



R数据科学教程

Fattroy胖春

2022年1月

R语言介绍

【一】什么是R语言

- R语言是继承自S语言的一种实现，而S语言是由AT&T贝尔实验室开发的一种用来进行数据探索、统计分析、作图的解释型语言。因此，R语言继承了S语言在统计计算、数据分析和绘图等方面的优势，成为一门在统计领域尤其热门的编程语言。
- R语言是一个基于GNU框架的“Free Language”，它高度自由、完全免费，并且充分开源，你几乎可以不需要任何成本的使用R语言并且融入其高度繁荣的用户社区。
- R语言是一门跨平台的解释性语言，不需要编译成为可执行文件即可运行程序，你甚至可以在PlayStation上面运行R语言的程序（别问我是怎么知道的，我甚至连PlayStation都没有）。

【二】为什么选择R语言

- 相比较于其他统计软件（SAS，SPSS等）R首先是完全免费的，能够节省大量的用来订阅软件的成本（~~或者是用来寻找破解版软件的时间~~），更重要的是R是高度自由和可扩展的，你能够利用R来进行设计和开发，实现你所期待的功能，这也是R被称为一门编程语言而非一个统计软件的原因。
- 相对于其他程序设计语言（C，Java等）R是简单而专业的，R拥有非常繁荣的社区，尤其是在统计和数据分析相关的领域，更多时候你在调用一个方法的时候并不需要考虑其和计算机是如何交互的（~~但这并不能成为你不愿意学统计学基础的理由~~），尤其是在分析和可视化方面，一行代码就能实现的报表和绘图节省了大量的时间。另一方面，R语言的上手也相对更加容易，因为在你了解基本的语法之后，它的程序结构往往相当直白。
- 与Python相比，emmmmm.....

【三】 R环境搭建

1. R的下载与安装

• Windows系统

- 访问 R语言官方网站 <https://www.r-project.org/>
- 在左侧的Downloads标签中点击 [CRAN](#)
- 在镜像列表向下拉找到China，选择任意镜像（推荐[清华源](#)）
- 点击 [Download R for Windows](#)
- 点击 [base](#)
- 点击 [Download R 4.1.2 for Windows](#) （如果有新的版本可能是不同的版本号）
- 打开下载好的.exe文件，按照指示进行安装即可



- Mac系统

- 访问 R语言官方网站 <https://www.r-project.org/>
- 在左侧的Downloads标签中点击 [CRAN](#)
- 在镜像列表向下拉找到China，选择任意镜像（推荐[清华源](#)）
- 点击 [Download R for macOS](#)
- 如果采用英特尔架构，点击 [R-4.1.2.pkg](#)；如果采用ARM架构（M1）点击 [R-4.1.2-arm64.pkg](#)
- 直接点击下载完成的.pkg文件按照指示进行安装即可



- Linux系统
 - 访问 [CRAN](#)
 - 点击 [Download R for Linux](#)
 - 找到自己对应的系统按照相应的介绍进行安装
- 或者
 - 目前主流的Linux发布都有安装R语言的包，可以利用对应的包管理指令进行快捷安装

```
$ sudo apt-get install r-base # for Ubuntu or Debian
```

```
$ sudo yum install R.i386      # for Red Hat or Fedora
```



2. RStudio的下载与安装

RStudio是为了R语言编程而开发的一个IDE，让我们在使用R语言进行编程和分析时更加快捷和明了。

- 访问 [RStudio下载页面](#)
- 下载提示的版本或者找到与自己所使用操作系统相匹配的版本进行下载
- 打开安装文件并像安装其他软件一样按照指示进行安装

注意：RStudio并不内置R语言，所以即使你已经下载安装了RStudio，也需要安装R语言

3. 直接启动RStudio进入R工作环境



RStudio IDE : : CHEAT SHEET

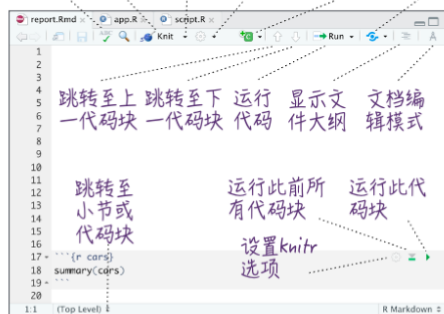


文档和应用

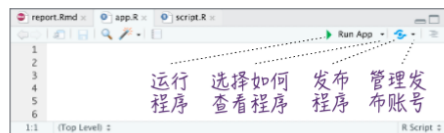


在源码编辑器中打开
Shiny, R Markdown, Knitr,
Sweave, LaTeX, Rd 文件

检查渲染选择输出配置渲染插入代发布至
拼写输出格式染选项码块服务器

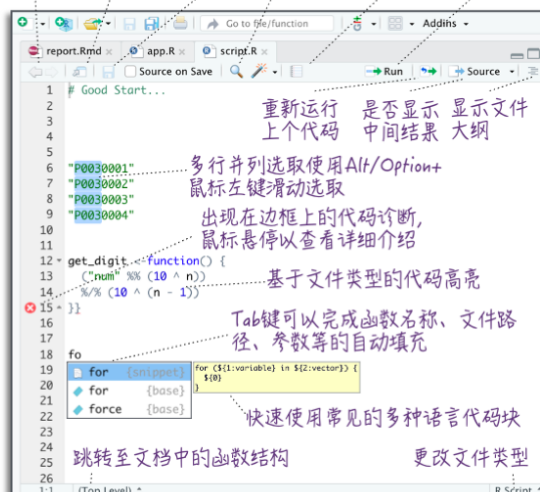


更多Rmarkdown教程可以查阅
Help > Markdown Quick Reference
文档编辑模式将源码即时渲染成文档进行显示
RStudio 将名为 app.R, server.R, ui.R, global.R 的
文档识别为一个 shiny 应用程序



