

《数据分析与R语言》

R语言介绍

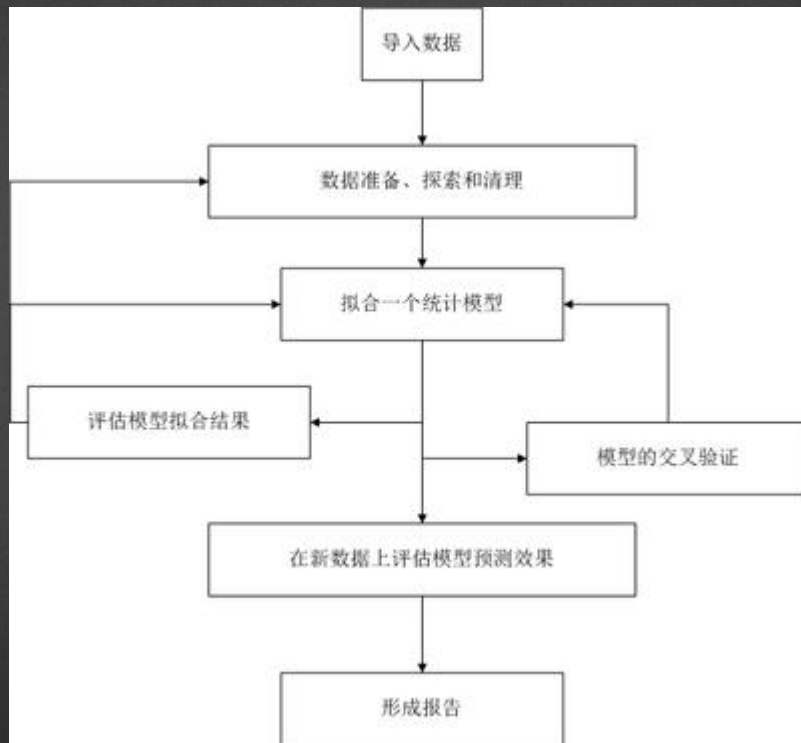


引入





引入





为何要使用R

❧ R语言由新西兰奥克兰大学的Ross Ihaka 和Robert Gentleman两人共同发明





为何要使用R

- 🦋 R是免费的开源软件
- 🦋 全面的统计研究平台，提供了各种各样的数据分析技术
- 🦋 R是一个程序设计语言，可以很容易地通过用户自定义的函数扩展
- 🦋 R拥有顶尖的制图功能
- 🦋 可以从多个数据源获取数据并将其转化为可用的形式
- 🦋 R可以运行在多种平台上



为何要使用R

R制图示例

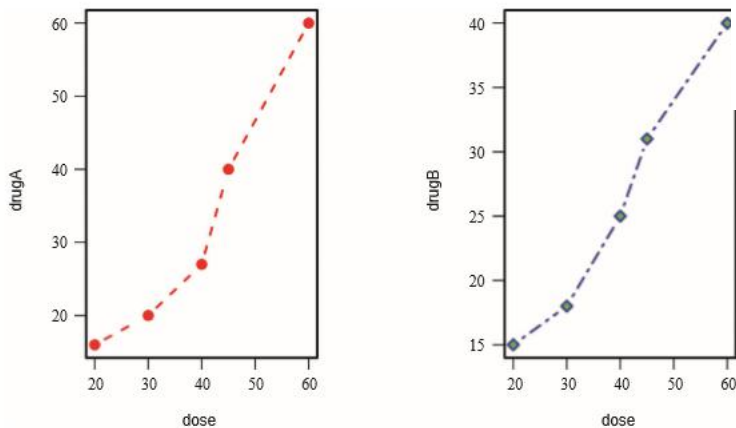


图 3-7 药物 A 和药物 B 剂量与响应的折线图

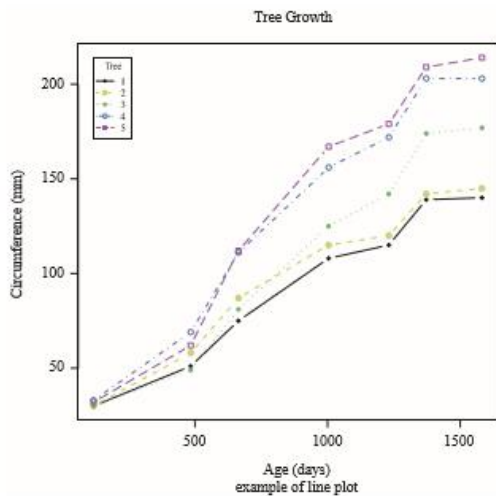


图 11-19 展示五种橘树生长状况的折线图

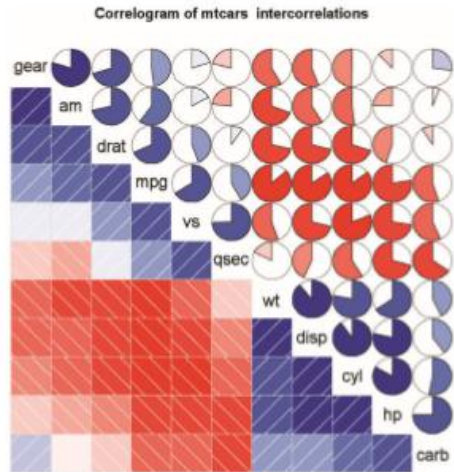


图 11-20 mtcars 数据框中变量的相关系数图。矩阵行和列都通过主成分分析法进行了重新

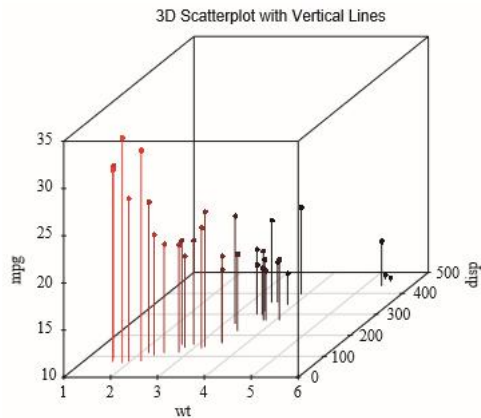


图 11-12 添加了垂直线和阴影的三维散点图



R的获取和安装

下载R : www.r-project.org

The screenshot shows the R Project website. A red arrow points to the 'CRAN' link under the 'Download' section. The website content includes the R logo, navigation links, and sections for 'Getting Started', 'News', and 'R Foundation'.

