

R语言零基础入门

复旦大学大数据学院 周之烁

R的安装

Step 1: 登陆R语言官方网站
<https://www.r-project.org>, 点击
download R。

Step 2: 在弹出的镜像 (Mirrors)
页面上选择合适的镜像入口, 比
如选择China下的清华镜像。

Step 3: 根据自己电脑的操作系统
选择。

Step 4: (Windows) 选择base版
本进行下载并完成安装, 安装成
功之后可在开始菜单中找到;
(Mac) 选择最新版本的pkg文件
进行下载, 安装后可以在
Applications文件夹下找到。

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 please choose your preferred CRAN mirror.

download R.

1

China

- <https://mirrors.tuna.tsinghua.edu.cn/CRAN/>
- <https://mirrors.bfsu.edu.cn/CRAN/>
- <https://mirrors.ustc.edu.cn/CRAN/>
- <https://mirror-hk.koddos.net/CRAN/>
- <https://mirrors.e-ducation.cn/CRAN/>
- <https://mirror.lzu.edu.cn/CRAN/>
- <https://mirrors.nju.edu.cn/CRAN/>
- <https://mirrors.tongji.edu.cn/CRAN/>
- <https://mirrors.sjtug.sjtu.edu.cn/cran/>
- <https://mirrors.sustech.edu.cn/CRAN/>

2

TUNA Team, T
Beijing Foreign
University of S
KoDDoS in Ho
Elite Educatior
Lanzhou Unive
eScience Cent
Tongji Univers
Shanghai Jiao
Southern Univ

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

3

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
- contrib
- old.contrib
- Rtools

Binaries for base distribution. This is what you want to **install**.

Binaries of contributed CRAN packages (for R >= 2.13.x; m. Ligges). There is also information on **third party software** at Windows services and corresponding environment and ma

Binaries of contributed CRAN packages for outdated versio 2.13.x; managed by Uwe Ligges).

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

R-4.0.4.pkg (notarized and signed)

SHA1-
hash: 0b2b3bc846f6ebc72a8bc0b53e6e85d600d95deb
(ca. 85MB)

Latest release:

R 4.0.4 binary for macOS 10.13 (High Sierra) and higher, signed and notarized package. Contains R 4.0.4 framework, R.app GUI 1.74 in 64-bit for Intel Macs, Tcl/Tk 8.6.6 X11 libraries and Texinfo 6.7. The latter two components are optional and can be omitted when choosing "custom install", they are only needed if you want to use the `tc1tk` R package or build package documentation from sources.

Note: the use of X11 (including `tc1tk`) requires **XQuartz** to be installed since it is no longer part of OS X.

4-Windows

4-Mac

RStudio的安装

!必须先安装R, 再安装RStudio

Step 1: 打开RStudio官方网站
<https://www.rstudio.com/products/rstudio/download/>, 点击RStudio
Desktop下的Download。

Step 2: 网站通常会自动检测到我
们的操作系统, 直接下载即可。

Step 3: 安装时, 64位系统可全选,
并接受默认选项。

Step 4: (Windows) 安装成功之
后可在开始菜单中找到, 双击打
开; (Mac) 安装后可以在
Applications文件夹下找到并打开。

RStudio的安装

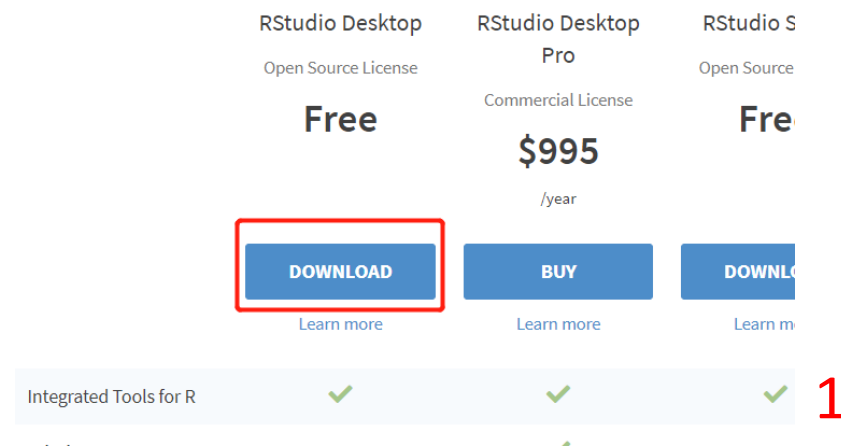
! 必须先安装R，再安装RStudio

Step 1: 打开RStudio官方网站
<https://www.rstudio.com/products/rstudio/download/>，点击RStudio Desktop下的Download。

