

R 语言简明教程

第九课、git 与版本控制

张金龙
jinlongzhang01@gmail.com

2022-04-13

《R 语言简明教程》内容安排 (I)

- 第一课、R 语言及其编程环境
- 第二课、数据类型、结构和基础操作
- 第三课、用 graphics 包绘图
- 第四课、用 ggplot2 包绘图
- 第五课、tidydata 以及 tidyverse 系列程序包的使用
- 第六课、基础统计方法的实现
- 第七课、编写 R 函数
- 第八课、Rmarkdown 以及可重复性数据分析
- 第九课、git 与版本控制

《R 语言简明教程》 内容安排（II）

- 第十课、R 程序包的编写、检查与发表
- 第十一课、生物多样性分析常用程序包
- 第十二课、空间数据处理和地图绘制
- 第十三课、系统发育研究的常用程序包
- 第十四课、课程总结

内容提要

Previous Releases of R for Windows

This directory contains previous binary releases of R for Windows.

The current release, and links to development snapshots, are available [here](#). So

In this directory:

[R 4.1.3](#) (March, 2022)
[R 4.1.2](#) (November, 2021)
[R 4.1.1](#) (August, 2021)
[R 4.1.0](#) (May, 2021)
[R 4.0.5](#) (March, 2021)
[R 4.0.4](#) (February, 2021)
[R 4.0.3](#) (October, 2020)
[R 4.0.2](#) (June, 2020)
[R 4.0.1](#) (June, 2020)
[R 4.0.0](#) (April, 2020)
[R 3.6.3](#) (February, 2020)
[R 3.6.2](#) (December, 2019)
[R 3.6.1](#) (July, 2019)
[R 3.6.0](#) (April, 2019)
[R 3.5.3](#) (March, 2019)
[R 3.5.2](#) (December, 2018)
[R 3.5.1](#) (July, 2018)
[R 3.5.0](#) (April, 2018)
[R 3.4.4](#) (March, 2018)
[R 3.4.3](#) (November, 2017)
[R 3.4.2](#) (September, 2017)
[R 3.4.1](#) (June, 2017)
[R 3.4.0](#) (April, 2017)
[R 3.3.3](#) (March, 2017)

图 1: R 的版本

R 程序包的版本

- 每个 R 程序包也有自己的版本
- 引用程序包时，需要指明使用的版本
- 软件的新版本可能增加了新函数，也可能增加了其他新的功能，去掉了一些旧函数等，各版本之间不一定完全兼容

R 程序包的版本

```
citation("vegan")
```

R 程序包的版本

版本号一般由三位组成 $x.y.z$

Bioconductor 对于 R 程序包版本的建议：

- x : 0 表示尚未完全开发好, 1.0.0 一般为首次公布。
- y : 偶数表示正式发布的版本, 奇数表示正在开发的版本。
- z : 新增加了一些功能, 较小的变化。

x,y,z 分别顺序增加

News

- **R version 4.2.0 (Vigorous Calisthenics) prerelease versions** will appear starting Tuesday 2022-03-22. Final release is scheduled for Friday 2022-04-22.
- **R version 4.1.3 (One Push-Up)** has been released on 2022-03-10.
- **R version 4.0.5 (Shake and Throw)** was released on 2021-03-31.

图 2: R 的版本

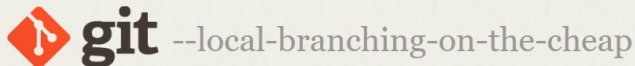
- 对软件源代码进行修改时，编程人员希望记录每一步重要修改。
- 类似于登山时，每向上走一步，都增加一个锚定点，在从这个锚定点出发，继续攀登。
- 如果没有版本控制软件，一旦做出的修改出现致命错误，就无法返回，也无法记录修改的过程。
- 大家熟悉的 MS office 软件，有前进后退，但是对于纯文本文件，如何记录修改历史？

版本控制软件

版本控制软件种类繁多

(https://en.wikipedia.org/wiki/List_of_version_control_software)。

目前在开源软件中，最常用为 git 和 svn。



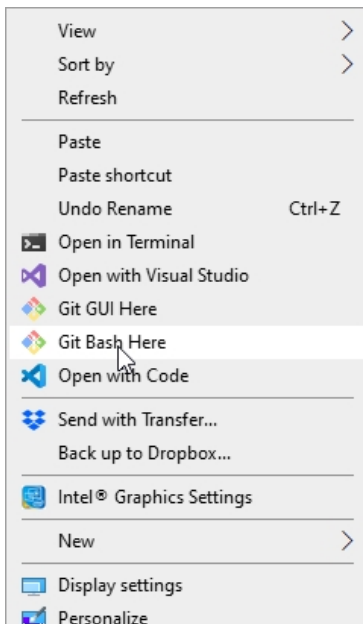
Git is a **free and open source** distributed version control system designed to handle everything from small to very large projects with speed and efficiency.