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

- **The R Journal Volume 9/1** is available.
- **R version 3.4.1 (Single Candle)** has been released on Friday 2017-06-30.
- **R version 3.3.3 (Another Canoe)** has been released on Monday 2017-03-06.
- **The R Journal Volume 8/2** is available.
- **useR! 2017** (July 4 - 7 in Brussels) has opened registration and more at <http://user2017.brussels/>
- Tomas Kalibera has joined the R core team.
- The R Foundation welcomes five new ordinary members: Jennifer Bryan, Dianne Cook, Julie Josse, Tomas Kalibera, and Balasubramanian Narasimhan.
- **The R Journal Volume 8/1** is available.
- The **useR! 2017** conference will take place in Brussels, July 4 - 7, 2017.

Help With R

Download

CRAN

R Project

About R
Logo
Contributors
What's New?
Reporting Bugs
Development Site
Conferences
Search

R Foundation

Foundation
Board
Members
Donors
Donate



The R Project for Statistical Computing

[Home]

Download

CRAN

R Project

About R

Logo

Contributors

What's New?

Reporting Bugs

Development Site

Conferences

Search

R Foundation

Foundation

Board

Members

Donors

Donate

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 3.5.0 (Joy in Playing)** has been released on 2018-04-23.
- **R version 3.4.4 (Someone to Lean On)** has been released on 2018-03-15.
- **useR! 2018** (July 10 - 13 in Brisbane) is open for registration at <https://user2018.r-project.org>
- **The R Journal Volume 9/2** is available.
- **useR! 2017** took place July 4 - 7 in Brussels <https://user2017.brussels>
- The **R Logo** is available for download in high-resolution PNG or SVG formats.

☐ 文件名

☐ R x64 3.4.3.lnk

☐ R-3.4.3-win.exe

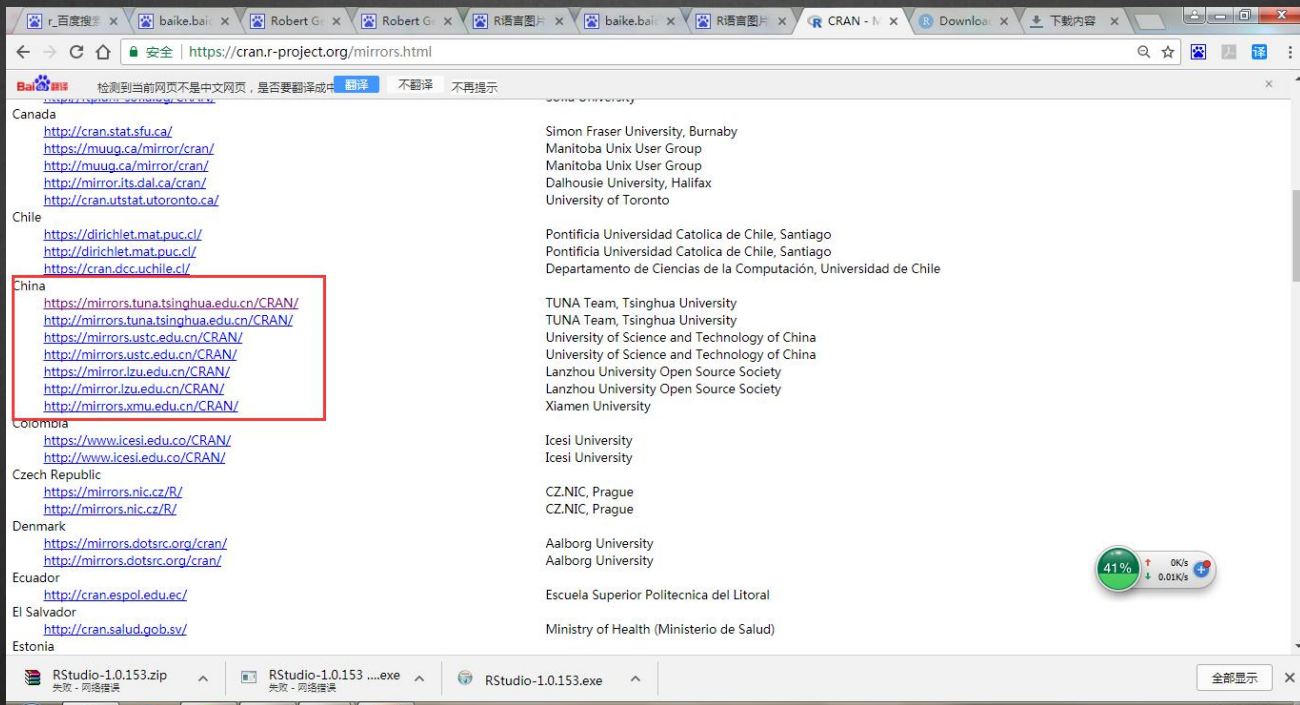
☐ R i386 3.4.3.lnk

☐ RStudio-1.1.383.exe



R的获取和安装

点击CRAN，选择离自己最近的镜像，点击下载





R的获取和安装

选择自己要下载的版本

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.

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 (Friday 2017-06-30, Single Candle) [R-3.4.1.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 [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](#)

Task Views Search

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

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

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

RStudio-1.0.153.zip
失败 - 网络错误

RStudio-1.0.153exe
失败 - 网络错误

RStudio-1.0.153.exe

全部显示



R的获取和安装

选择base

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.