Step 2: 网站通常会自动检测到我们的操作系统，直接下载即可。

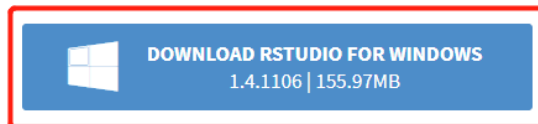
Step 3: 安装时，64位系统可全选，并接受默认选项。

Step 4: (Windows) 安装成功之后可在开始菜单中找到，双击打开；(Mac) 安装后可以在Applications文件夹下找到并打开。



RStudio Desktop 1.4.1106 - [Release Notes](#)

1. Install R. RStudio requires R 3.0.1+.
2. Download RStudio Desktop. Recommended for your system:



Requires Windows 10/8 (64-bit)

2-Windows

RStudio Desktop 1.4.1106 - [Release Notes](#)

1. Install R. RStudio requires R 3.0.1+.
2. Download RStudio Desktop. Recommended for your system:



Requires macOS 10.13+ (64-bit)

2-Mac

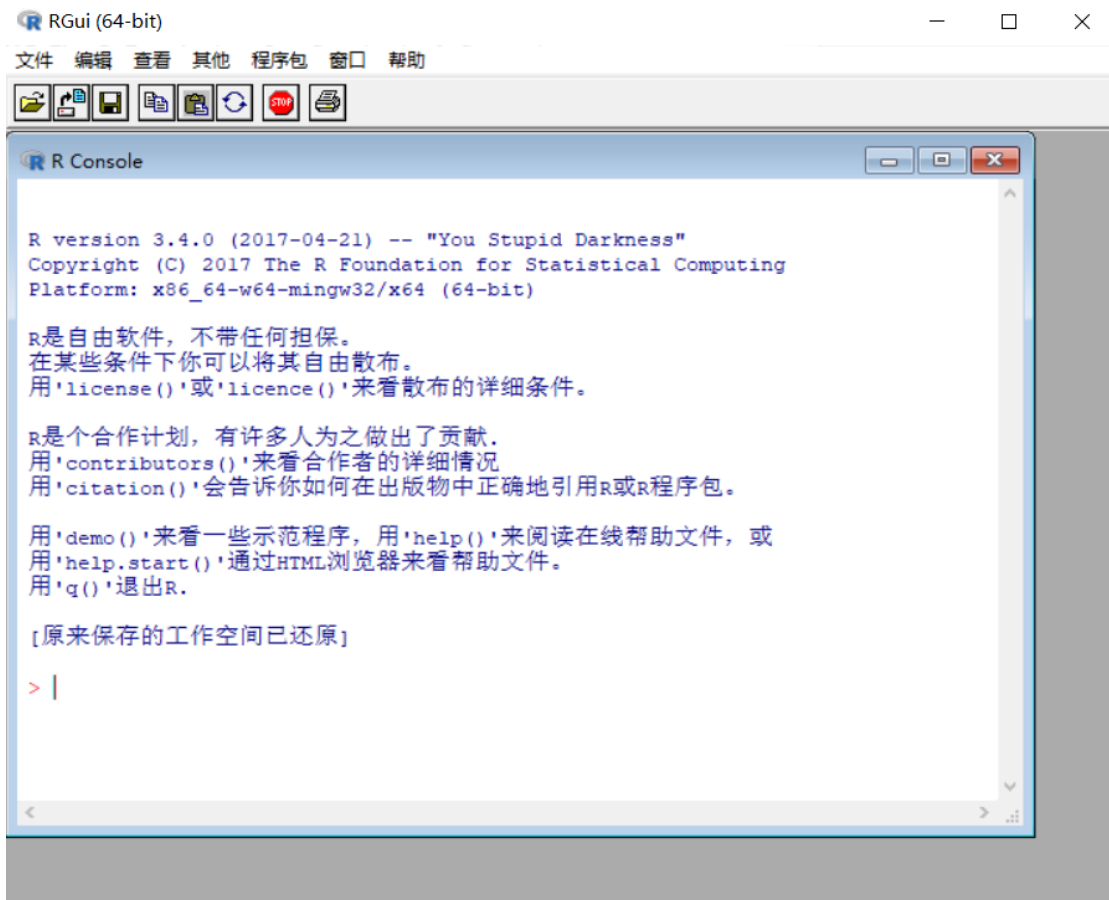
目录

1. R与R studio的安装与介绍
2. 工作路径与数据读入
3. R包的安装与载入
4. 赋值与变量
5. 数据类型与运算符
6. 向量与数据框
7. 上手陌生代码（以R作图为例）

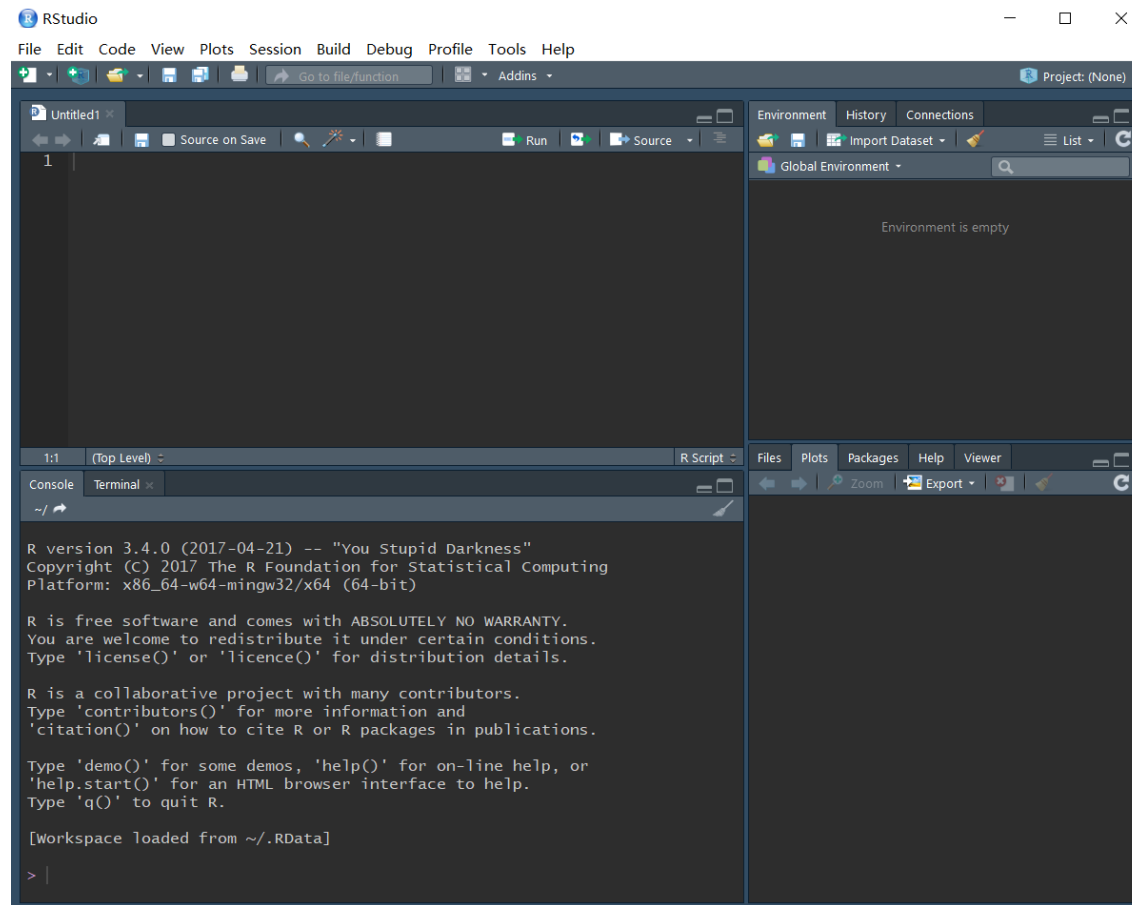


语言优点

1. 免费
2. 颜值高
3. 覆盖广
4. 社区支持强 `stack overflow`



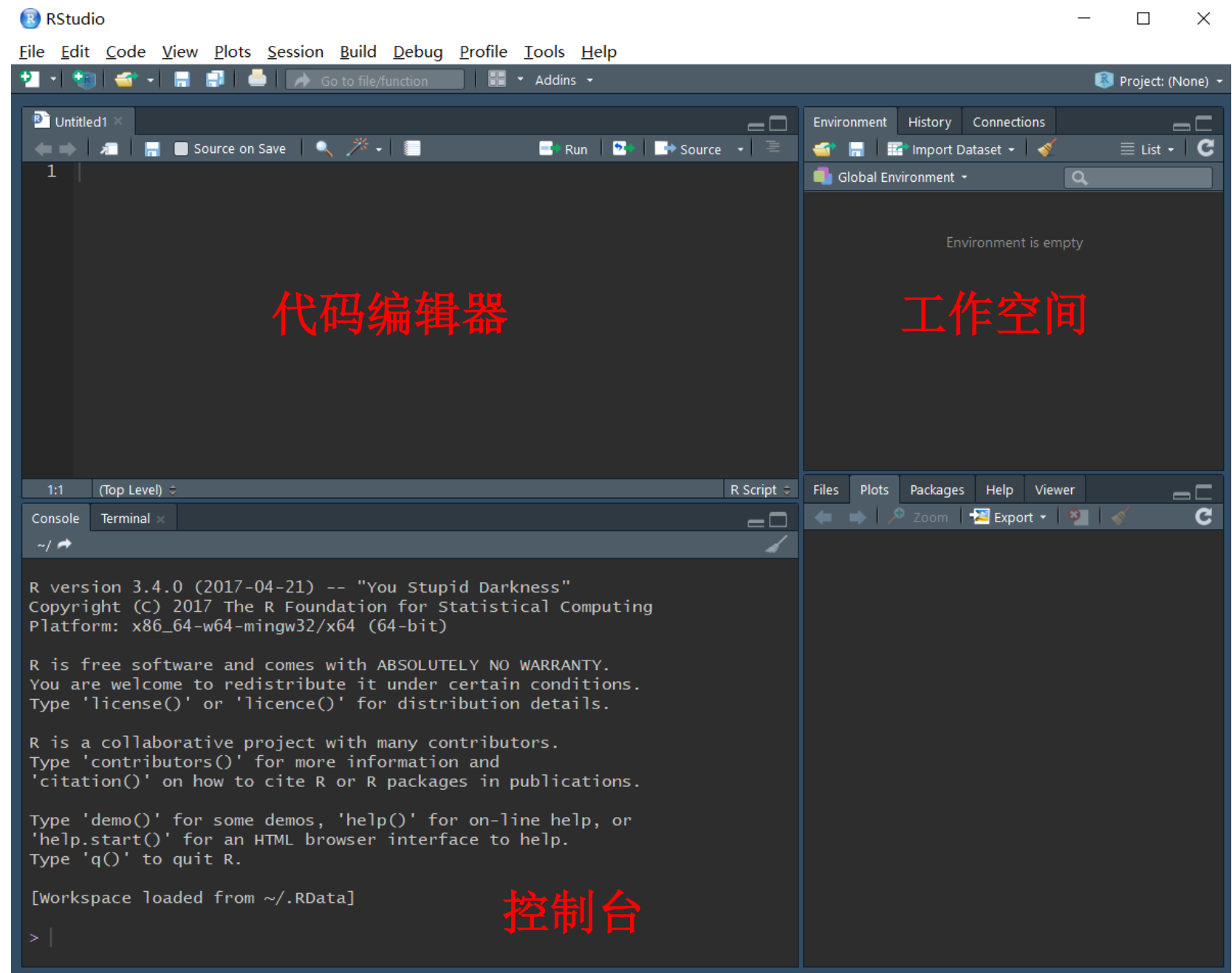
R



Rstudio

RStudio优点

1. 直观的界面使我们能够跟踪保存的对象，脚本和图形
2. 具有颜色编码语法等功能的文本编辑器，可帮助我们编写简洁的脚本
3. 自动完成功能可节省时间
4. 用于创建包含项目代码，注释和视觉效果文档的工具
5. 专用的Project文件夹可将所有内容保存在一个地方



工作路径

getwd(): 当前的工作路径

setwd("..."): 设置当前的工作路径（记得加双引号！）

只有明确当前的工作路径，才能保证文件的顺利读取 or
保存在自己想要的位置

路径的分隔符需用“/”，无论是Windows还是Mac:

```
setwd("E:\讲座\0324") # No  
setwd("E:/讲座/0324") # Yes  
setwd("E://讲座/0324") # No
```

```
> setwd("E:\讲座\0324")  
Error: '\? is an unrecognized escape in character string starting ""E:\?
```



