



第1章 引言



内容导航

CONTENTS

1.1 ● R的起源与发展

1.2 ● 安装与运行R系统

1.3 ● 安装与使用包

1.4 ● 工作空间管理

1.5 ● R语言的集成开发环境RStudio

1.6 ● 使用帮助系统

- R语言是一种开源的脚本语言，在数据分析与预测，以及数据可视化等方面享有良好的声誉。
- 1993年，R的最初版本发布给研究人员使用，解决复杂数据统计分析任务，并用多样化的图形来展示结果。
- 据说，R的名字来源于它当时的两名开发者，新西兰奥克兰大学的Ross Ihaka和Robert Gentleman。
- R是对S语言的继承与发展。R系统是开源、免费的，由R开发核心小组负责其开发及维护。



01 适用于统计计算和机器学习

02 简单易学，具有高度的灵活性

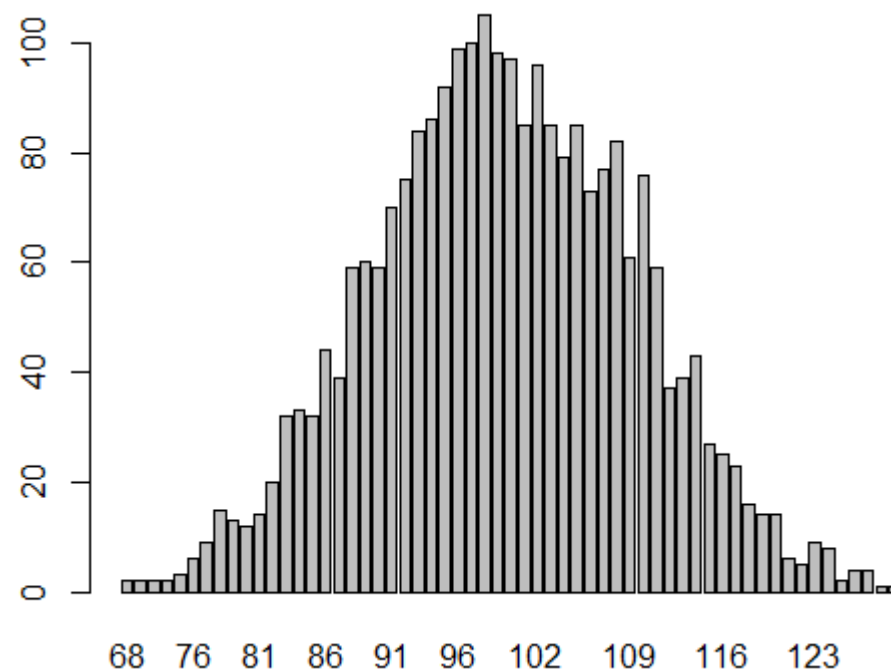
03 支持高质量的图形和强大的
可视化功能

04 代码形式简洁

```
> x <- 1:20                #把x赋值为一个从1到20的整型数向量
> y <- x + rnorm (20, 0, 2) #在x上添加均值为0、标准差为2的正态分布随机数，然后赋给y
> fit <- lm (y ~ x)         #完成y~x的线性回归，结果保存到变量fit
> summary (fit)            #概要显示线性回归的结果
```

- 在R语言里，符号 “<-” 的意义代表着赋值。
- 这个例子的前两行分别准备了两列数据—自变量x与因变量y。
- 第三行的函数lm即线性模型，根据提供的样本数据进行线性回归计算
- 第四行summary()显示模型的计算结果。

```
> n <- floor(rnorm(2500,100,10)) #产生2500个正态分布随机数并取整  
> t<- table(n)    #统计n中所有整数出现的次数  
> barplot(t)      #画出柱状图
```





内容导航

CONTENTS

1.1 ● R的起源与发展

1.2 ● 安装与运行R系统

1.3 ● 安装与使用包

1.4 ● 工作空间管理

1.5 ● R语言的集成开发环境RStudio

1.6 ● 使用帮助系统

- R可以在CRAN(Comprehensive R Archive Network) 上免费下载。
- <https://mirrors.tuna.tsinghua.edu.cn/CRAN/>
- 最新版本 R-4.3.1, 2023-06-16发布; 实验室安装版本4.2.1 (差异不大)
- Linux、Mac OS X和Windows都有相应编译好的二进制版本。



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

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

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

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

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 \(Debian, Fedora/Redhat, Ubuntu\)](#)
- [Download R for macOS](#)
- [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!