Please do not submit binaries to CRAN. Package developers might want to contact Duncan Murdoch or 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](#)

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

41% 0.06K/s OK/s

https://mirrors.tuna.tsinghua.edu.cn/CRAN/bin/windows/base/rw-FAQ.html

RStudio-1.0.153.zip
RStudio-1.0.153exe
RStudio-1.0.153.exe



R的获取和安装

🔗 下载后得到R-3.4.4-win.exe，直接安装即可（现在是3.5.0）。

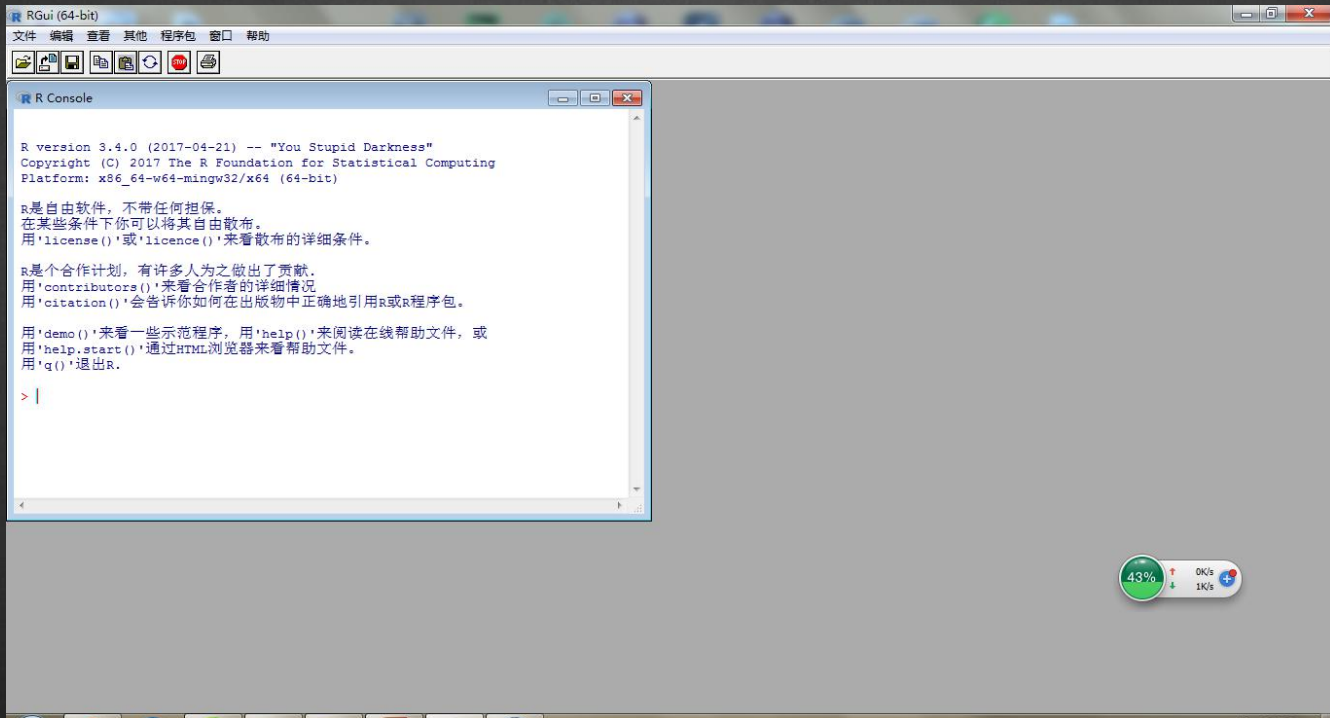
🔗 关于R说明：

▣ 上面下载的是R的运行环境，要在电脑上跑R，必须先安装R



1.3启动R

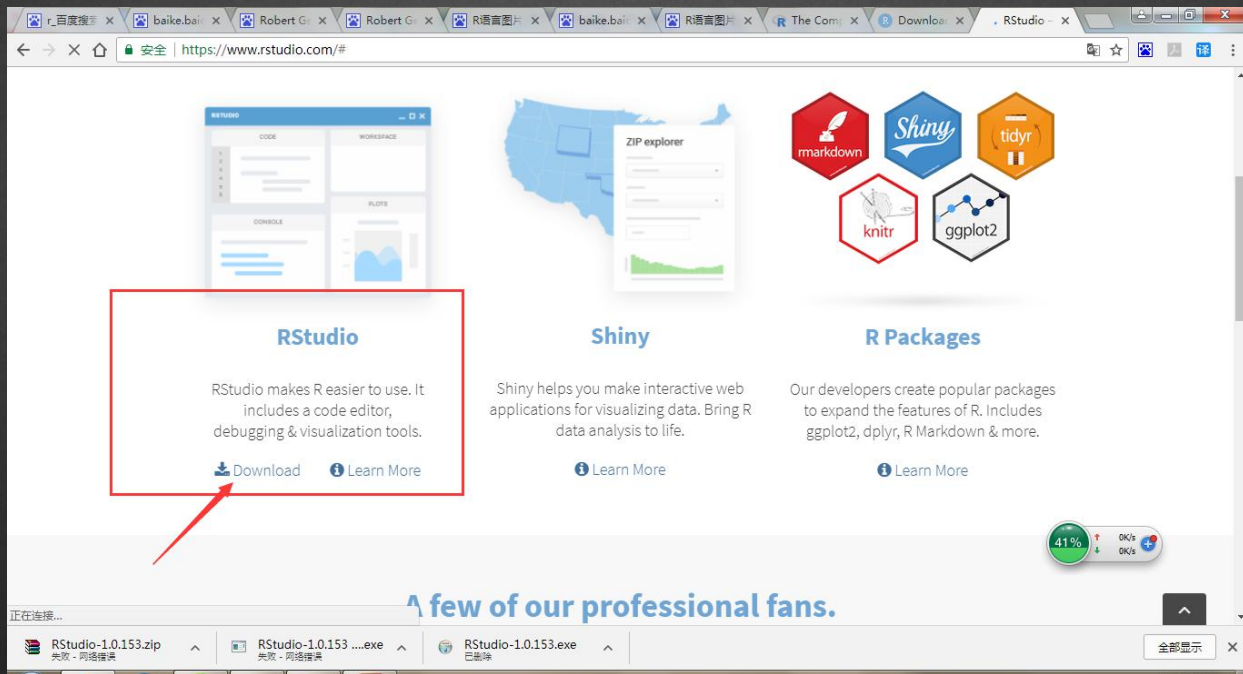
启动R你会看到：





下载安装RStudio

Rstudio是R的整合窗口环境，如果你想要使用这个窗口，可以到<https://www.rstudio.com/>进行下载。





下载安装RStudio

点击download，找到如下界面：

The screenshot shows the RStudio download page at <https://www.rstudio.com/products/rstudio/download/>. A red box highlights the AGPL license option, which is free. A red arrow points to the 'DOWNLOAD' button for this option.

