

植物学家怎么学习 R

张金龙

2021 年 7 月 20 日

嘉道理农场暨植物园

什么是 R?

什么是 R?

- R 是一个统计、计算、绘图的环境，可以运行在 Windows、MacOS、Linux、UNIX 等多种操作系统中运行
- R 既可以看作一个软件，也可以看作指运行在这个软件上的计算机语言

哪个软件跟 R 的关系最近?

- Excel: 数据输入与分析
- SPSS: 统计
- SAS: 统计
- Matlab: 数学建模
- Origin: 绘图
- SigmaPlot: 绘图
- Mathematica: 数学计算和绘图
- Maple: 数学计算

哪个语言跟 R 的关系最近？

- Perl: 生物信息学分析
- Python: 大数据、人工智能
- C: 计算机语言, 编写应用程序
- Cpp (C++): 计算机语言, 编写应用程序
- Fortran: 高性能计算
- Splus: 统计

R 语言是怎么来的？

- 1970 年代，美国贝尔实验室的 John Chambers, Rick Becker, Trevor Hastie, Allan Wilks 等人研发了用于统计分析的 S 语言
- 1988 年，有公司开发了 Splus 商业软件，执行 S 语言，大受欢迎
- 90 年初，新西兰奥克兰大学的 Robert Gentleman 和 Ross Ihaka 开发了一个软件，用于 S 语言的教学
- 1995 年，两人将该软件开源，并命名为 R
- 1997 年，R 由核心管理开始管理，CRAN 上线
- 2000 年，R1.0.0 版本发布

目前版本 R4.1.0

近年来的重要进展

- 2007 年, ggplot2 上线
- 2011 年, Rstudio 发布
- 2014 年, rmarkdown 包发布
- 2016 年, tidyverse 包发布

为什么要学习 R?

植物学家正在从多个角度“研究”植物

1. 系统分类 (描述、命名、分类修订)
2. 编目、保护 (红色名录评估)
3. 基因组学 (鉴定、推断系统发育关系)
4. DNA 条形码 (鉴定、推断系统发育关系)
5. 群体遗传学 (推断种群历史以及多种指标)
6. 物种潜在分布区预测
7. 标本数字化 (数据的清洁和标准化)
8. 标本馆管理 (标签、名录更新)
9. “科普” (微信公众号、博客) 跟普通公众交流

植物学研究的一些新变化

1. 数据量巨大 (Big data) :

动辄几万、几十万条数据，一般难以手动处理

2. 复杂的工作流程 (working flow)

越来越依靠 R、Python、Perl、Ruby 等计算机语言，代码本身包括了工作流

3. 可重复性 (Reproducibility):

数据收集的方法以及各种细节，保证其他人在看过材料与方法之后，就能将研究重复出来。

举例：不仅仅要提供软件，还要提供软件的版本和运行该软件的代码

4. 开放科学 open science:

开放获取文章、开放数据、开放代码、开放审稿意见

植物学家需要掌握的工具越来越多

1. 数据分析汇总: Excel、SPSS 或者 SAS
2. 建立进化树, 如 BEAST、MEGA、paup*、MrBayes、RAxML...
3. 绘图软件: sigmaPlot、Excel、PhotoShop、Adobe Illustrator、Origin、FigTree、Geneious、paint.net、Inkscape
4. GIS: ArcGIS、QGIS
5. 物种潜在分布区分析: Maxent
6. 系统发育比较分析: Mesquite、PAML...
7. 编写植物名录: Access、MySQL
8. 文献管理: Zotero、Mendeley、Endnote 等等

如果能有一个软件, 能将上述大部分的工作做了, 而且, 过程可重复, 不好吗?

对植物学家来说，R 能做什么？

1. 计算器
2. 基本的统计分析 (t 检验、方差分析、回归)
3. 绘图 (各种复杂的图形)
4. 空间信息处理以及地图绘制 (GIS)
5. 做幻灯片、文档
6. 编写植物名录
7. 开发自己的模型、算法
8. 还有上一页软件的几乎所有功能！

植物学家真的需要 R 吗？

植物学家真的需要 R 吗？

- 用 R 处理植物图库
- 植物分布以及区系分区
- 进化树与系统发育分析
- 人工智能和机器学习（如果真有兴趣……）
- Rmarkdown 与可重复研究

用 R 处理植物信息？

- 批量重命名照片
- 批量移动照片到所在的科、属的文件夹
- 批量提取地理坐标并生成 kml 文件
- 批量裁剪照片添加水印
- 批量生成物种分布图
- 获取 GBIF 上的植物标本信息

上述都可通过编写简单的 R 脚本实现

用 plantlist 包查询科属

1. 批量查询科属：你熟悉 APG 系统吗？
2. 根据中文名查询学名
3. 根据拉丁名查询中文名和分布
4. 根据拉丁名查询接受状态

用 plantlist 包生成植物名录

根据中文名，查询出学名，分布地点，并按照科属顺序排列

```
library(plantlist)
make_checklist(
  checklist_dat = CTPL(
    c("狼毒", "粗叶卷柏",
      "侧金盏花", "土沉香")),
  outfile = "output_make_checklist.md",
  theme = "minimal")
```

用 plantlist 包生成植物名录

2 Ferns and lycophytes

3 卷柏科 Selaginellaceae

粗叶卷柏

4 Angiosperms

114 毛茛科 Ranunculaceae

侧金盏花

Adonis amurensis Regel et Radde

252 瑞香科 Thymelaeaceae

土沉香

Aquilaria sinensis (Lour.) Spreng.

图 1: Markdown 格式的植物名录

用 herblabel 包生成标本标签

Kadoorie Farm and Botanic Garden Herbarium (KFBG)

EUPHORBIACEAE

Excoecaria agallocha L.

HONG-KONG, New Territories, Sai Kung District, Yung Shue O
22°25'45"N, 114°17'11"E; 2m

Tree 3m tall. Sea shore, along the sea shore, within the mangroves. Not common. Growth with *Kandelia candel*, *Bruguiera gymnorhiza*, *Avicennia marina*.

Jin-Long Zhang, #JL0019 · · · · · → 13 · August · 2014

→ Det.: Jin-Long Zhang, 04 February 2015

图 2: KFBG 主题

用 herblabel 包生成标本标签

Kadoorie·Farm·and·Botanic·Garden·Herbarium·(KFBG)

Collector:·Jin-Long·Zhang

No.:·JL0019·····**Date·of·Coll.:**·13·August·2014

Location:·HONG·KONG, New·Territories, Sai·Kung·District, ·Yung·Shue·O

Description:·Tree·3m·tall··

Remarks:·Sea·shore, ·along·the·sea·shore, ·within·the·mangroves. ·Not·common. ·Growth with: *Kandelia candel*, *Bruguiera gymnorhiza*, *Avicennia marina*. ·

Alt.:·2m; ·**Lat.:**·22°25'45"N, ·**Lon.:**·114°17'11"E

Family:·Euphorbiaceae·····**Local·Name:**·

Species:·*Excoecaria agallocha*·L·

图 3: PE 主题

用 herblabel 包生成标本标签

Euphorbiaceae

Excoecaria agallocha L.

New Territories, Sai Kung District, Yung Shue O. 22°25'45"N, 114°17'11"E; 2m. Tree 3m tall. Sea shore, along the sea shore, within the mangroves. Not common. Growth with *Kandelia candel*, *Bruguiera gymnorhiza*, *Avicennia marina*.

Jin-Long Zhang

JL0019

→

13 August 2014

**Kadoorie Farm and Botanic Garden Herbarium
(KFBG)**

图 4: HU 主题标本标签

植物分布以及区系分区

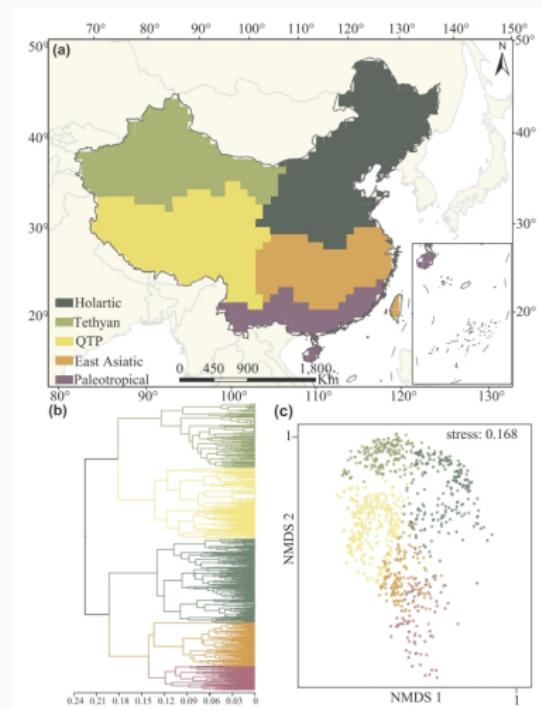


Fig. 3. Map (a) and dendrogram (b) based on UPGMA hierarchical clustering

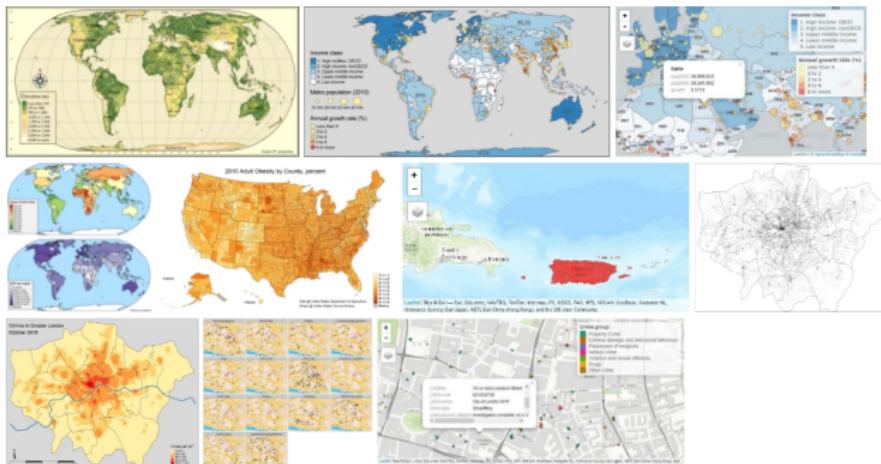
图 5: 叶建飞等

绘制地图 (sf + ggplot2 + tmap + raster + rgdal + spatial)

tmap: thematic maps in R

[build](#) [passing](#) License [GPL v3](#) CRAN [3.3-2](#) CRAN [ERROR](#) downloads [24K/month](#)

`tmap` is an actively maintained open-source R-library for drawing thematic maps. The API is based on [A Layered Grammar of Graphics](#) and resembles the syntax of `ggplot2`, a popular R-library for drawing charts.



See [below](#) the source code for these images.

phyloregion



An R package for biogeographic regionalization (the classification of geographical areas in terms of their biotas) and macroecology. With macroecological datasets of ever increasing size and complexity, `phyloregion` offers the possibility of handling and executing very large scale analysis of biogeographic regionalization. It also allows fast and memory efficient analysis of standard conservation measures such as phylogenetic diversity, phylogenetic endemism, evolutionary distinctiveness and global endangerment. `phyloregion` can run on any operating system (Mac, Linux, Windows or even high performance computing cluster) with R 3.6.0 (or higher) installed.