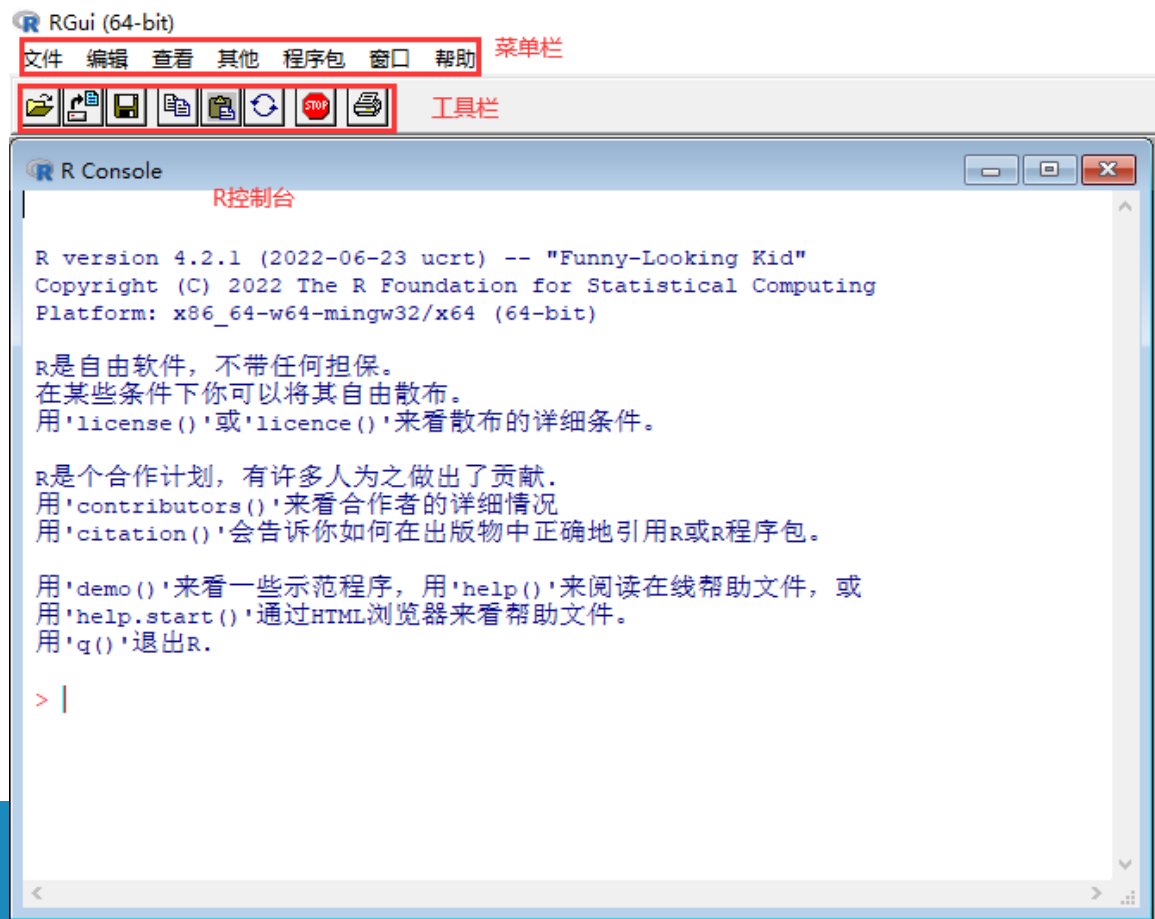
- The latest release (2022-06-23, Funny-Looking Kid) [R-4.2.1.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [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

- 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.

1. R的图形用户界面

在R的GUI窗口里，有菜单栏、工具栏和R的控制台。





内容导航

CONTENTS

- 1.1 ● R的起源与发展
- 1.2 ● 安装与运行R系统
- 1.3 ● 安装与使用包
- 1.4 ● 工作空间管理
- 1.5 ● R语言的集成开发环境RStudio
- 1.6 ● 使用帮助系统

- 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。
- R中的包存储在计算机上名为library的目录下，使用函数libPaths ()可以查看该文件夹在计算机中的具体路径，函数library ()和search ()则可分别显示已安装和加载的包列表。
- R已经预装了一组标准的包，其他包则可以通过下载安装来使用。
- 目前CRAN上提供了上万个可使用的包。