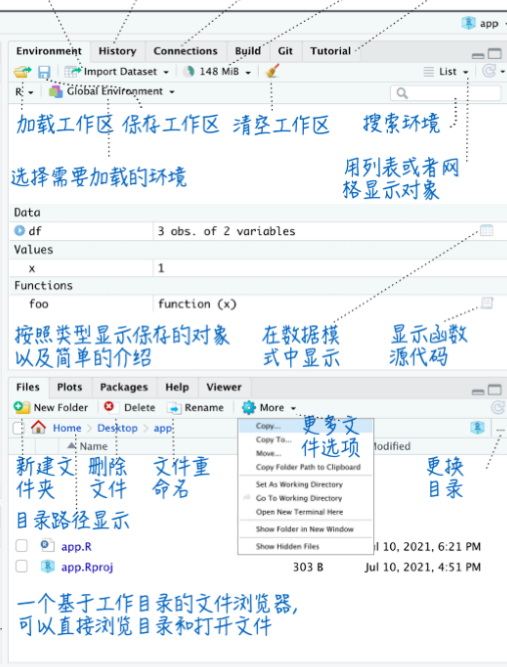
源码编辑器

撤销/重做 在新窗口中打开 保存 查找和替换 编译成笔记 运行所选代码



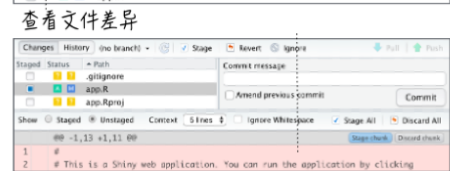
窗口面板

导入数据 操作记录 管理外部数据库 查看内存占用 教程



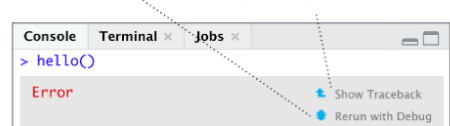
版本管理

在Tools > Project Options > Git/SVN中设置打开



调试模式

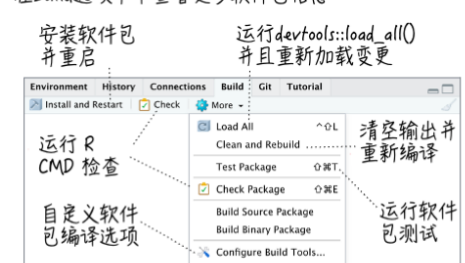
调用 debug(), browser(), 或者添加中断点并且运行程序以开启调试模式
从错误源开始进行调试 打开回溯, 检查错误发生前调用的函数



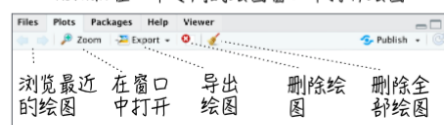
软件包开发

File > New Project > New Directory > R Package
进行新的软件包的开发
Tools > Project Options > Build Tools 添加Roxxygen
文档

更多Roxxygen教程可以查阅Help > Roxxygen Quick Reference
在Build选项卡中查看更多软件包信息



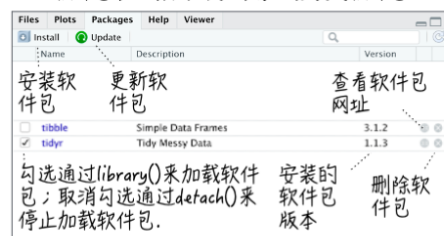
RStudio在一个专门的绘图窗口中打开绘图



RStudio在一个专门的帮助窗口中打开文档



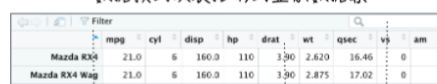
软件包管理器列出了所有已安装的软件包



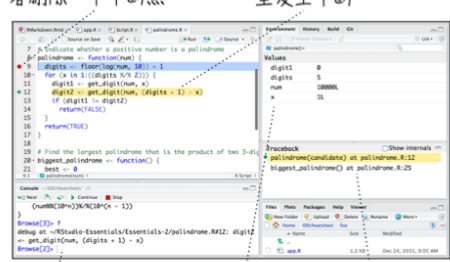
查看器窗口显示HTML内容, 如Shiny应用程序、RMarkdown报告和交互式可视化。



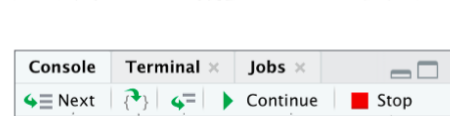
数据模式以表格形式显示数据集



点击行号左侧, 添加或者删除一个中断点



在中断执行的环境 检查执行环境中的变量 选择回溯中的函数来调试



Hello World R!

```
r <- "R"  
message("Hello World ", r, "!")
```

```
## Hello World R!
```

```
a <- 2  
b <- 3  
message("a + b = ", a + b)
```

```
## a + b = 5
```