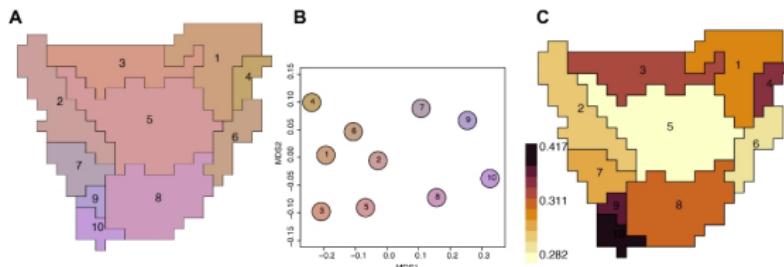


图 7: phyloregion 包

物种潜在分布区预测

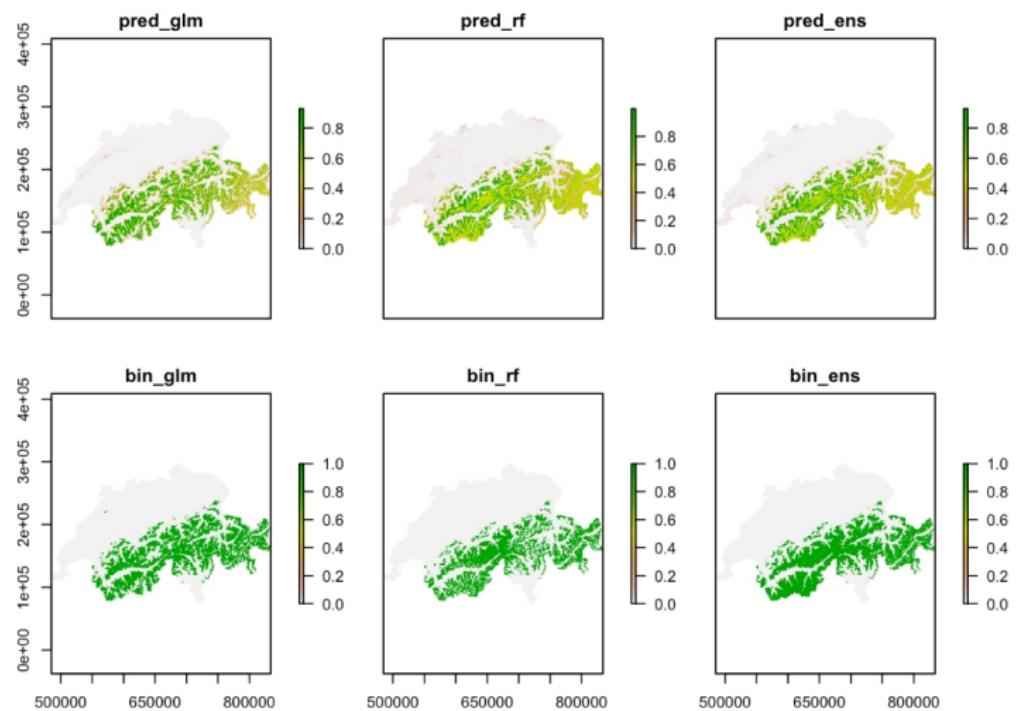


图 8: 物种潜在分布区预测

(<https://damariszurell.github.io/SDM-Intro/>)

绘制进化树

B. Xue, et al.

Molecular Phylogenetics and Evolution 142 (2020) 106659

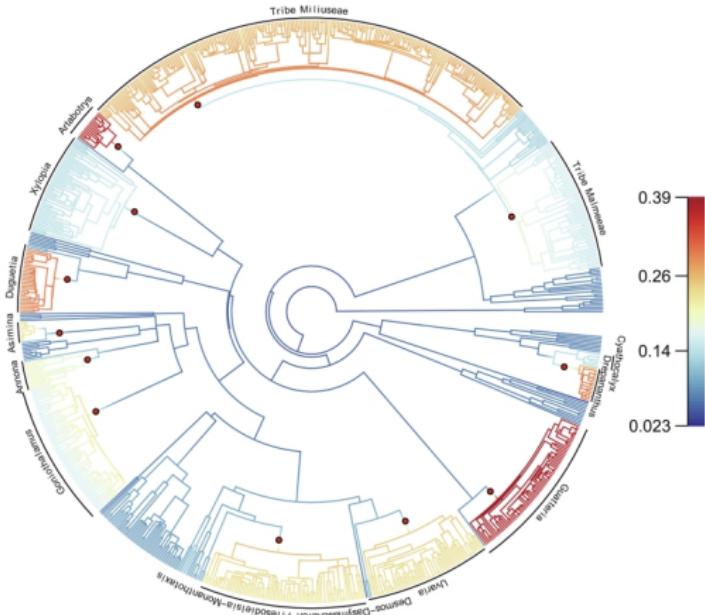


Fig. 1. Phylorate plot of Annonaceae with branches colored according to net diversification rate (Myr^{-1}), resulting from BAMM analysis. Red dots indicate diversification rate shifts.

重建祖先性状 ape

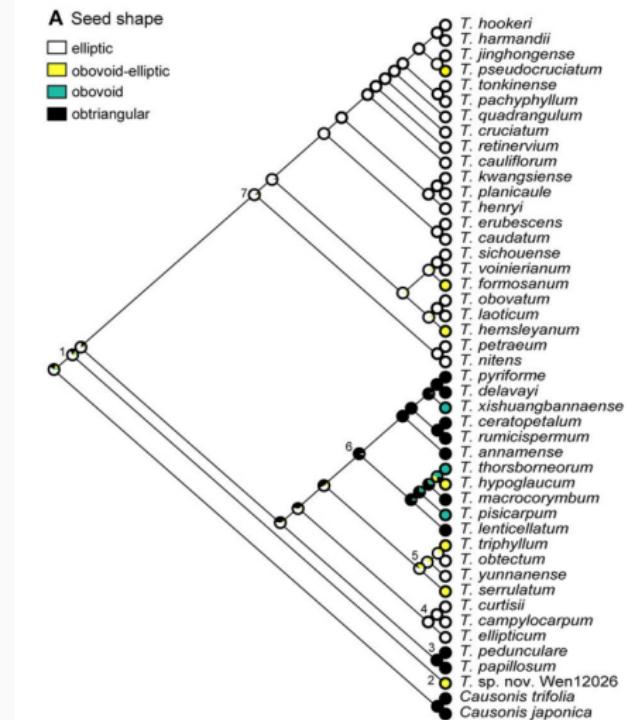


图 10: Habib et al

进化树与其他信息组合 ggtree

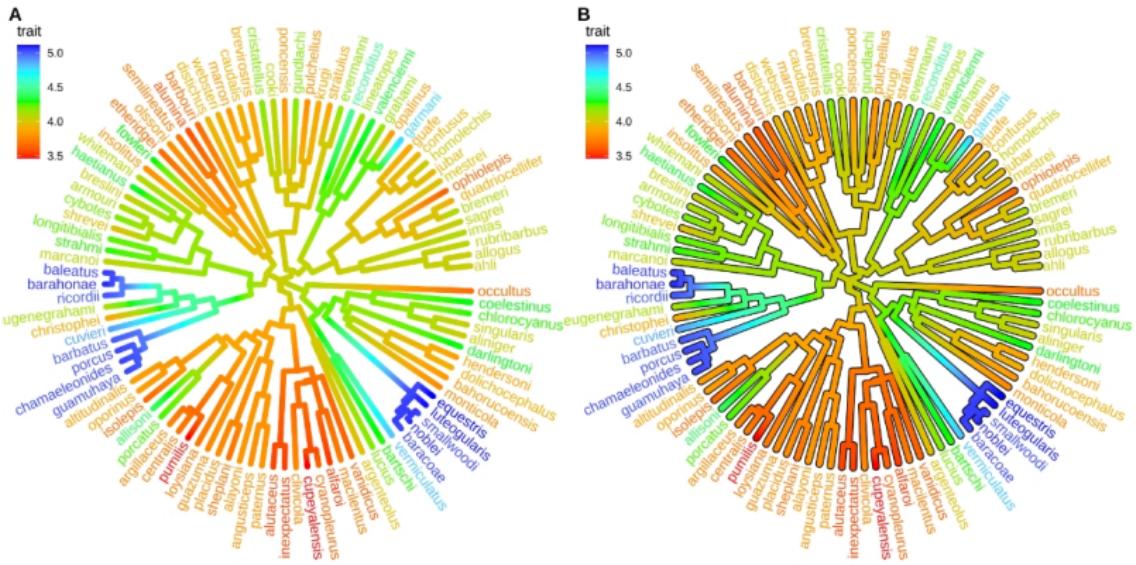


Figure 4.11: **Continuous state transition in edges.** Edges are colored by the values from ancestral trait to offspring.

祖先分布区推断 biogeobears

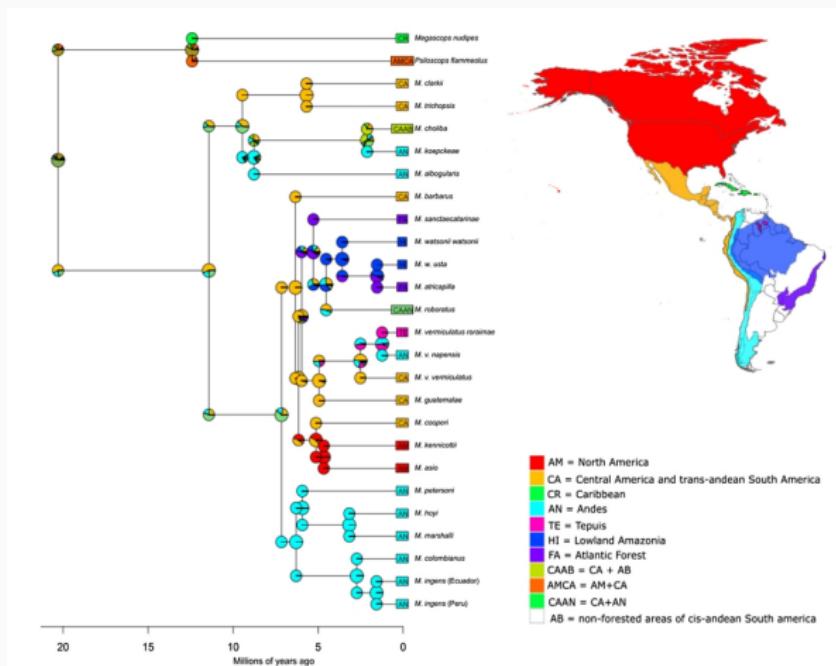


Fig. 3. Ancestral area reconstruction from BioGeoBEARS, derived from the Bayesian chronogram. The best-fit model was DIVALIKE+. Most likely biogeographic areas are shown in the circles, and the colors in the squares indicate the current species distribution. The map shows the location of the areas used. Combinations of two areas (for example CAAN) are not shown on the map. We considered Mexico as Central America for this study.

图 12: Dantas et al. 2015

关键性状进化速率

B. Xie, et al.

Molecular Phylogenetics and Evolution 142 (2020) 106659

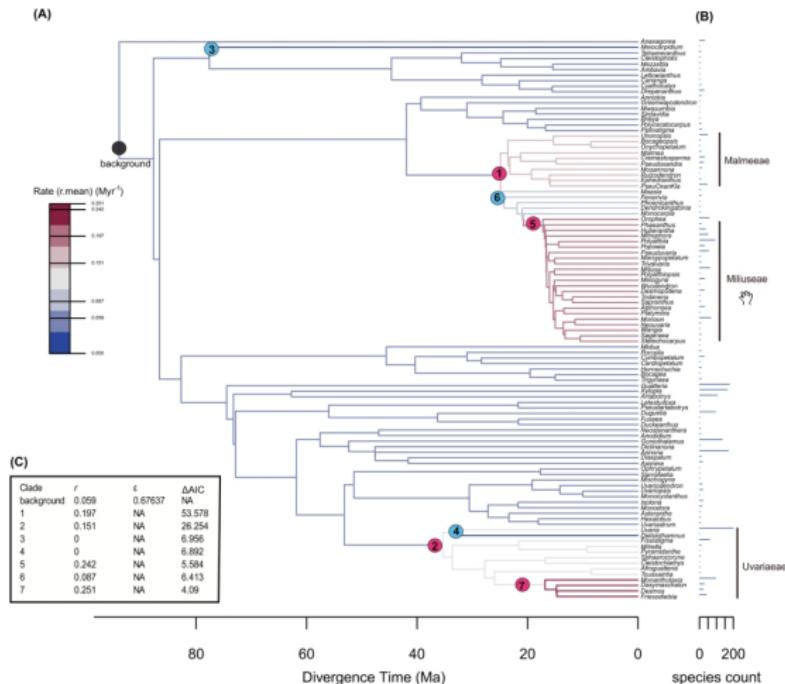


图 13: 薛彬娥等 2020

植物学家怎么学习 R?