License	Commercial	AGPL	Commercial	Commercial
Pricing	\$995/yr	FREE	\$9,995/yr	\$29,995/yr
Download	BUY NOW	DOWNLOAD	DOWNLOAD	CONTACT SALES
Learn More	Learn More	Learn More	Learn More	Learn More about RStudio Connect

Have Questions?

正在连接...

RStudio-1.0.153.zip
失败 - 网络错误

RStudio-1.0.153exe
失败 - 网络错误

RStudio-1.0.153.exe
已取消

全部显示



下载安装RStudio

选择你要下载的对的版本：

RStudio Desktop 1.0.153 — Release Notes

RStudio requires R 2.11.1+. If you don't already have R, download it here.

Installers for Supported Platforms

Installers	Size	Date	MD5
RStudio 1.0.153 - Windows Vista/7/8/10	81.9 MB	2017-07-20	b3b4bbce82865ab105c21cb70b17271b3
RStudio 1.0.153 - Mac OS X 10.6+ (64-bit)	71.2 MB	2017-07-20	8773610566b74ec3e1a88b2fcb10c8b5
RStudio 1.0.153 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	85.5 MB	2017-07-20	981be44f91fc07e5f69f52330da32659
RStudio 1.0.153 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	91.7 MB	2017-07-20	2d0769bea2bf6041511d6901a1cf69c3
RStudio 1.0.153 - Ubuntu 16.04+/Debian 9+ (64-bit)	61.9 MB	2017-07-20	d584cbab01041777a15d62cbef69a976
RStudio 1.0.153 - Fedora 19+/Red Hat 7+/openSUSE 13.1+ (32-bit)	84.7 MB	2017-07-20	8dfee96059b05a063c49b705eca0ceb4
RStudio 1.0.153 - Fedora 19+/Red Hat 7+/openSUSE 13.1+ (64-bit)	85.7 MB	2017-07-20	16c2c8334f961c65d9bfaf5b813ad7e7

Zip/Tarballs

Zip/tar archives	Size	Date	MD5
RStudio 1.0.153 - Windows Vista/7/8/10	117.6 MB	2017-07-20	024b5714fa0ef337fe0c06f5e2894cbcb
RStudio 1.0.153 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	86.2 MB	2017-07-20	f8e0ffa7ec62665524f9e2477facd346
RStudio 1.0.153 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	92.7 MB	2017-07-20	2077c181311d1aad6fb8d435f8f1f45f
RStudio 1.0.153 - Fedora 19+/Red Hat 7+/openSUSE 13.1+ (32-bit)	85.4 MB	2017-07-20	92e1a22d14952273ec389e5a55b614f
RStudio 1.0.153 - Fedora 19+/Red Hat 7+/openSUSE 13.1+ (64-bit)	86.6 MB	2017-07-20	0b71c5a7fc53c84b3fe67242240b3531

Source Code

Taskbar: RStudio-1.0.153.zip (失败 - 网络错误), RStudio-1.0.153exe (失败 - 网络错误), RStudio-1.0.153.exe (已安装)