表格数据读入

```
read.table(file_name, header = FALSE, sep = "")
```

`file_name` 表示文件名；`header`用来告诉计算机是否把数据的第一行识别为变量名（默认第一行不是变量名）；`sep`则用来指定文件中的分隔符。

```
read.csv(file_name, header=TRUE)
```

`file_name` 表示文件名；`header`用来告诉计算机是否把数据的第一行识别为变量名（默认第一行是变量名）；无需指定`sep`，因为`csv`文件的分隔符一定是逗号。

```
library(readxl)
```

```
read_excel(file_name, col_names=TRUE)
```

`file_name` 表示文件名；`col_names`用来告诉计算机是否把数据的第一行识别为变量名（默认第一行是变量名）

R包（R package）

1. 什么是R包？
 - 把R函数、数据、预编译代码以一种定义完善的格式组成在一起的集合
2. R在安装时会自带一系列默认包
 - 包括base, datasets, stats, methods, graphics等包
3. 如果需要装备更多、更厉害的拓展技能，就需要安装新包

R包 (R package)

两个概念:

安装——`install.packages("...")`——只需安装一次，显示在右下角Packages中——买回家

载入——`library(...)`——在不同代码文件中，一般需要重新载入——拿到桌子上

载入时发现包不存在，则需要安装

```
> install.packages("factoextra")
```

最后出现以下字样则安装成功，此时也可在右下角Packages中查看

```
The downloaded source packages are in  
  'C:\Users\lenovo\AppData\Local\Temp\Rtmpqa60QX\downloaded_packages'
```

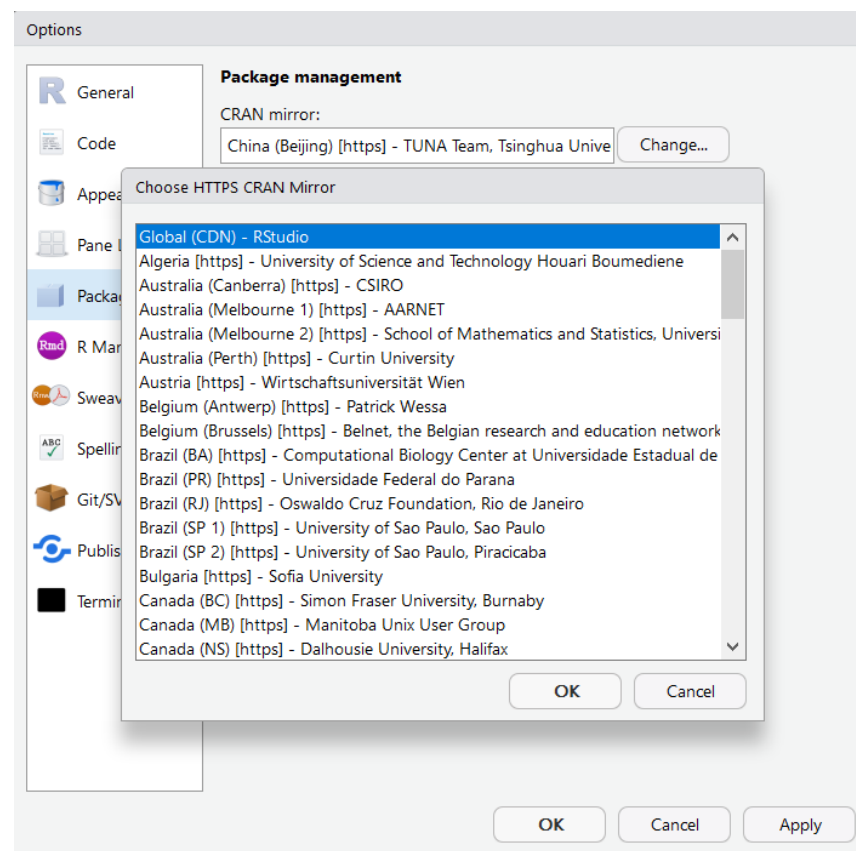
安装后记得重新载入!