怎么学习 R?

植物学家：如果没有很好的数学和计算机基础，能学好吗？

答：可以

1. 搭建好 R 的运行环境
2. 参加培训班或者看看 R 的学习视频
3. 理解数据的读入和 R 中数据的类型以及操作的方法
4. 练习、练习、练习

R 数据分析平台的搭建（技术性很强）

R 语言平台的搭建

安装以下软件

1. R 软件: 运行 R 语言
2. Rtools: Windows 编译 R 程序包的软件
3. Rstudio: R 的集成开发环境
4. tinytex: texlive 的子集, 用于编译 pdf 等
5. git: 控制版本
6. vscode: 编写文本

安装 R 软件



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

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 4.1.0 \(Camp Pontanezen\)](#) has been released on 2021-05-18.
- [R version 4.0.5 \(Shake and Throw\)](#) was released on 2021-03-31.
- Thanks to the organisers of useR! 2020 for a successful online conference. Recorded tutorials and talks from the conference are available on the [R Consortium YouTube channel](#).
- You can support the R Foundation with a renewable subscription as a [supporting member](#)

图 14: R 首页

安装 R

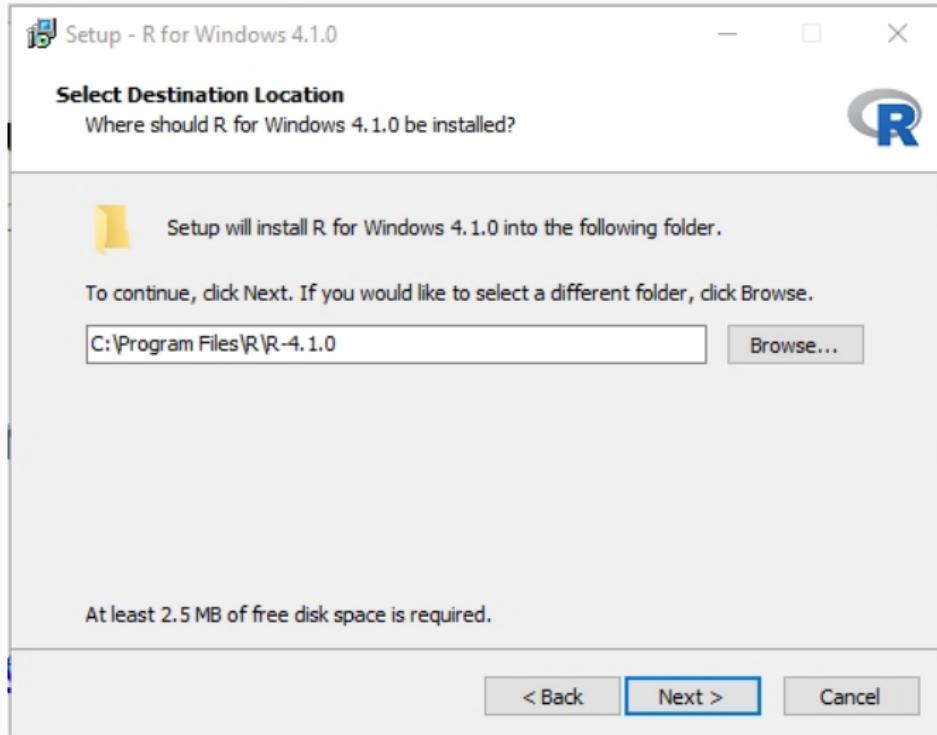
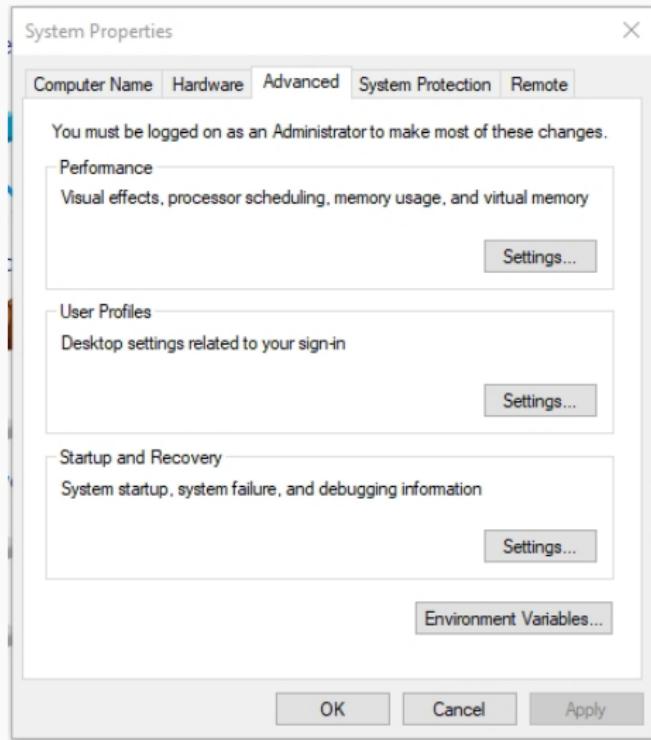


图 15: 默认路径

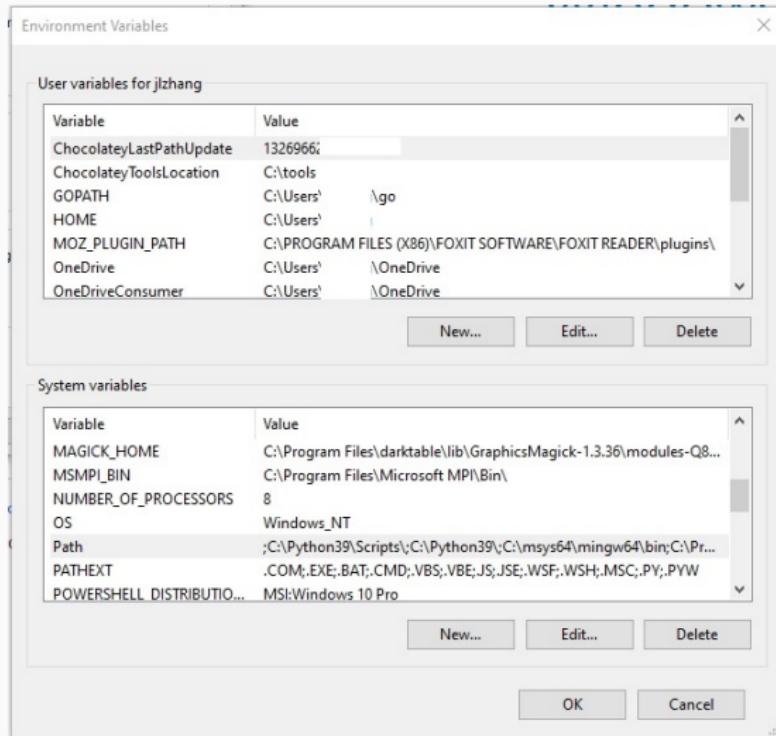
将 R 添加到系统路径

This PC > Local Disk (C:) > Program Files > R > R-4.1.0 > bin > x64			
Name	Date modified	Type	Size
open.exe	18/5/2021 9:41 am	Application	20 KB
R.dll	18/5/2021 9:41 am	Application exten...	30,855 KB
R.exe	18/5/2021 9:41 am	Application	105 KB
Rblas.dll	18/5/2021 9:41 am	Application exten...	497 KB
Rcmd.exe	18/5/2021 9:41 am	Application	105 KB
Rfe.exe	18/5/2021 9:41 am	Application	88 KB
Rgraphapp.dll	18/5/2021 9:40 am	Application exten...	640 KB
Rgui.exe	18/5/2021 9:41 am	Application	88 KB
Riconv.dll	18/5/2021 9:40 am	Application exten...	359 KB
Rlapack.dll	18/5/2021 9:42 am	Application exten...	2,477 KB
Rscript.exe	18/5/2021 9:41 am	Application	93 KB
RSetReg.exe	18/5/2021 9:41 am	Application	89 KB
Rterm.exe	18/5/2021 9:41 am	Application	89 KB

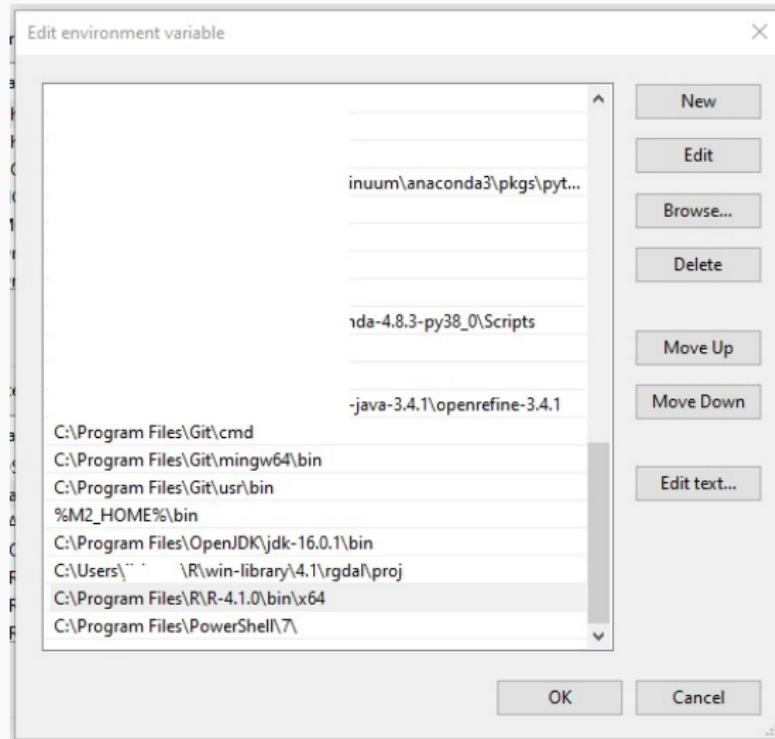
将 R 添加到系统路径



将 R 添加到系统路径



将 R 添加到系统路径



将 R 添加到系统路径

```
C:\Users\██████>R

R Under development (unstable) (2021-02-27 r80043) -- "Unsuffered
Copyright (C) 2021 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.

Natural language support but running in an English locale

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.

> -
```

安装开发工具 Rtools (编译程序包)

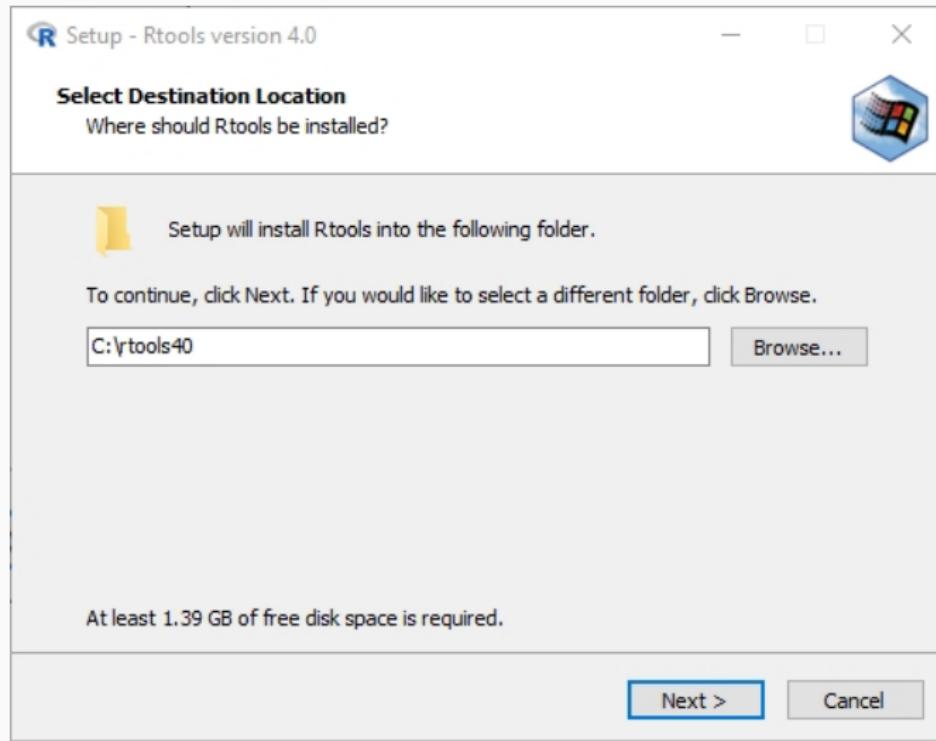
Installing Rtools40

Note that rtools40 is only needed build R packages with C/C++/Fortran code from source. By default, R for Windows installs the precompiled “binary packages” from CRAN, for which you do not need rtools!