下载安装RStudio

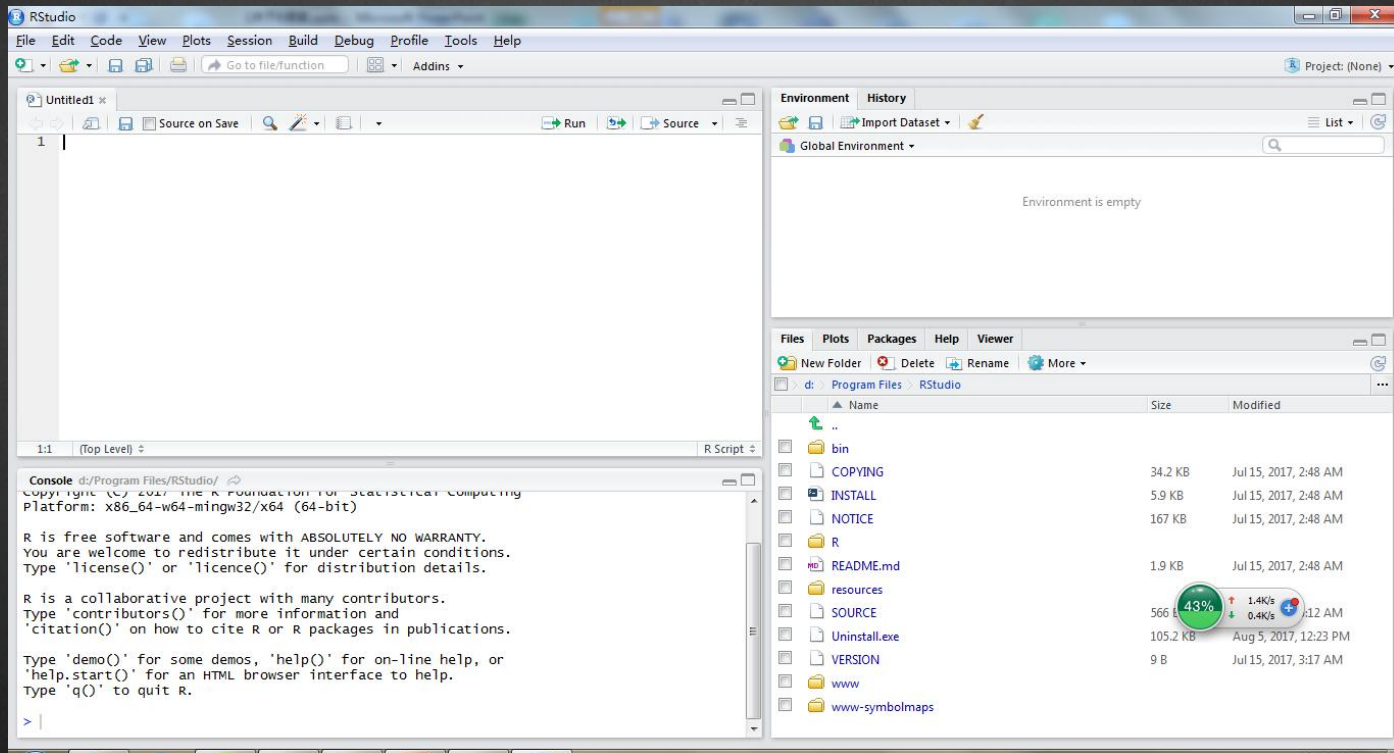
🔗 下载后得到RStudio-1.1.383.exe，点击进行安装





启动RStudio

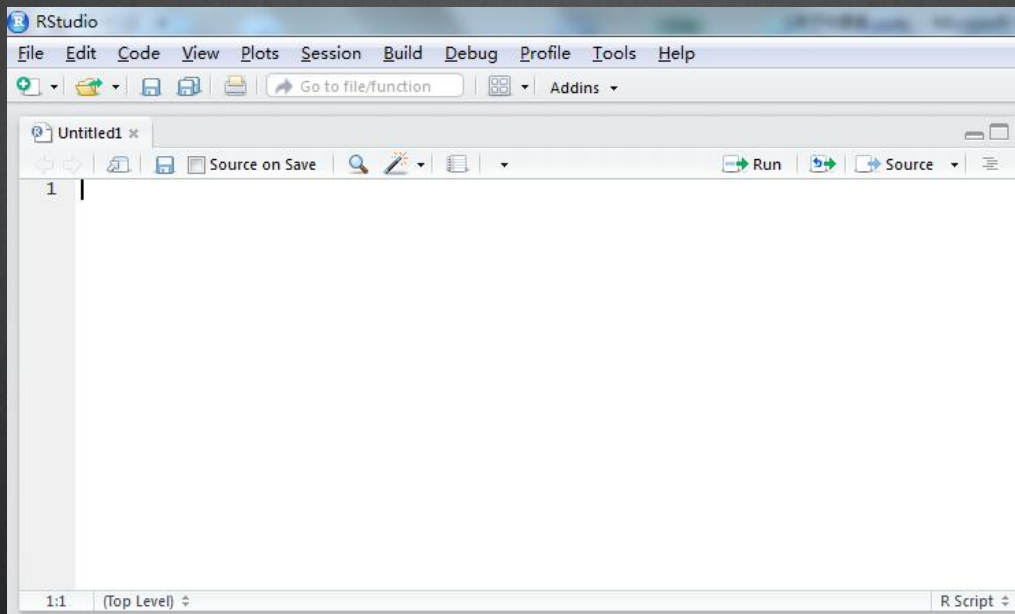
启动RStudio你会看到：





认识Rstudio环境

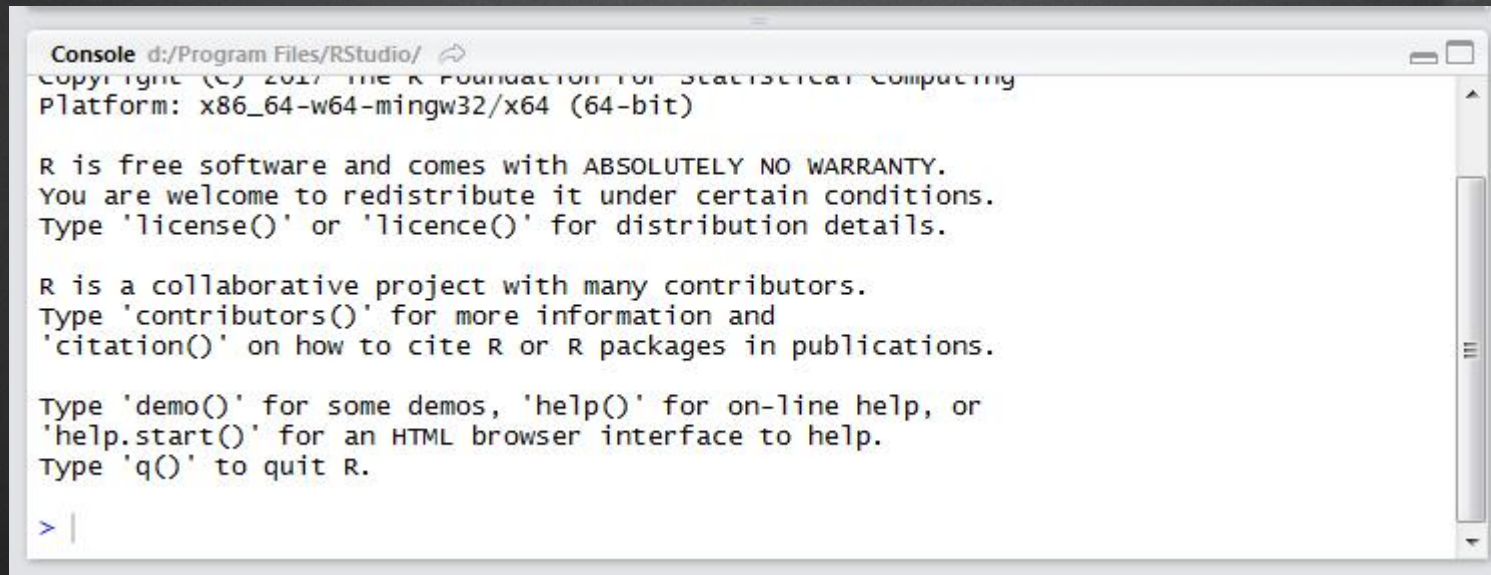
- Source Editor: Source Editor区域位于Rstudio窗口的左上角，这是R语言的程序代码的编辑区域，在这里进行R语言代码的编写，存储，最后在运行。