最受欢迎的5个包分别是：

- dplyr：一种数据操作的语法
- devtools：一组用于包开发的工具
- foreign：读取用其他软件如Minitab、S、SAS、SPSS和Stata等存储的数据
- cluster：聚类分析方法
- ggplot2：R语言图形工具

R

◆安装包：在R环境中执行 > `install.packages("xxx")`

◆载入包：在R环境中执行 > `library("xxxx")`

VS

Python

◆安装包：终端命令行中执行 `pip install xxx`

◆载入包：在开发环境中执行 `import xxx`



内容导航

CONTENTS

- 1.1 ● R的起源与发展
- 1.2 ● 安装与运行R系统
- 1.3 ● 安装与使用包
- 1.4 ● 工作空间管理
- 1.5 ● R语言的集成开发环境RStudio
- 1.6 ● 使用帮助系统

- 启动R环境之后，交互地运行一些语句，直到退出R，这就构成了一个R会话；
- 会话中创建的所有的对象都被临时保存在**全局环境.GlobalEnv**，称为**工作空间**；
- 当使用`q()`函数或直接关闭RGui窗口退出R环境时，就结束了当前会话。系统提示是否保存工作空间，如果选择“是”，当前环境中的所有对象都会被写入到一个叫做“.RData”的文件中，使用过的命令行历史则会保存在名为“.Rhistory”的文件中；
- 这些文件会被默认地保存在当前R会话的工作目录中。下次从此目录启动R会话时，软件会自动将对象、命令历史载入到工作空间。

- R提供了一些管理工作空间的函数
- 利用这些函数可以方便地完成目录、对象与包的管理

函 数	说 明
getwd ()	显示当前工作目录
setwd ()	修改当前工作目录
ls ()	显示当前工作空间中的所有对象
str ()	显示对象的结构
ls.str ()	显示对象中每一个变量的结构
exists ()	当前工作空间内是否存在某个对象
rm ()	删除一个或多个对象
q ()	退出R。在这之前会询问是否保存工作空间
install.packages ()	安装包
library ()	载入包

直接调用相应的管理函数

```
> getwd ()           #查看当前工作目录
[1] "C:/Users/Temp"
> ls ()              #查看工作空间所有对象
[1] "all_data"          "classifier"      "count"
"error"
[5] "i"                 "index"
"iris.Test"          "iris.Training"
[9] "lm_fit"            "positions"
"predictions"        "testing"
[13] "training"          "us.cities"       "x"
"x1"
[17] "x2"                "x3"              "y"
```

如果希望删除内存中的对象，可以使用rm ()函数。嵌套执行rm (list = ls ()) 则会清除当前工作空间中的全部对象

```
> rm(i,x1,x2,x3)      #删除I,x1,x2,x3等
> rm (list = ls ())
> ls ()
character(0)
```