To use rtools40, download the installer from CRAN:

- On Windows 64-bit: [rtools40v2-x86_64.exe](#) (recommended: includes i386, x64, and x64-ucrt compilers)
- On Windows 32-bit: [rtools40-i686.exe](#) (i386 compilers only)

安装 Rtools



将 Rtools 添加到系统路径

Putting Rtools on the PATH

After installation is complete, you need to perform **one more step** packages: you need to put the location of the Rtools *make utilities* on the system's PATH. The easiest way to do so is create a text file `.Renvirons` in your `Documents` folder with the following line:

```
PATH="${RTOOLS40_HOME}\usr\bin;${PATH}"
```

You can do this with a text editor, or from R like so (note that in R backslashes):

```
writeLines('PATH="${RTOOLS40_HOME}\\\\usr\\\\bin;${PATH}"', con = "~/.Renvirons")
```

将 Rtools 添加到系统路径（安装成功了吗？）

```
Command Prompt
Microsoft Windows [Version 10.0.19041.1052]
(c) Microsoft Corporation. All rights reserved.

C:\Users\████████>make
make: *** No targets specified and no makefile found. Stop.

C:\Users\jlzhang>█
```

安装开发工具 tinytex (编译文档)



TinyTeX: A lightweight and easy-to-maintain LaTeX distribution

安装 tinytex

```
> install.packages("tinytex")
Installing package into ‘C:/Users/_/_/_/R/win-library/4.1’
(as ‘lib’ is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.1/tinytex_0.32.zip'
Content type 'application/zip' length 123465 bytes (120 KB)
downloaded 120 KB

package ‘tinytex’ successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\_\_\_\AppData\Local\Temp\RtmpAhazeh\downloaded_packages
> tinytex()
Error in tinytex() : could not find function "tinytex"
> library(tinytex)
> tinytex()
Error in tinytex()
> ?tinytex
No documentation available for 'tinytex' in package(s) tinytex and libraries:
you could try ‘??tinytex’
> install_tinytex()
trying URL 'https://yihui.org/tinytex/TinyTeX-1.zip'
Content type 'application/octet-stream' length 103777374 bytes (99.0 MB)
```

怎样知道成功安装了 tinytex?

```
Command Prompt - pdflatex
Microsoft Windows [Version 10.0.19041.1052]
(c) Microsoft Corporation. All rights reserved.

C:\Users\██████>pdflatex
This is pdfTeX, Version 3.141592653-2.6-1.40.22 (TeX Live 2021/W32TeX)
 restricted \write18 enabled.

**
```

版本控制软件 git



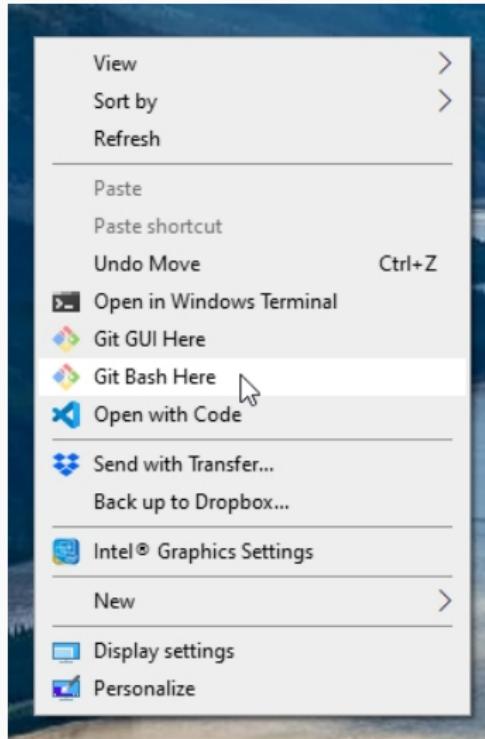
git --distributed-even-if-your-workflow-isnt

Git is a **free and open source** distributed version control system designed to handle everything from small to very large projects with speed and efficiency.

Git is **easy to learn** and has a **tiny footprint with lightning fast performance**. It outclasses SCM tools like Subversion, CVS, Perforce, and ClearCase with features like **cheap local branching**, convenient staging areas, and **multiple workflows**.



git 安装成功了么？



git 安装成功了!

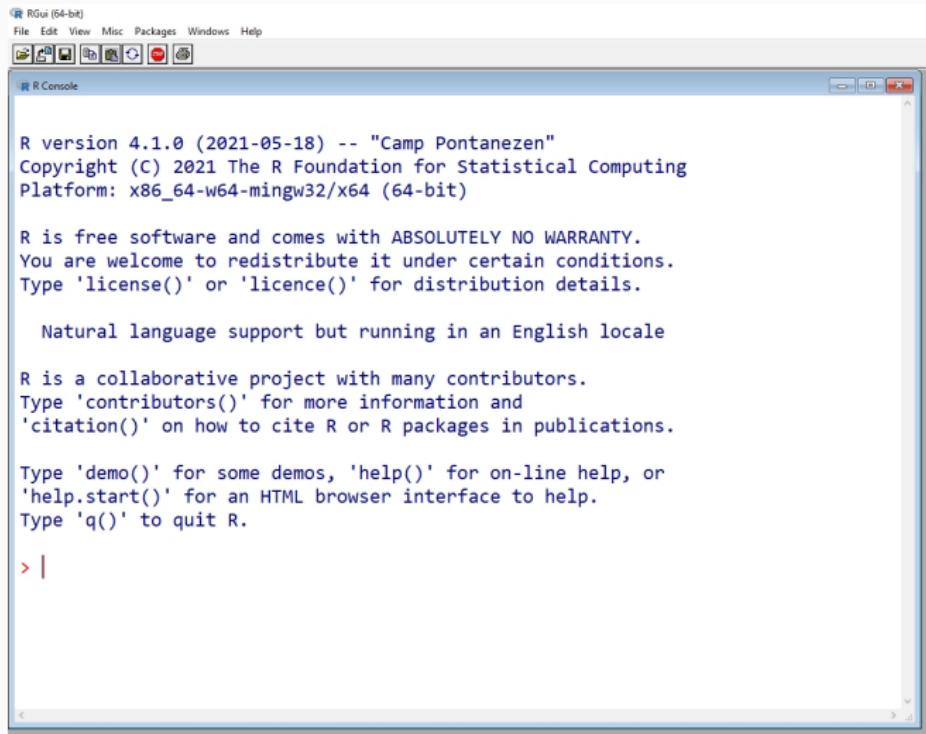
```
MINGW64:/c/Users/jlzhang/Desktop
MINGW64 ~/Desktop
$ git
usage: git [--version] [--help] [-C <path>] [-c <name>=<value>
   [--exec-path[=<path>]] [--html-path] [--man-path]
   [-p | --paginate | -P | --no-pager] [--no-replace-branch]
   [--git-dir=<path>] [--work-tree=<path>] [--namespace=<name>]
   [--super-prefix=<path>] [--config-env=<name>=<envvar>]
   <command> [<args>]

These are common Git commands used in various situations:

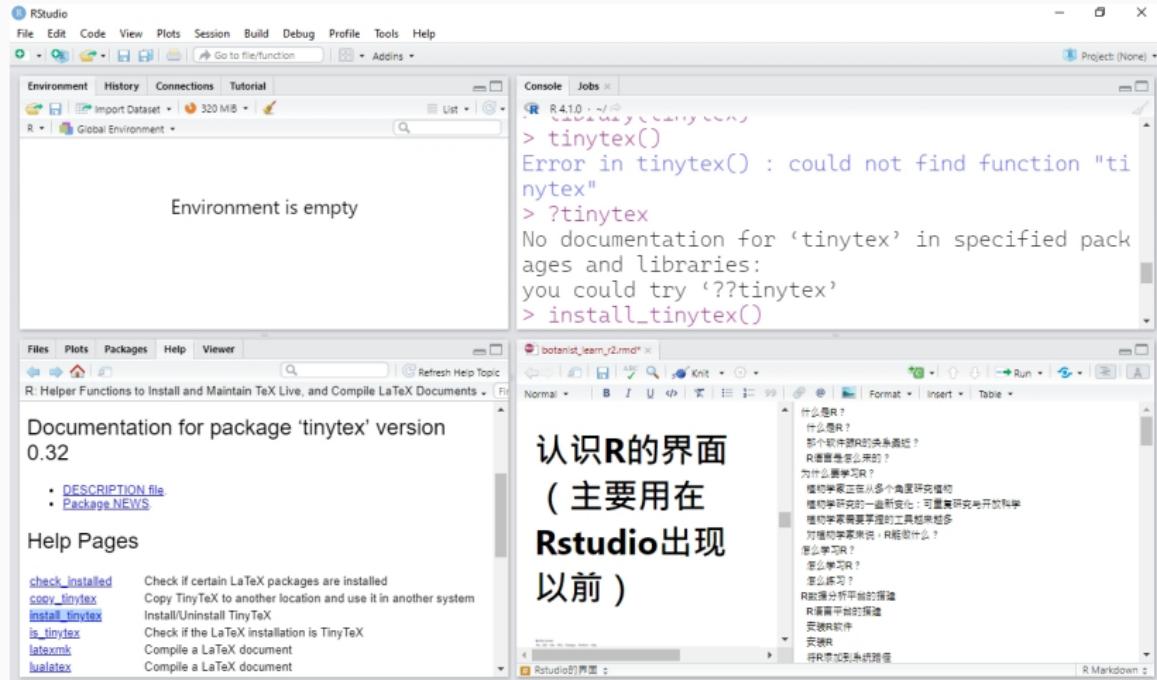
start a working area (see also: git help tutorial)
  clone          Clone a repository into a new directory
  init           Create an empty Git repository or reinit
  one
```

R 基础

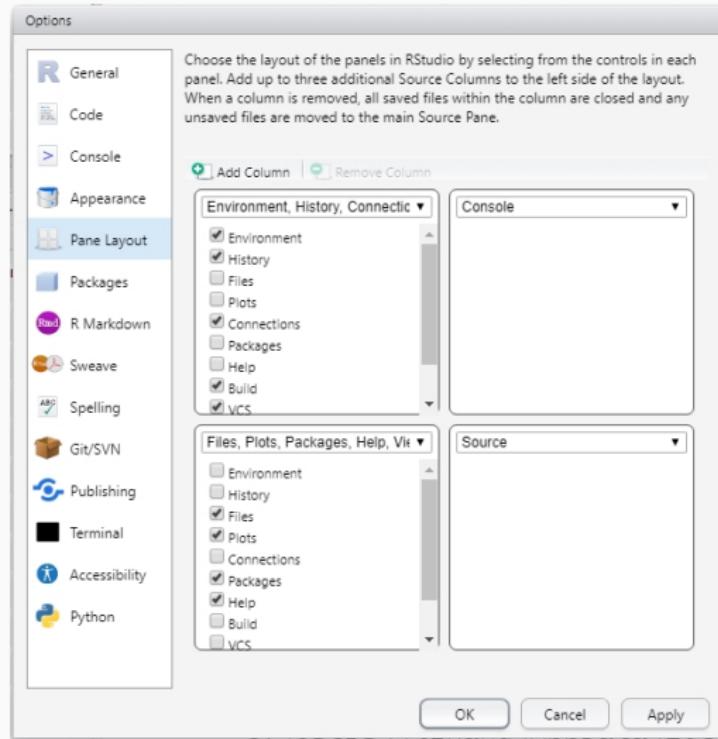
认识 R 的界面 (主要用在 Rstudio 出现以前)