认识Rstudio环境

- ❧ Console：Console区域位于Rstudio的左下角。支持直译器的功能，也可在此窗口直接输入命令，同时获得执行结果。



```
Console d:/Program Files/RStudio/
Copyright (C) 2017 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.

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.

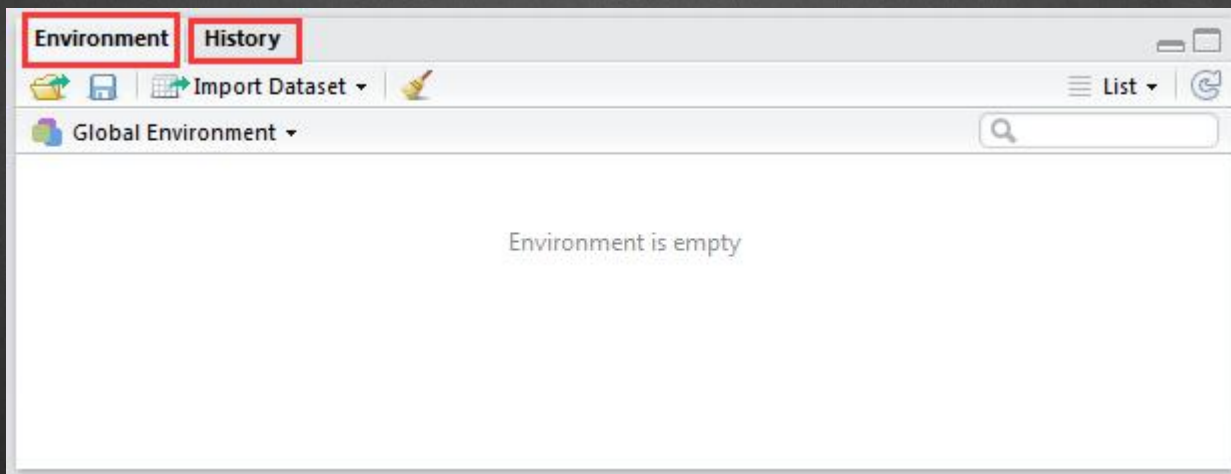
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```



认识Rstudio环境

- Workspace：位于Rstudio窗口的右上角，如果选择Environment标签，那么次区域会记录在Console输入的所有指令的相关对象的变量名称和值。如果选择的是History标签，则可以在此看到Console窗口所有执行指令的记录。

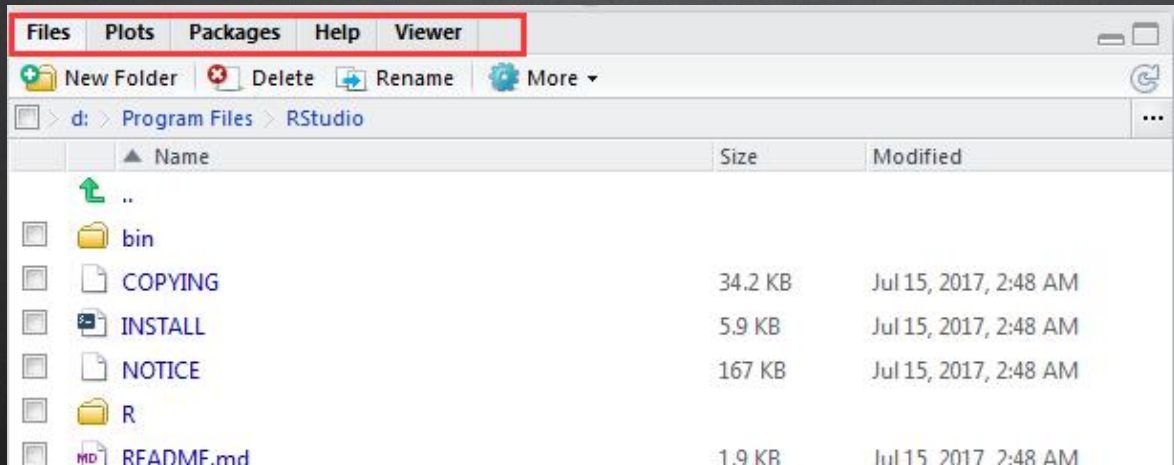




认识Rstudio环境

Files、Plots、Packages、Help和Viewer：该区域位于右下角。

- Files：在此可以查看各文件夹的内容；
- Plots：在此可以呈现图表；
- Packages：在此可以看到已经安装的R的扩充套件；
- Help：在此可以浏览辅助说明文件内容；





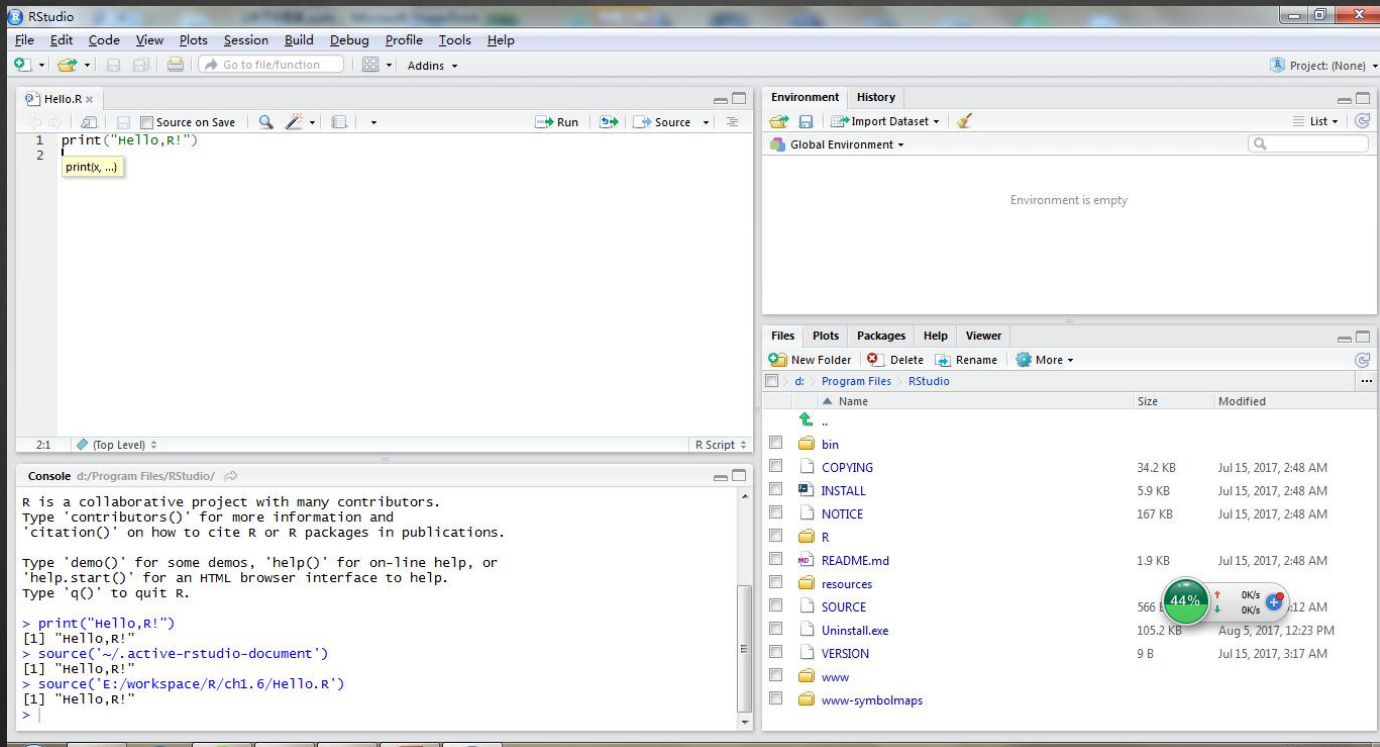
R的使用

- ❧ 可以直接在Rstudio中输入指令，也可以直接在R的GUI命令Console中进行输入，不过日后的的编码我们统一在Rstudio中进行。



R的使用

- 我们也可以使用Source Editor编辑程序，然后再执行
- 首先在编辑器中输入指令，
- 然后进行代码的保存，选择适当的文件夹，输入适当的文件名。R语言默认的文件扩展名是R。

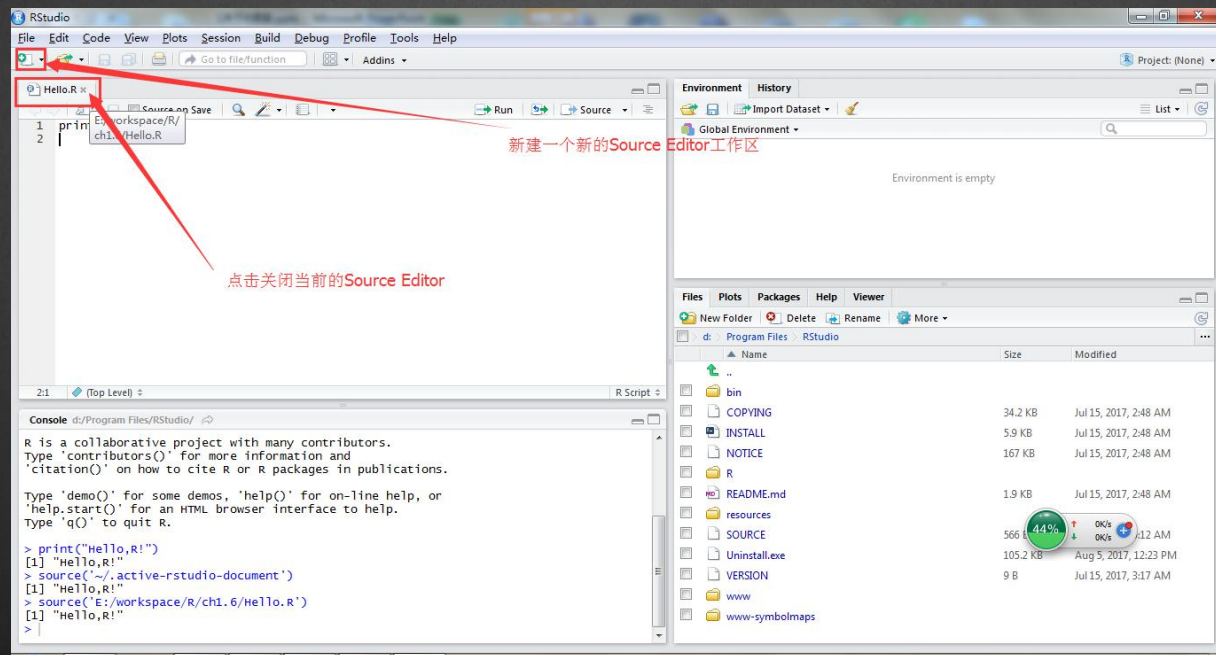




R的使用



在编辑完毕之后，你可以关闭当前的Source Editor，同时可以新建一个新的Source Editor，只需要点击Rstudio左上角的按钮即可新建。





R的使用

- R语言是一种面向对象的语言，在程序中输入一下语句：

`x=5`

- 对于上述的x可以称为变量，也可以直接以对象称之。
- 注意：在R语言中，可以直接使用“=”，但是更多的R语言程序设计师会使用“<-”，与“=”一样，将变量x设定为5。
- 在以后的开发中也会见到这样的符号“->”例如“5->x”，一般没有人这样去写，也不建议大家这样写程序。