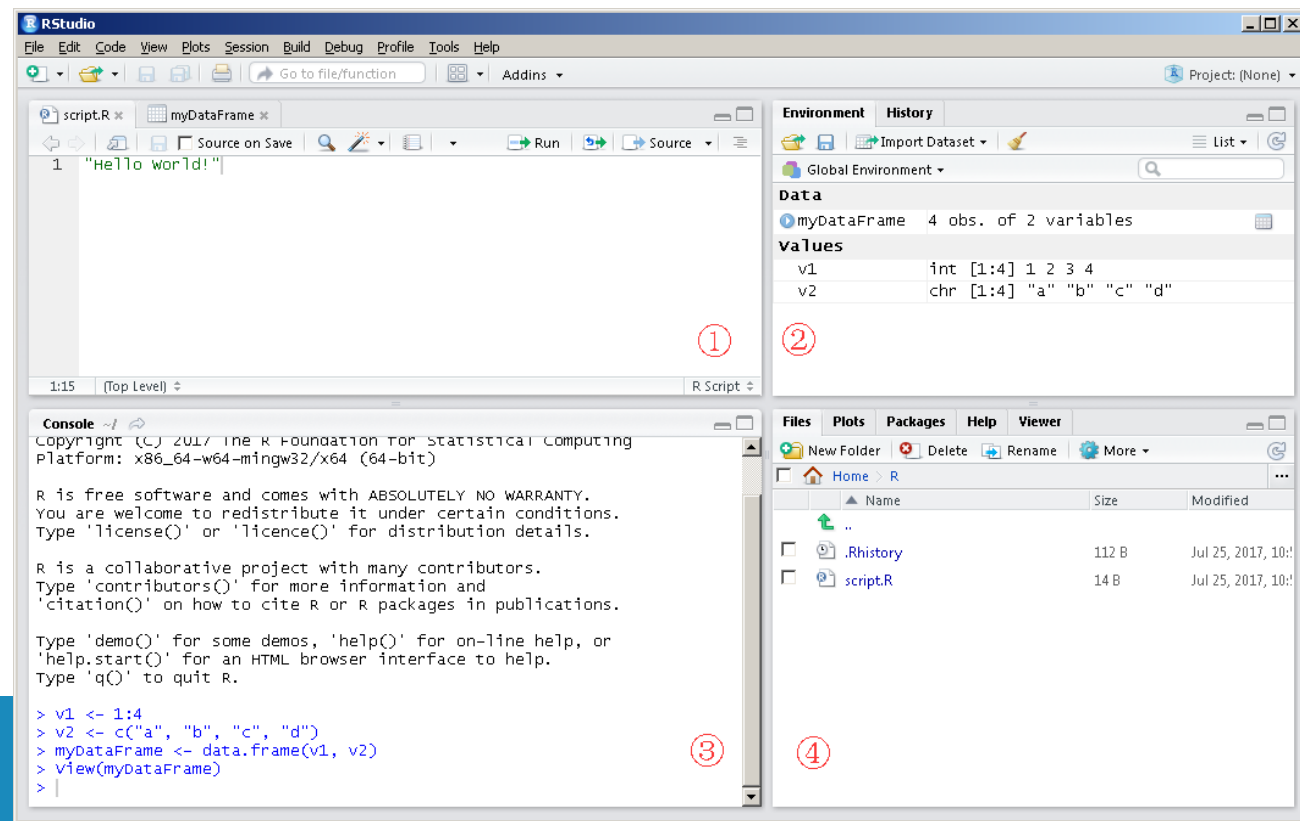
内容导航

CONTENTS

- 1.1 ● R的起源与发展
- 1.2 ● 安装与运行R系统
- 1.3 ● 安装与使用包
- 1.4 ● 工作空间管理
- 1.5 ● R语言的集成开发环境RStudio
- 1.6 ● 使用帮助系统

- 集成开发环境 (Integrated Development Environment, IDE) , 是一个集代码编辑器、调试器、图形用户界面等一系列工具为一体的应用程序
- 许多优秀的IDE都将开发所需的主要功能集成在一个桌面环境中, 大大提高了生产效率
- 使用IDE的优点包括:
 1. 节省时间和精力
 2. 统一代码标准
 3. 完善的调试环境

- Rstudio是一个免费的R语言IDE
- 它具有**语法高亮、命令补全、对象浏览、语法错误检查以及断点调试**等功能，可以大大提高开发效率
- Rstudio可以在<https://www.rstudio.com/products/rstudio/download/#download> 获取，选择对应的系统平台进行下载安装
- 默认情况下窗口被分成如下四个区域：
 - ① 代码编辑、数据预览窗口；
 - ② 工作空间、命令历史；
 - ③ 控制台；
 - ④ 文件浏览、绘图、包管理等。





内容导航

CONTENTS

- 1.1 ● R的起源与发展
- 1.2 ● 安装与运行R系统
- 1.3 ● 安装与使用包
- 1.4 ● 工作空间管理
- 1.5 ● R语言的集成开发环境RStudio
- 1.6 ● 使用帮助系统

- R的内置帮助系统提供了当前已安装包中所有的函数细节和使用示例
- 使用`help.start ()` 打开帮助HTML的帮助文档首页，包括了关于R语言的一整套学习和帮助材料
- R提供了不同的函数高效地帮助使用者查找帮助信息

函 数	说 明
<code>help.start ()</code>	打开帮助文档首页
<code>help (), ?</code>	查看
<code>help.search (), ??</code>	以关键字搜索的形式搜索本地帮助文档
<code>RSiteSearch ()</code>	以关键字搜索的形式搜索在线帮助文档
<code>apropos ()</code>	列出名称中含有某字符串的所有可用函数
<code>example ()</code>	查看某个函数的使用示例

- 在RStudio中，资源管理窗口的Help子窗口则可以直接通过函数的输入查看相关的帮助文档