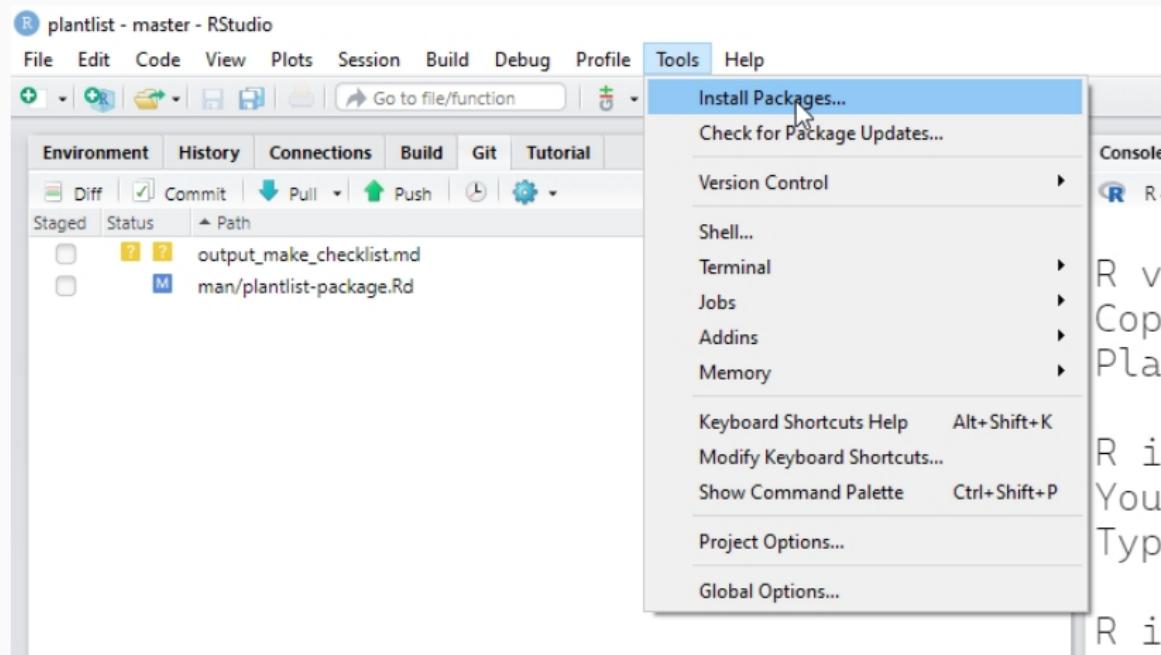
Rstudio 的界面



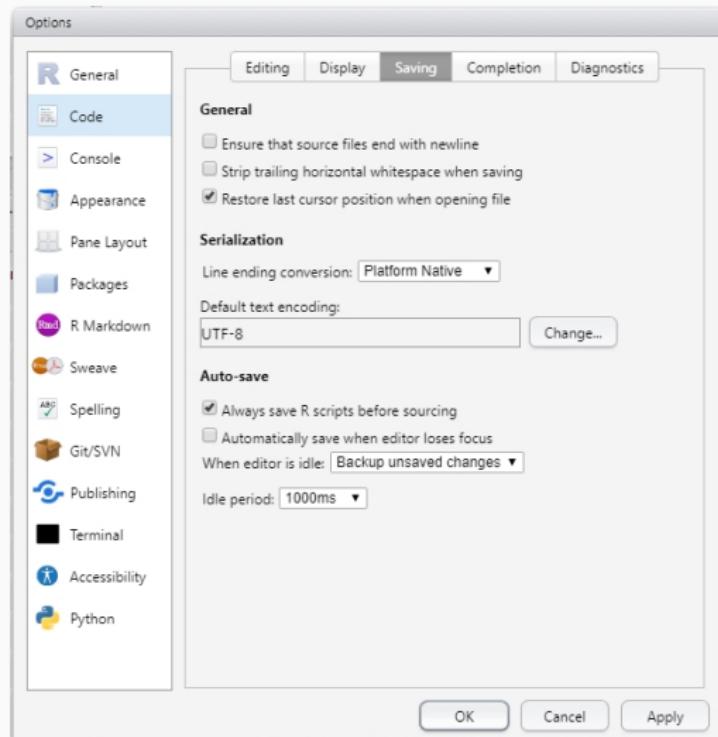
Rstudio 常用的选项 (页面布局)



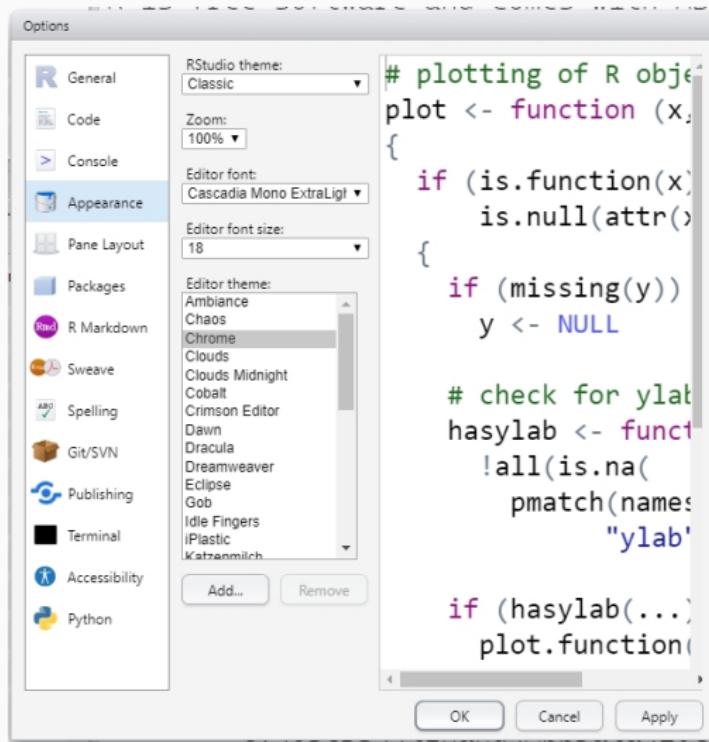
Rstudio 常用的选项 Tools



Rstudio 常用的选项保存和读取 UTF8 编码的文本文件



Rstudio 常用的选项 (改变主题和字体)



文本编辑器 VScode

Visual Studio Code Docs Updates Blog API Extensions FAQ Learn [Search Docs](#) [Download](#)

Version 1.58 is now available! Read about the new features and fixes from June.

Code editing. Redefined.

Free. Built on open source. Runs everywhere.

[Download for Windows](#) Stable Build [Other platforms and Insiders Edition](#)

By using VS Code, you agree to its license and privacy statement.

EXTENSIONS MARKETPLACE

- Python 2019.6.24221 ⚡ 54.9M • 45 Microsoft Linting, Debugging (multi-threaded, ... [Install](#)
- GitLens — Git super... 9.8.5 ⚡ 23.1M • 5 Eric Amoako Supercharge the Git capabilities built-in [Install](#)
- C/C++ 0.24.0 ⚡ 23M • 35 Microsoft C/C++ + IntelliSense, debugging, and ... [Install](#)
- ESLint 1.0.0 ⚡ 21.9M • 45 Dirk Beamer Integrates ESLint JavaScript linter into VS... [Install](#)
- Debugger for Ch... 4.11.6 ⚡ 20.8M • 4 Microsoft Debug your JavaScript code in the C... [Install](#)
- Language Supp... 0.47.0 ⚡ 18.7M • 45 Red Hat Java Linting, Intellisense, formatting, ... [Install](#)
- vscode-icons 8.8.0 ⚡ 17.2M • 5 VSCode Icons Team Icons for Visual Studio Code [Install](#)
- Vetur 0.21.1 ⚡ 1.7M • 45 Pine Wu Vue tooling for VS Code [Install](#)
- CP 1.21.0 ⚡ 15.6M • 4 Microsoft CP for Visual Studio Code (powered by ... [Install](#)

App.js JS index.js JS serviceWorker.js

```
src > JS serviceWorker.js > @Register > @window.addEventListener('load') callback
    checkValidServiceWorker(swUrl, config);
    // Add some additional logging to localhost, p...
    navigator.serviceWorker.ready.then(() => {
        ...
        product
        productSub
        ...
        removeSiteSpecificTrackingException
        ...
        removeWebWideTrackingException
        ...
        requestMediaKeySystemAccess
        ...
        sendBeacon
        ...
        serviceWorker (property) Navigator.serviceWork...
        ...
        storage
        ...
        storeSiteSpecificTrackingException
        ...
        storeWebWideTrackingException
    })
    ...
    vendor
}

function registerValidSw(swUrl, config) {
    navigator.serviceWorker
        .register(swUrl)
        .then(registration => {
            ...
            Local: http://localhost:3000/
            On Your Network: http://10.211.55.8:3000/
        })
}

```

TERMINAL

You can now view `create-react-app` in the browser.

Local: <http://localhost:3000/>
On Your Network: <http://10.211.55.8:3000/>

Note that the development build is not optimized.

In 43, Col 19 | Spaces 2 | UTF-8 | LF | JavaScript

R 程序包

什么是 R 程序包?

- R 程序包是包含 R 函数、文档、数据以及相应说明的一个文件夹或者编译好的压缩文件。
- 一个程序包在 R 环境内执行一个或者多个特定的功能。

R 的功能靠程序包实现

- **base 系列程序包**: R 的基础程序包，可以实现基本的统计、绘图、文字处理、文件操作等
- **tidyverse 系列程序包**: 主要用于数据科学
- **某一领域专用的系列程序包**: 如系统发育分析、生物多样性分析、空间分析绘图等

一些常用的程序包

1. ape: 系统发育与进化分析
2. tidyverse: 包括 ggplot2、purrr、tibble、dplyr、tidyR、stringr、readr、forcats 等，主要用于数据科学
3. sf+ggplot2+tmap: 空间数据的读取、分析以及绘图
4. ggtree: 绘制进化树

一些常用的程序包

8. knitr + rmarkdown: 编译 markdown 文档, 生成 word、html、pdf、powerpoint 等
9. testthat、devtools、remotes: R 程序包的安装和测试
10. dismo、ENMeval、wallace: 物种潜在分布区分析
11. meta analysis: 荟萃分析
12. shiny: 制作动态网页
13. Rsqlite + dbplyr: 操作数据库
14. plantlist: 查询科属和生成植物名录
15. herblabel: 生成植物标本标签

在哪里找程序包？

1. CRAN: 各种类别，参考 CRAN task views，用 `install.packages()` 安装
2. Bioconductor: 主要是基因组学和生物信息学工具
3. github: 多为开发阶段的程序包，用 `devtools` 安装
4. Rforge: 多为开发阶段的程序包，现在已不太活跃

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

清华大学 R 镜像

China

<https://mirrors.tuna.tsinghua.edu.cn/CRAN/> TUNA Team, Tsinghua University

<https://mirrors.bfsu.edu.cn/CRAN/> Beijing Foreign Studies University

<https://mirrors.ustc.edu.cn/CRAN/> University of Science and Technology of China

<https://mirror-hk.koddos.net/CRAN/> KoDDoS in Hong Kong

<https://mirrors.e-ducation.cn/CRAN/> Elite Education

<https://mirror.lzu.edu.cn/CRAN/> Lanzhou University Open Source Society

<https://mirrors.nju.edu.cn/CRAN/> eScience Center, Nanjing University

<https://mirrors.tongji.edu.cn/CRAN/> Tongji University

<https://mirrors.sjtug.sjtu.edu.cn/cran/> Shanghai Jiao Tong University

<https://mirrors.sustech.edu.cn/CRAN/> Southern University of Science and Technology (SUSTech)

Bioconductor

[Home](#)[Install](#)[Help](#)

[Home](#) » About

Bioconductor

The *Bioconductor* mission is to promote the statistical analysis and comprehension of current and emerging high-throughput biological assays. *Bioconductor* is based on packages written primarily in the [R](#) programming language. Bioconductor is committed to open source, collaborative, distributed software development and literate, reproducible research. Enabling user and developer communities is an essential part of our mission. [Scientific](#), [Technical](#) and [Community](#) Advisory Boards provide project oversight.

https://github.com/helixcn

Search or jump to...