```
1 x=5
2 x<-5
3 5->x
4
```

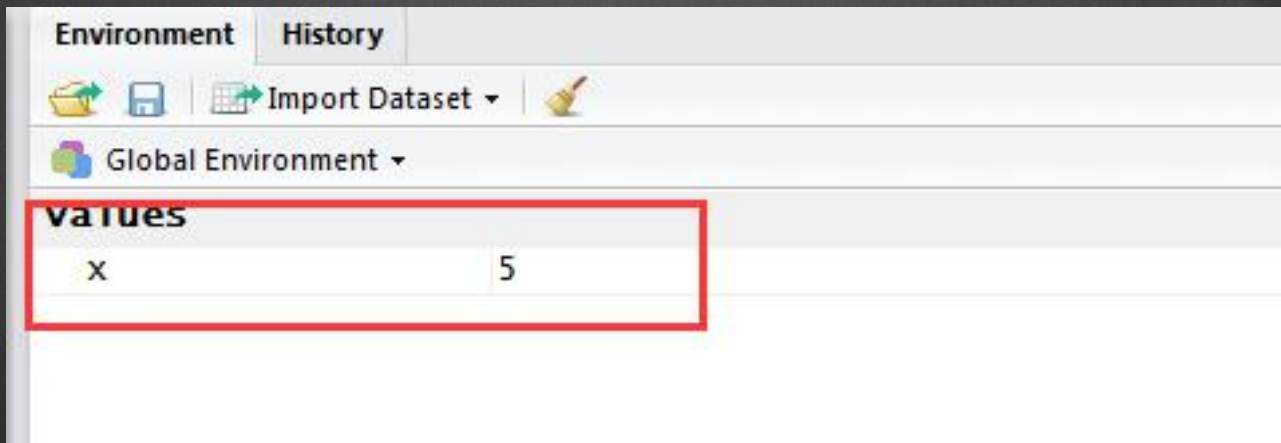
建议的书写方式

不建议



R的使用

- 在Workspace窗口，点击“Environment”标签，则可以看到至今所使用的对象变量及此对象变量的值。





R语言的注释

- 程序注释的主要功能是提高程序的可读性。程序加上注释可以方便他人阅读程序。
- 在R中可以通过“#”后面在文字的方式增加注释。

The screenshot shows the RStudio application window. The menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar. The main editor window, titled 'Untitled1*', contains the following R code:

