStudio
Products
Resources
Pricing
About Us
Blogs
RStudio Desktop 1.2.1335 — 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.
RStudio 1.2 requires a 64-bit operating system, and works exclusively with the 64 bit version of R. If you are on a 32 bit system or need the 32 bit version of R, you can use an older version of RStudio.
Installers for Supported Platforms
Installers Size Date MD5
RStudio 1.2.1335 - Windows 7+ (64-bit) 126.9 MB 2019-04-08 d0e2470f1f8ef4cd35a669aa323a2136
RStudio 1.2.1335 - macOS 10.12+ (64-bit) 121.1 MB 2019-04-08 6c570b0e2144583f7c48c284ce299eef
RStudio 1.2.1335 - Ubuntu 14/Debian 8 (64-bit) 92.2 MB 2019-04-08 c1b07d0511469abfe582919b183eee83
RStudio 1.2.1335 - Ubuntu 16 (64-bit) 99.3 MB 2019-04-08 c142d69c210257fb10d18c045fff13c7
RStudio 1.2.1335 - Ubuntu 18/Debian 10 (64-bit) 100.4 MB 2019-04-08 71a8d1990c0d97939804b46cfb0aea75
RStudio 1.2.1335 - Fedora 19/RedHat 7 (64-bit) 114.1 MB 2019-04-08 296b6ef88969a91297fab6545f256a7a
RStudio 1.2.1335 - Debian 9 (64-bit) 100.6 MB 2019-04-08 1e32d4d6f6e216f086a81ca82ef65a91
RStudio 1.2.1335 - OpenSUSE 15 (64-bit) 101.6 MB 2019-04-08 2795a63c7efd8e2aa2dae86ba09a81e5
RStudio 1.2.1335 - SLES/OpenSUSE 12 (64-bit) 94.4 MB 2019-04-08 c65424b06ef6737279d982db9eefcae1
Zip/Tarballs
R-3.6.1-win.exe
취소됨
전체 보기

2.3 Rstudio interface



2.4 R code basics

- Basics: Console: Where you enter in commands. Running code: The act of telling R to perform an act by giving it commands in the console. Objects: Where values are saved in R. We'll show you how to assign values to objects and how to display the contents of objects. Data types: Integers, doubles/numerics, logicals, and characters. Vectors: A series of values. These are created using the c() function, where c() stands for "combine" or "concatenate." For example: c(6, 11, 13, 31, 90, 92). Factors: Categorical data are represented in R as factors. Data frames: Data frames are like rectangular spreadsheets: they are representations of datasets in R where the rows correspond to observations and the columns correspond to variables that describe the observations. We'll cover data frames later in Section 1.4. Conditionals: Testing for equality in R using == (and not = which is typically used for assignment). Ex: $2 + 1 == 3$ compares $2 + 1$ to 3 and is correct R code, while $2 + 1 = 3$ will return an error. Boolean algebra: TRUE/FALSE statements and mathematical operators such as $<$ (less than), $<=$ (less than or equal), and != (not equal to). Logical operators: & representing "and" as well as | representing "or." Ex: $(2 + 1 == 3)$ & $(2 + 1 == 4)$ returns FALSE since both clauses are not TRUE (only the first clause is TRUE). On the other hand, $(2 + 1 == 3)$ | $(2 + 1 == 4)$ returns TRUE since at least one of the two clauses is TRUE. Functions, also called commands: Functions perform tasks in R. They take in inputs called arguments and return outputs. You can either manually specify a function's arguments or use the function's default values.

Chapter 3

introduction of data

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter **??**. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter **??**.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 3.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 3.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2019) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

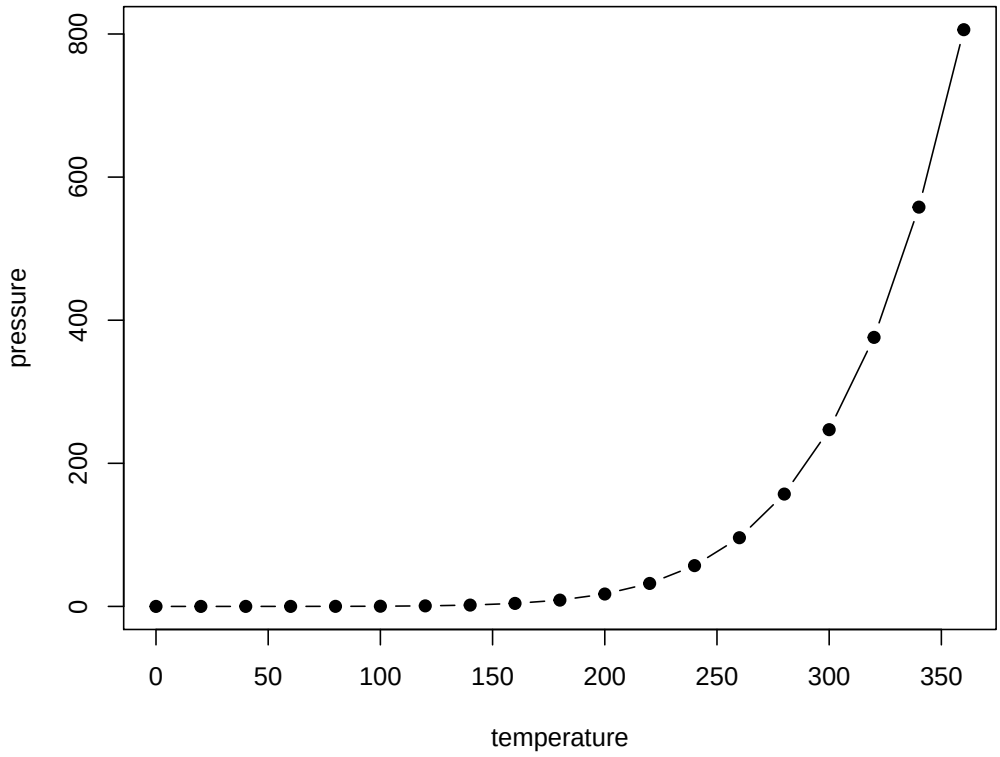


Figure 3.1: Here is a nice figure!

Table 3.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2019). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.13.