Pull requests Issues Marketplace Explore

Overview Repositories 233 Projects Packages

Jinlong Zhang
helixcn

Ecologist working on species distribution

Edit profile

133 followers · 305 following · 694

EcoEvoLinks
Links for Ecology and Evolution
19 stars · 14 forks

phylotools
Phylogenetic tools for Eco-phylogenetics
R package · 4 stars · 5 forks

R_phylo_refcard
PDF Reference Card for Phylogenetic Comparative Analysis in R
TeX · 4 stars · 5 forks

安装 CRAN 上的程序包

```
install.packages("ape")
```

安装一个

```
install.packages("tidyverse",
```

```
  dependencies = TRUE)
```

安装一系列，并安装所有依赖的程序包

安装 Github 上的程序包

从源代码安装 github 上的 phylotools 包

```
library(devtools)
install_github("helixcn/phylotools",
  build_vignettes = TRUE)
```

如果还没有安装 devtools，则输入

```
install.packages("devtools")
```

安装 Bioconductor 上的程序包

```
if (!requireNamespace("BiocManager",
  quietly = TRUE))
  install.packages("BiocManager")

BiocManager::install("ggtree")
```

安装 CRAN 上的一类程序包

安装 CRAN 上某一领域（如 Phylogenetics）的相应程序包

```
library(ctv)
install.views("Phylogenetics")
```

植物学研究一般参考 Phylogenetics 和 Environmetrics

[Optimization](#)

[Pharmacokinetics](#)

[Phylogenetics](#)

[Psychometrics](#)

[ReproducibleResearch](#)

[Robust](#)

[SocialSciences](#)

[Spatial](#)

[SpatioTemporal](#)

[Survival](#)

[TeachingStatistics](#)

Optimization and Mathematical Programming

Analysis of Pharmacokinetic Data

Phylogenetics, Especially Comparative Methods

Psychometric Models and Methods

Reproducible Research

Robust Statistical Methods

Statistics for the Social Sciences

Analysis of Spatial Data

Handling and Analyzing Spatio-Temporal Data

Survival Analysis

Teaching Statistics

CRAN Task View: Phylogenetics

CRAN Task View: Phylogenetics, Especially Comparative Methods

Maintainer: Brian O'Meara

Contact: omeara.brian at gmail.com

Version: 2021-02-25

URL: <https://CRAN.R-project.org/view=Phylogenetics>

The history of life unfolds within a phylogenetic context. Comparative phylogenetic methods analyze historical patterns along phylogenetic trees. This task view describes R packages for comparative phylogenetic methods. This is an active research area and much of the information is that many important packages are not on CRAN: either they were formerly on CRAN, for example, if they failed to incorporate necessary changes as R is updated) or they are developed on CRAN yet. Such packages may be found on GitHub, R-Forge, or authors' websites.

有多少程序包？

Contributed Packages

Available Packages

Currently, the CRAN package repository features 17837 available packages.

[Table of available packages, sorted by date of publication](#)

[Table of available packages, sorted by name](#)

截至今天 2021-07-10

CRAN 提供的 R 程序包的页面

picante: Integrating Phylogenies and Ecology

Functions for phylocom integration, community analyses, null-models, traits and evolution. Implements numerous ecophylogenetic approaches including measures of community phylogenetic and trait diversity, phylogenetic signal, estimation of trait values for unobserved taxa, null models for community and phylogeny randomizations, and utility functions for data input/output and phylogeny plotting. A full description of package functionality and methods are provided by Kembel et al. (2010)
[<doi:10.1093/bioinformatics/btq166>](https://doi.org/10.1093/bioinformatics/btq166).

Version:	1.8.2
Depends:	ape , vegan , nlme , methods
Suggests:	brglm , circular , corpcor , quantreg
Published:	2020-06-10
Author:	Steven W. Kembel, David D. Ackerly, Simon P. Blomberg, Will K. Cornwell, Peter D. Cowan, Matthew R. Helmus, Helene Morlon, Campbell O. Webb
Maintainer:	Steven W. Kembel <steve.kembel at gmail.com>
License:	GPL-2
NeedsCompilation:	yes
Citation:	picante citation info
Materials:	README
In views:	Phylogenetics
CRAN checks:	picante results

CRAN 提供的 R 程序包的页面：以 picante 包为例

Downloads:

Reference manual: [picante.pdf](#)

Vignettes: [Data and analysis in picante](#)

Package source: [picante_1.8.2.tar.gz](#)

Windows binaries: r-devel: [picante_1.8.2.zip](#), r-devel-UCRT: [picante_1.8.2.zip](#), r-release: [picante_1.8.2.zip](#),
r-oldrel: [picante_1.8.2.zip](#)

macOS binaries: r-release (arm64): [picante_1.8.2.tgz](#), r-release (x86_64): [picante_1.8.2.tgz](#), r-oldrel:
[picante_1.8.2.tgz](#)

Old sources: [picante archive](#)

Reverse dependencies:

Reverse depends: [CommEcol](#), [idar](#), [RPANDA](#)

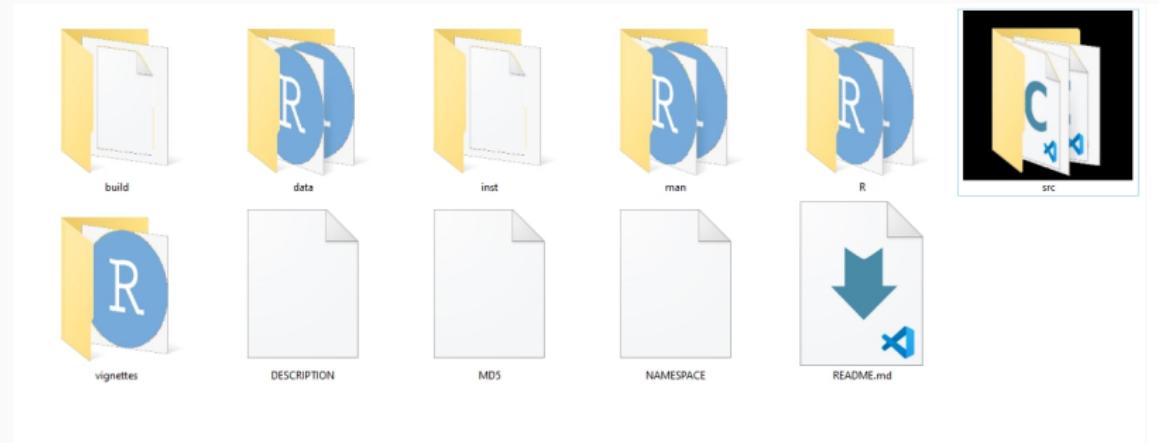
Reverse imports: [betapart](#), [CountClust](#), [DAMOCLES](#), [nodiv](#), [PCPS](#), [pez](#), [recluster](#)

Reverse suggests: [cati](#), [EcoIndR](#), [ecolottery](#), [enveomics.R](#), [fishtree](#), [microeco](#), [phyr](#), [Rarefy](#), [RRphylo](#)

R 程序包的内部结构

名称	大小	压缩后大小	修改时间
build	215	512	2020-06-10 22:53
data	2 765	3 072	2018-05-01 22:47
inst	179 125	180 224	2020-06-10 22:53
man	86 181	97 280	2018-05-01 22:47
R	105 966	114 176	2019-03-21 02:58
src	4 618	5 120	2020-06-10 22:53
vignettes	15 477	15 872	2020-06-10 22:53
DESCRIPTION	1 253	1 536	2020-06-10 23:20
MD5	4 281	4 608	2020-06-10 23:20
NAMESPACE	669	1 024	2018-05-02 01:25
README.md	1 438	1 536	2020-02-13 03:48

C 语言的代码放在 src 文件夹里



Description 文件的内容

```
1 Package: picante
2 Type: Package
3 Title: Integrating Phylogenies and Ecology
4 Version: 1.8.2
5 Date: 2020-06-08
6 Author: Steven W. Kembel <steve.kembel@gmail.com>, David D. Ackerly <dackerly@berkeley.edu>, Samuel Blomberg <s.blomberg1@uq.edu.au>, Will K. Cornwell <cornwell@zoology.ubc.ca>, Peter C. D'Antonio <pdc@berkeley.edu>, Matthew R. Helmus <mrhelmus@wisc.edu>, Helene Morlon <morlon@berkeley.edu>, Campbell O. Webb <cwebb@oeb.harvard.edu>
7 Maintainer: Steven W. Kembel <steve.kembel@gmail.com>
8 Depends: ape, vegan, nlme, methods
9 Suggests: brglm, circular, corpcor, quantreg
10 Description: Functions for phylocom integration, community analyses, null-models. Implements numerous ecophylogenetic approaches including measures of community p
diversity, phylogenetic signal, estimation of trait values for unobserved taxa, community and phylogeny randomizations, and utility functions for data input/out
plotting. A full description of package functionality and methods are provided by the package vignettes. See also the package manual <doi:10.1093/bioinformatics/btq166>.
11 License: GPL-2
12 NeedsCompilation: yes
13 Packaged: 2020-06-10 14:53:27 UTC; steve
14 Repository: CRAN
15 Date/Publication: 2020-06-10 15:20:06 UTC
16
```

加载程序包

```
> library(tidyverse)
-- Attaching packages -----
v ggplot2 3.3.5      v purrr    0.3.4
v tibble   3.1.2      v dplyr    1.0.7
v tidyr    1.1.3      v stringr  1.4.0
v readr    1.4.0      vforcats  0.5.1
-- Conflicts -----
x dplyr::collapse() masks nlme::collapse()
x dplyr::filter()   masks stats::filter()
x dplyr::lag()     masks stats::lag()
> |
```

R 的对象及基本操作

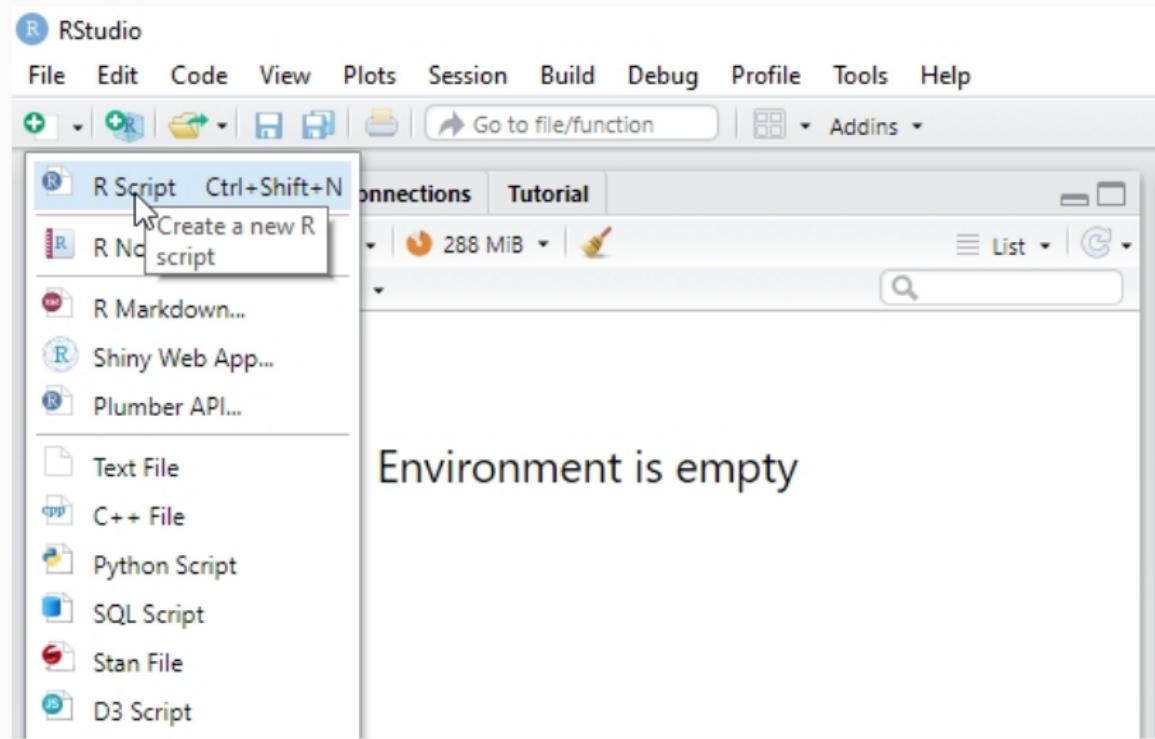
运行代码的方式

- 贴入 R console 运行 (原始社会?)
- 在 Rstudio 的 Console 运行
- 在 Rstudio 的 Source 窗口运行
- 用 CMD 中 Rscript 运行