```
1 x<-5#对象x的值为5
2 y<-10
3
```



❧ 示例：

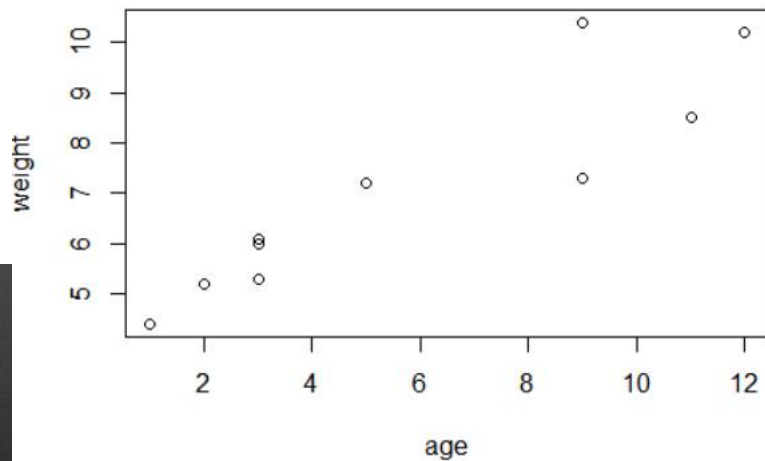
表1-1 10名婴儿的月龄和体重			
年龄（月）	体重（kg）	年龄（月）	体重（kg）
01	4.4	09	7.3
03	5.3	03	6.0
05	7.2	09	10.4
02	5.2	12	10.2
11	8.5	03	6.1

以上为虚构数据。



新手上路

```
> age <- c(1,3,5,2,11,9,3,9,12,3)
> age
[1] 1 3 5 2 11 9 3 9 12 3
> weight <- c(4.4,5.3,7.2,5.2,8.5,7.3,6.0,10.4,10.2,6.1)
> mean(weight)
[1] 7.06
> sd(weight)
[1] 2.077498
> cor(age,weight)
[1] 0.9075655
> plot(age,weight)
```





❧ R中的帮助函数

函 数	功 能
<code>help.start()</code>	打开帮助文档首页
<code>help("foo")</code> 或 <code>?foo</code>	查看函数 <code>foo</code> 的帮助（引号可以省略）
<code>help.search("foo")</code> 或 <code>??foo</code>	以 <code>foo</code> 为关键词搜索本地帮助文档
<code>example("foo")</code>	函数 <code>foo</code> 的使用示例（引号可以省略）
<code>RSiteSearch("foo")</code>	以 <code>foo</code> 为关键词搜索在线文档和邮件列表存档
<code>apropos("foo", mode="function")</code>	列出名称中含有 <code>foo</code> 的所有可用函数
<code>data()</code>	列出当前已加载包中所含的所有可用示例数据集
<code>vignette()</code>	列出当前已安装包中所有可用的vignette文档
<code>vignette("foo")</code>	为主题 <code>foo</code> 显示指定的vignette文档



用于管理R工作空间的函数

函 数	功 能
<code>getwd()</code>	显示当前的工作目录
<code>setwd("mydirectory")</code>	修改当前的工作目录为mydirectory
<code>ls()</code>	列出当前工作空间中的对象
<code>rm(objectlist)</code>	移除（删除）一个或多个对象
<code>help(options)</code>	显示可用选项的说明
<code>options()</code>	显示或设置当前选项
<code>history(#)</code>	显示最近使用过的#个命令（默认值为25）
<code>savehistory("myfile")</code>	保存命令历史到文件myfile中（默认值为.Rhistory）
<code>loadhistory("myfile")</code>	载入一个命令历史文件（默认值为.Rhistory）
<code>save.image("myfile")</code>	保存工作空间到文件myfile中（默认值为.RData）
<code>save(objectlist, file="myfile")</code>	保存指定对象到一个文件中
<code>load("myfile")</code>	读取一个工作空间到当前会话中（默认值为.RData）
<code>q()</code>	退出R。将会询问你是否保存工作空间



- ❧ 包是R函数、数据、预编译代码以一种定义完善的格式组成的集合
- ❧ 下载地址：<http://cran.r-project.org/web/packages>
- ❧ 包的安装：第一次安装一个包，使用命令`install.packages("包名")`即可
- ❧ 一个包仅需安装一次,但和其他软件类似，包经常被其作者更新
- ❧ 载入包`library(包名)`，包只需载入一次。



- ❧ 命令`help(package="package_name")`可以输出某 个包的简短描述以及包中的函数名称和数据集名称的列表
- ❧ 使用函数`help()`可以查看其中任意 函数或数据集的更多细节



示例

- ❧ (1) 打开帮助文档首页，并查阅其中的“Introduction to R”。
(2) 安装vcd包。(3) 列出此包中可用的函数和数据集。(4) 载入这个包并阅读数据集Arthritis的描述。(5) 显示数据集Arthritis的内容。(6) #运行数据集Arthritis自带的示例。(7) 退出。

```
1 help.start()
2 install.packages("vcd")
3 help(package="vcd")
4 library(vcd)
5 help(Arthritis)
6 Arthritis
7 example(Arthritis)
8 q()
```



小结

- 🔗 R的安装
- 🔗 熟悉R语言
- 🔗 运行R程序





Thankyou !





输入和输出

输入

函数source("filename")可在当前会话中执行一个脚本

输入

函数sink("filename")将输出重定向到文件filename

图形输出

函 数	输 出
pdf("filename.pdf")	PDF文件
win.metafile("filename.wmf")	Windows图元文件
png("filename.png")	PBG文件
jpeg("filename.jpg")	JPEG文件
bmp("filename.bmp")	BMP文件
postscript("filename.ps")	PostScript文件



输入和输出

❧ 示例

```
> source("script1.R")  
# 将工作空间路径中的脚本文件script1.R运行结果显示在屏幕上。  
> sink("myoutput", append=TRUE, split=TRUE) # 保存文本运行结果  
> pdf("mygraphs.pdf") # 保存图片至pdf  
> source("script2.R") # 调用脚本script2.R  
> sink()  
> dev.off() # 输出图片  
> source("script3.R") # 这次调用内容将仅显示在屏幕上
```