更换R包的下载源

如果下载的很慢，大概率是由于此时RStudio采用的是默认的国外Global镜像。

切换为本地镜像：

1. 点击“Tools → Global Options”
2. 点击Packages选项卡
3. 更换为China开头的镜像，比如清华大学的镜像



第一阶段结束！

安装R与Rstudio

初识Rstudio

新建代码文件并保存

初步导入数据

学会问号一下+百度/谷歌一下

安装R包



初级调包侠

赋值运算符

`<-` 被称为赋值运算符，该运算符将值分配给变量。

Rstudio中，键入赋值运算符的键盘快捷键：

1. Mac OS X: 同时按Option和-
2. Windows : 同时按Alt和-

变量

变量命名规则：

1. R 语言的有效变量名称由字母，数字以及点号 . 或下划线 _ 组成。
2. 变量名称以字母或点开头（. 号开头后面不能跟着数字）

习惯性规则：

1. 下划线/点号分隔法
2. 驼峰法

变量名	是否正确	原因
var_name2.	正确	字符开头，并由字母、数字、下划线和点号组成
var_name%	错误	% 是非法字符
2var_name	错误	不能数字开头
.var_name, var.name	正确	可以 . 号开头，但是要注意 . 号开头后面不能跟着数字
.2var_name	错误	. 号开头后面不能跟着数字
_var_name	错误	不能以下划线 _ 开头

数据类型 & 对象类型

数据类型

1. 数字：123、1.23e2
2. 逻辑（布尔型）：TRUE、FALSE
3. 文本：可以用单引号或双引号包含

对象类型

1. 向量（vector）：c()函数
2. 列表（list）
3. 矩阵（matrix）
4. 数组（array）
5. 因子（factor）
6. 数据框（data.frame）

运算符

优先级	符号	含义
1	()	括号
2	^	乘方运算
3	%%	整除求余
	%/%	整除
4	*	乘法
	/	除法
5	+	加法
	-	减法

数学运算符

运算符	描述
>	判断第一个向量的每个元素是否大于第二个向量的相对对应元素。
<	判断第一个向量的每个元素是否小于第二个向量的相对对应元素。
==	判断第一个向量的每个元素是否等于第二个向量的相对对应元素。
!=	判断第一个向量的每个元素是否不等于第二个向量的相对对应元素。
>=	判断第一个向量的每个元素是否大于等于第二个向量的相对对应元素。
<=	判断第一个向量的每个元素是否小于等于第二个向量的相对对应元素。

关系运算符

运算符

运算符	描述
&	元素逻辑 与 运算符，将第一个向量的每个元素与第二个向量的相对应元素进行组合，如果两个元素都为 TRUE ，则结果为 TRUE ，否则为 FALSE 。
	元素逻辑 或 运算符，将第一个向量的每个元素与第二个向量的相对应元素进行组合，如果两个元素中有一个为 TRUE ，则结果为 TRUE ，如果都为 FALSE ，则返回 FALSE 。
!	逻辑 非 运算符，返回向量每个元素相反的逻辑值，如果元素为 TRUE 则返回 FALSE ，如果元素为 FALSE 则返回 TRUE 。

逻辑运算符

运算符	描述
:	冒号运算符，用于创建一系列数字的向量。
%in%	用于判断元素是否在向量里，返回布尔值，有的话返回 TRUE ，没有返回 FALSE 。

向量运算符

对象类型

对象类型

1. **向量 (vector)** : `c()`函数
2. 列表 (list)
3. 矩阵 (matrix)
4. 数组 (array)
5. 因子 (factor)
6. **数据框 (data.frame)**

向量相关函数

函数名	含义
sort	返回排序后的向量（默认从小到大）
rev	返回逆序向量
order	返回向量排序后的下标向量

函数名	含义
seq	生成有间隙的等差数列
rep	生成重复出现的数字序列

函数名	含义
sum	求和
mean	求平均值
median	求中位数
var	方差
sd	标准差
min	最小值
max	最大值
range	取值范围（二维向量，最大值和最小值）
length	向量长度

循环

```
for (value in vector) {  
    statements  
}
```

for循环

```
while (condition) {  
    statements  
}
```

while循环

数据框（data.frame）

head()：查看前6行

str()：查看数据框每列的变量类型

summary()：查看数据框的数据分布基础信息

取某行所有值：

1. 指明第几行

取某列所有值：

1. 用列名来提取（加双引号）
2. 用“\$”符号后接列名来提取（不加双引号）
3. 指明第几列

如何上手陌生的代码

1. 观察原代码对应的数据结构，思考手里的数据是否需要进行处理
2. 查询每个函数的具体作用
3. 通过增添/删除/修改代码中函数的参数，确认各个参数的实际作用

第二阶段结束！

赋值

变量

数据类型

运算符

向量

数据框

作图



中级调包侠

狗熊会: R公众号

Runoob: 工具型网站

End.