运行代码的方式

```
> library(tidyverse)
-- Attaching packages -----
v ggplot2 3.3.5      v purrr    0.3.4
v tibble   3.1.2      v dplyr    1.0.7
v tidyr    1.1.3      v stringr  1.4.0
v readr    1.4.0      vforcats  0.5.1
-- Conflicts -----
x dplyr::collapse() masks nlme::collapse()
x dplyr::filter()   masks stats::filter()
x dplyr::lag()     masks stats::lag()
> |
```

在 Rstudio 新建 R 脚本



运行 R 脚本

The screenshot shows the RStudio interface with an R script editor. The code in the editor is:

```
1
2 1
3 1 + 2
4 5 * 5
5 1:10
6 print(1:10)
7
8 ## · 注释
9
10
```

The code consists of several arithmetic operations (addition and multiplication) and a call to the `print` function. Line 8 contains a comment starting with `##`. The RStudio toolbar at the top includes icons for back, forward, source control, search, and run.

R 脚本举例

```
1 setwd("C:/Users/_ /Desktop/shuisong_project/working/modi
2 #·load("ENMeval20210709.RData")·
3 #·load("output_sdm_20210709.RData")·
4 load("output_sdm_20210709_with_buried_ancient_trees.RData")·
5 library(sf)·
6 library(tmap)·
7 library(raster)·
8 ·
9 world·<-·read_sf("./shapefiles/world.shp")·
10 china·<-·read_sf("./shapefiles/bou1_4l.shp")·
11 provinces·<-·read_sf("./shapefiles/province_polygon.shp")·
12 ·
13 show_map·<-·
14 ·function(r,·
15 ······title·=·"",·
16 ······palette·=·NULL,·
17 ······legend.show·=·TRUE,·
```

完整 R 脚本包含的内容

1. 说明此脚本的目的、作者以及日期、版本等
2. 设定工作目录 `setwd()`
3. 加载程序包 `library()`
4. 读取数据 `read.csv()`
5. 数据处理 `lm, glm...`
6. 建立模型
7. 绘图 `plot`
8. 保存数据 `write.csv, save.image()`

R 中的对象及其基本操作

R 中的对象

R 中一切都是对象 (Object)

但是这些对象有一些类似的结构: - 有些类似 Excel 表格的数据 (数据框: data.frame) - 有些类似表格的一列 (向量: vector) - 有些则将多种数据类型整合到一起 (列表 list)。

对数据操作的代码, 如果有一个简单的名字可以调用, 就称为函数 function。

而所有上述的数据, 都是由一些基本的元素组成的 (类似 Excel 的一个单元格)

R 中对象的命名

R 中的对象最好用英文字母开头，不能用数字开头。

四则运算

- +: 加
- -: 减
- *: 乘
- /: 除

```
(2+5)*(15-4.1)
```

```
## [1] 76.3
```

赋值

- 尽量使用 <-
- 设定参数时 =
- 不要用 ->
- 管道符号 |>

赋值符号左右均应留一个空格

元素的类型

- **数值型 Numeric**: 用户不必区分整数和小数, R 会自动判断, 如 12, 5000, 3.1415926
- **字符串型 Character**: 如 "Guangzhou"
- **因子型 Factor**: 如某种实验处理, 增加光照, 不增加光照, 就属于因子类型
- **逻辑型 Logical**: 如导师是否同意报销午餐, 是 TRUE 或否 FALSE
- **复数型 Complex**: 较少用到

其他类型 NaN, NA, NULL

- NA: 缺失值
- NULL: 不存在
- NaN: Not a number 非数值
- Inf: 正无穷

基本数据类型

以上基本元素按照一定规律组合，就构成了 R 中的基本数据类型：

- **向量 vector**: 元素按照一定顺序组合，就构成了向量。注意，单独一个元素也是向量。
- **矩阵 matrix**: 同一种类型数据按照行和列排列。
- **数组 array**: 同一种类型数据，按照二维或者多个维度排列，在 R 中较少用到。
- **数据框 data.frame**: 当多个类型不同，但是长度相同的向量按照列合并，类似于常见的数据记录表。
- **列表 list**: 向量、矩阵、数据框均可放入列表 (list) 中。列表中也可以放列表。

向量的基本操作

向量中的元素必须是相同类型的，如果不相同，一般会自动转换为字符串类型或者逻辑型。

- `c()`: 生成向量，合并向量
- `length()`: 向量的元素个数，或称为长度。
- `paste()`: 合并字符串向量
- `rep()`: 某一个或者几个元素重复指定的次数

```
aaa <- c("a", "b", "c", 1)  
length(aaa)
```

```
## [1] 4
```

```
paste(aaa, collapse = "-")
```

```
## [1] "a-b-c-1"
```

矩阵 matrix 与数组 array

定义 matrix 有两种方式

- `dim()` 查看 matrix 的行列数，也可以用来给向量指定行列，生成 matrix
- `matrix()` 生成 matrix

```
test <- 1:20  
dim(test) <- c(4, 5)
```

查看行列

- `nrow()` 查看行数
- `ncol()` 查看列数

```
dim(test)
```

```
## [1] 4 5
```

```
nrow(test)
```

```
## [1] 4
```

```
ncol(test)
```

```
## [1] 5
```

数据框 data.frame

data.frame(): 生成数据框

cbind(): 将某一向量添加到现有的数据框中

```
library(vegan)
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

```
## This is vegan 2.5-7
```

```
data(dune.env)
```

```
# head(dune.env, 5)
```

数据框 data.frame

```
head(dune.env, 5)
```

	A1	Moisture Management		Use Manure	
## 1	2.8	1	SF Haypastu	4	
## 2	3.5	1	BF Haypastu	2	
## 3	4.3	2	SF Haypastu	4	
## 4	4.2	2	SF Haypastu	4	
## 5	6.3	1	HF Hayfield	2	

查看 data.frame 的属性

- nrow() 查看行数
- ncol() 查看列数
- rownames() 行名
- colnames() 列名提取某一列，用 \$ 加列名即可

在数据框中，各列的长度必须相等

```
dim(dune.env)
```

```
## [1] 20 5
```

```
colnames(dune.env)
```

```
## [1] "A1"           "Moisture"     "Management"   "Use"
```

```
dune.env$Moisture # 提取 Moisture 这一列
```

list 列表

列表中，使用 \$ 提取某一个已经命名的组件，列表中的组件可以是向量、矩阵、数组、数据框以及列表的任意一种。

```
library(vegan)
data(dune.env)
list(dune.env$A1, dune.env$Use)
```

```
## [[1]]
## [1] 2.8 3.5 4.3 4.2 6.3 4.3 2.8 4.2 3.7 3.5
## [16] 5.7 4.0 4.6 3.7 3.5
##
## [[2]]
## [1] Haypastu Haypastu Haypastu Haypastu Hayfield Ha
## [9] Hayfield Hayfield Pasture Haypastu Haypastu 104Pa
## [17] Hayfield Hayfield Hayfield Hayfield
```

列表内元素的命名

```
list(A1 = dune.env$A1, Use = dune.env$Use)
```

```
## $A1
## [1] 2.8 3.5 4.3 4.2 6.3 4.3 2.8 4.2 3.7 3
## [16] 5.7 4.0 4.6 3.7 3.5
##
## $Use
## [1] Haypastu Haypastu Haypastu Haypastu Hayfield Ha
## [9] Hayfield Hayfield Pasture Haypastu Haypastu Pa
## [17] Hayfield Hayfield Hayfield Hayfield
## Levels: Hayfield < Haypastu < Pasture
```

下标 index 与索引 indexing

下标是一个向量，用来提取基本数据类型中的某一些符合条件的值。

- 对于 vector, 使用 []
- 对于 matrix, 使用 [,]
- 对于 data.frame, 使用 [,]
- 对于 list, 使用 [()]

下标 index 与索引 indexing (II)

可以放置整数型或者逻辑型两种向量：

1. 放置 1, 2, 3, 4 等，表示要提取元素的位置
2. 放置 TRUE, FALSE 等，表示是否提取对应的元素

```
letters[1:6]
```

```
## [1] "a" "b" "c" "d" "e" "f"
```

```
ind <- letters %in% c("a", "c", "f")
letters[ind]
```

```
## [1] "a" "c" "f"
```

比较/判断

比较数值的大小，结果均为逻辑向量

- == 判断数值是否相等，是为 TRUE, 否为 FALSE

```
3 == 5
```

```
## [1] FALSE
```

- >: 判断左侧值是否大于右侧值
- <: 判断左侧值是否小于右侧值
- >=: 判断左侧值是否大于或等于右侧值
- <=: 判断左侧值是否小于或等于右侧值
- !=: 判断左侧值是否不等于右侧值

逻辑运算

- `!:` 取反
 - `!TRUE` 结果为 `FALSE`
 - `!FALSE` 结果为 `TRUE`
- `&:` 与, 表示同时发生
 - `TRUE & TRUE` 结果为 `TRUE`
 - `TRUE & FALSE` 结果为 `FALSE`
 - `FALSE & FALSE` 结果为 `FALSE`
- `|:` 或
 - `TRUE|TRUE` 结果为 `TRUE`
 - `TRUE|FALSE` 结果为 `TRUE`
 - `FALSE|FALSE` 结果为 `FALSE`

提取部分数据

使用下标，或者 subset 函数

```
library(vegan)
data(dune.env)
subset(dune.env, Use == "Haypastu")[1:3,]
```

```
##      A1 Moisture Management      Use Manure
## 1  2.8          1           SF Haypastu     4
## 2  3.5          1           BF Haypastu     2
## 3  4.3          2           SF Haypastu     4
```

```
dune.env[dune.env$A1 > 6,]
```

```
##      A1 Moisture Management      Use Manure
## 5  6.3          1           HF Hayfield    2
```

数据类型判断 `is.xxx`

返回结果为 TRUE 或者 FALSE

```
apropos("is\\\.")
is.character() # 是否为字符串?
is.data.frame() # 是否为 data.frame?
is.matrix() # 是否为矩阵?
is.vector() # 是否为向量?
is.list() # 是否为列表?
is.logical() # 是否为逻辑类型?
```

类型转换 as.xxx

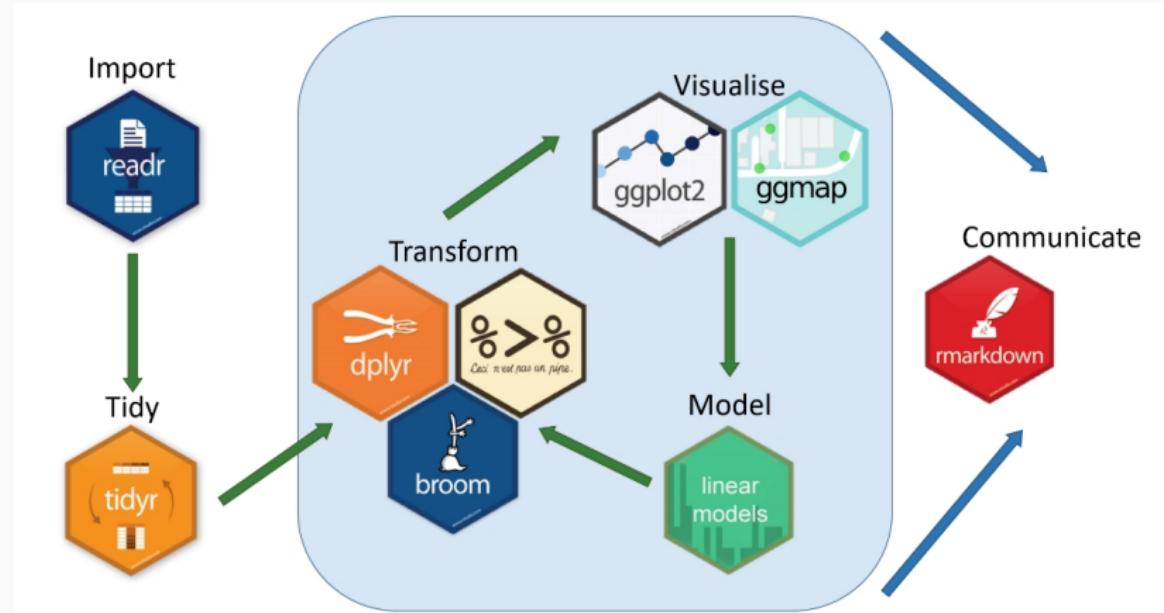
```
apropos("as\\\.")  
as.character() # 转换为字符串  
as.data.frame() # 转换为 data.frame  
as.matrix() # 转换为 matrix  
as.vector() # 转换为 vector  
as.list() # 转换为 list  
as.logical() # 转换为逻辑类型
```

Tidyverse 支持者对此强烈反对

Tidyverse 是什么?



Tidyverse 的数据科学工作流程

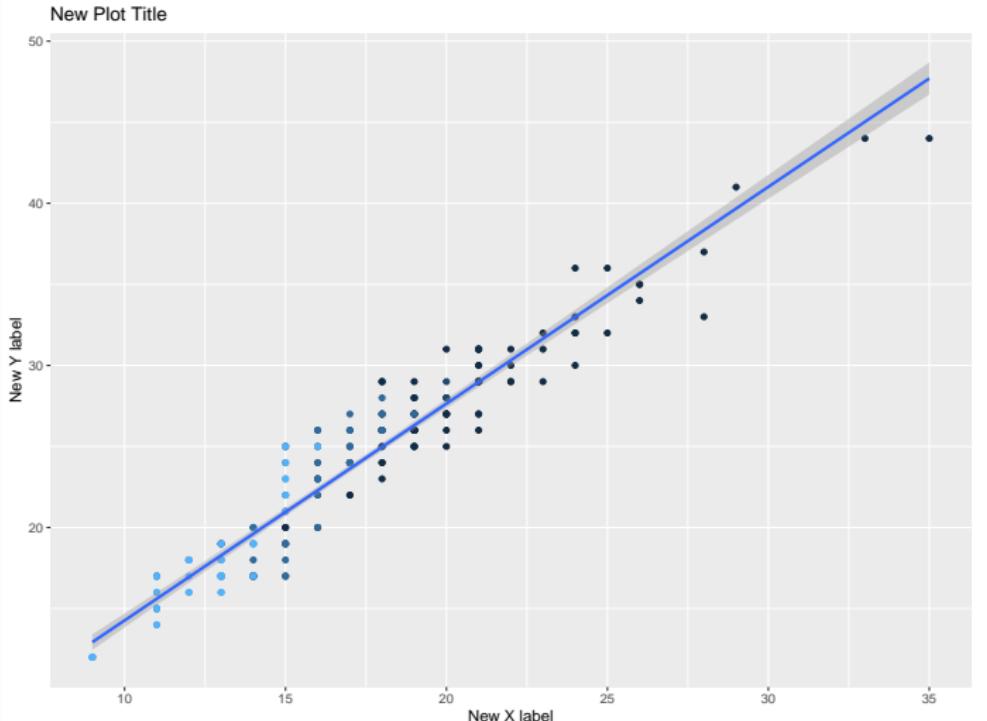


ggplot2 绘图

```
library(ggplot2)
ggplot(data = mpg, aes(x = cty, y = hwy)) + ### 设置 x
  geom_point(aes(color = cyl)) +           ### 添加点
  geom_smooth(method = "lm") +             ### 添加拟合
  coord_cartesian() +                     ### 笛卡尔坐标系
  scale_color_gradient() +                ### 图例的映射
  scale_shape(solid = FALSE) +            ### 点的类型
  ggtitle("New Plot Title") +            ### 标题
  xlab("New X label") +                  ### 横轴
  ylab("New Y label")                   ### 纵轴
```

ggplot2 绘图

```
## `geom_smooth()` using formula 'y ~ x'
```



dplyr 转换为 tibble

```
dplyr::tbl_df(iris)

## Warning: `tbl_df()` was deprecated in dplyr 1.0.0.
## Please use `tibble::as_tibble()` instead.

## # A tibble: 150 x 5
##   Sepal.Length Sepal.Width Petal.Length Petal.Width
##       <dbl>        <dbl>        <dbl>        <dbl>
## 1         5.1        3.5        1.4        0.2
## 2         4.9        3.0        1.4        0.2
## 3         4.7        3.2        1.3        0.2
## 4         4.6        3.1        1.5        0.2
## 5         5.0        3.6        1.4        0.2
## 6         5.4        3.9        1.7        0.4
```

dplyr 选取部分数据

```
dplyr::filter(iris, Sepal.Length > 7)
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
## 1	7.1	3.0	5.9	2.1
## 2	7.6	3.0	6.6	2.1
## 3	7.3	2.9	6.3	1.8
## 4	7.2	3.6	6.1	2.5
## 5	7.7	3.8	6.7	2.2
## 6	7.7	2.6	6.9	2.3
## 7	7.7	2.8	6.7	2.0
## 8	7.2	3.2	6.0	1.8
## 9	7.2	3.0	5.8	1.6
## 10	7.4	2.8	6.1	1.9
## 11	7.9	3.8	6.4	2.0

dplyr 去重复

```
dplyr::distinct(iris)
```

```
##      Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1          5.1        3.5         1.4         0.0
## 2          4.9        3.0         1.4         0.0
## 3          4.7        3.2         1.3         0.0
## 4          4.6        3.1         1.5         0.0
## 5          5.0        3.6         1.4         0.0
## 6          5.4        3.9         1.7         0.0
## 7          4.6        3.4         1.4         0.0
## 8          5.0        3.4         1.5         0.0
## 9          4.4        2.9         1.4         0.0
## 10         4.9        3.1         1.5         0.0
## 11         5.4        3.7         1.5         0.0
```

R 是一种向量化的语言，要尽量避免循环

- 很慢
- 不容易掌握

apply 家族? For 循环? 下标?

Tidyverser 说：打住！

交给 tidyverse 处理吧！

Markdown、RMarkdown 与可重复 研究

Markdown 是什么？

Markdown 是一种简单的标记语言。R markdown 是嵌入了 R 代码的 markdown 语言，可以转换为多种格式。

为什么要发明 Markdown?

Markdown 文档用纯文本书写，通过精心设计的简单字符标记，就能标出文档格式要素。

Markdown 由 John Gruber 设计出来以后，很快受到程序员、科学家、作家的欢迎。与 Markdown 相关的软件与文档也迅速增加。

参考：*

<https://markdown-zh.readthedocs.io/en/latest/> *

<http://wowubuntu.com/markdown/>

Markdown 能干什么？

1. 可重复性研究 Reproducible research

- <https://cran.r-project.org/web/views/ReproducibleResearch.html>
- <https://ropensci.org/>

2. 编写 R 程序包的使用指南。R Markdown 是编写 Vignettes 的首选。

3. 分享文档 <https://rpubs.com/>

4. 制作电子书，博客，个人网站等

- <https://bookdown.org/>
- <https://bookdown.org/yihui/blogdown/>

Markdown 文档中的元素

- 题目：通过 YAML 文件头设置
- 作者：通过 YAML 文件头设置
- 标题：# 一级标题，## 二级标题
- 段落：上下有空行的文字自动成为一段
- 文字引用：每行前面添加 > 作为块引用
- 列表：文字之前加 1. , * , - , + 即可自动识别为列表
- 代码块：自动高亮
- 水平线：连续三个 ***
- 链接：[an example] ([http://example.com/ "Title"](http://example.com/))
- 强调：包围在两个 * , _ , ** , __ 之间，会分别转为斜体或加粗
- 图片：![图片标题] (/path/to/img.jpg)
- 链接：将网址放入 <> 即可自动转换为链接

R Markdown 文档

R Markdown 是 Markdown 语言的进一步发展，借助 knitr 和 rmarkdown 程序包，整合了 R 脚本的 markdown 文档可以很好地编译，并转换为 html、pdf、docx 等格式。

R Markdown 文档中，嵌入的 R 代码，称为 chunk。花括号中可以设置各种参数，控制 R 代码是否显示，是否运行，要显示图片的大小等。

```
1 ---  
2 title: "Untitled"  
3 author: "Jinlong Zhang"  
4 date: "21 June 2018"  
5 output: html_document  
6 ---  
7  
8 ````{r setup, include=FALSE}  
9 knitr::opts_chunk$set(echo = TRUE)  
10 ````
```

图 16: R Markdown 的 trunk

R Markdown 文档的 YAML 文件头设定输出格式、作者与样式等

RMarkdown 有自己的 YAML 表头，控制标题，作者，日期，文档编译的方式和类型以及相应参数等。



R Markdown Reference Guide

Learn more about R Markdown at rmarkdown.rstudio.com
Learn more about Interactive Docs at shiny.rstudio.com/articles

Templates	Basic YAML	Template options	Latex options
html_document	---	---	---
pdf_document	title: "A Web Doc"	title: "Chapters"	title: "My PDF"
word_document	author: "John Doe"	output:	output: pdf_document
md_document	date: "May 1, 2015"	html_document:	fontsize: 11pt
ioslides_presentation	output: md_document	toc: true	geometry: margin=1in
slidy_presentation	---	toc_depth: 2	---
beamer_presentation	---	---	---

图 17: R Markdown 的 YAML 表头

植物学家怎么学习 R?

植物学家怎么学习 R?

1. 安装好 R，搭建好 R 的运行环境
2. 理解数据的读入和 R 中数据的类型以及操作的方法
3. 练习、练习、练习

怎么练习？

1. 看入门书籍，明白计算的原理，并将书中的例子都重复出来
2. 跟着视频学习，学什么？视频怎么组织，怎么讲课
3. 阅读函数的说明文档，编写程序包的文档
4. 阅读别人的代码，特别是脚本，熟悉什么情况用什么方法
5. 自己写代码，注意遵循代码编写的规范

学习资源

- github: R 生态学学习资源
<https://github.com/helixcn/awesome-R-ecology-evolution>
- youtube: R 语言编程合集 (英文)
https://pan.baidu.com/s/1_wyj1sTeb7wp4cWFwk6Vzw 提取码: 6pfj
- bioone 视频系列 (<https://www.bio-one.org.cn/>, 目前正在准备中): 张金龙、胡文浩、崔煜文、李勤、徐洲锋 (Python)期待更多同仁加入
- ecoinformatics 公众号、Y 叔的公众号、生信小白鱼的公众号、庄闪闪的 R 语言手册、R learning、R 语言与藻类生态学.....
- 赖江山博士的博客和培训班

结论

1. R 是一个免费开源的数据分析和绘图的环境，也是一种计算机语言
2. R 能做的内容很多，包括了绝大部分现代进化生物学的分析，掌握了 R 就掌握了海量的科研资源
3. R 的代码能保存工作流程，可让研究更容易重复，这也是当代研究的需要
4. R 简单易学，从今天开始学习 R 并不晚

致谢

- 感谢马老师邀请
- 感谢马老师和李老师组织植物系统学家园微信群作为交流平台
- 感谢各位关注！

问题？