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**.

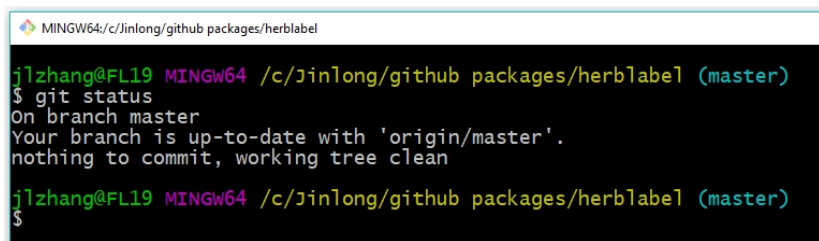
为什么要用 git?

- git 支持多分支同时开发
- git 支持复杂的文件结构和大文件，且十分稳定。在没有网络时，仍然可以工作
- git 软件是开源的，不收取费用
- git 操作检验，很容易学会
- github 网站支持用 git 分享代码和课程资料
- 别人也在用 git
-

git 的基本操作



git 的查看代码的状态（修改前）

A screenshot of a Windows command prompt window. The title bar reads 'MINGW64; c:/jinlong/github/packages/herblabel'. The prompt is 'j1zhang@FL19 MINGW64 /c/jinlong/github/packages/herblabel (master)'. The user enters '\$ git status'. The output is: 'On branch master', 'Your branch is up-to-date with 'origin/master'.', and 'nothing to commit, working tree clean'. The prompt returns to '\$'.

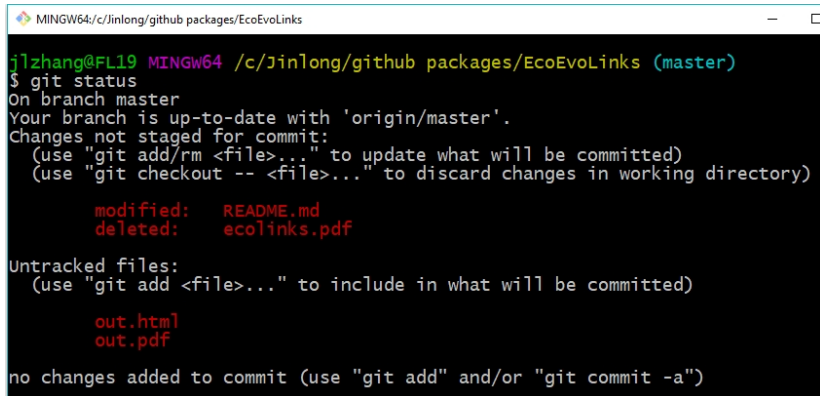
```
MINGW64; c:/jinlong/github/packages/herblabel

j1zhang@FL19 MINGW64 /c/jinlong/github/packages/herblabel (master)
$ git status
On branch master
Your branch is up-to-date with 'origin/master'.
nothing to commit, working tree clean

j1zhang@FL19 MINGW64 /c/jinlong/github/packages/herblabel (master)
$
```

图 5: git 查看代码的状态（修改前）

git 的查看代码的状态（修改后）

A screenshot of a Windows terminal window with a black background and white text. The title bar at the top shows the path 'MINGW64:/c:/Jinlong/github packages/EcoEvoLinks'. The terminal content shows the execution of 'git status' and its output. The output indicates the user is on the 'master' branch, which is up-to-date with 'origin/master'. It lists changes not staged for commit: 'README.md' is modified and 'ecolinks.pdf' is deleted. It also lists untracked files: 'out.html' and 'out.pdf'. At the bottom, it states 'no changes added to commit'.

```
MINGW64:/c:/Jinlong/github packages/EcoEvoLinks
jlzhang@FL19 MINGW64 /c:/Jinlong/github packages/EcoEvoLinks (master)
$ git status
On branch master
Your branch is up-to-date with 'origin/master'.
Changes not staged for commit:
  (use "git add/rm <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working directory)

        modified:   README.md
        deleted:    ecolinks.pdf

Untracked files:
  (use "git add <file>..." to include in what will be committed)

        out.html
        out.pdf

no changes added to commit (use "git add" and/or "git commit -a")
```

图 6: git 查看代码的状态（修改后）

常用 git 命令

- `git clone`: 从某处克隆一个 git repository
- `git status`: 修改之后, 查看当前分支的状态
- `git add`: 添加修改
- `git commit -m "XXXX"`: 提交修改
- `git checkout`: 选择工作分支
- `git rm`: 删除文件
- `git diff`: 显示版本的差异
- `git pull`: 从远程仓库更新代码
- `git push`: 将修改更新到远程仓库

前不久，微软 75 亿美元收购了 github ...



图 7: 微软收购了 github

Github 的主要功能是分享和收藏资源，也可以建立个人网站。对开源项目完全免费。

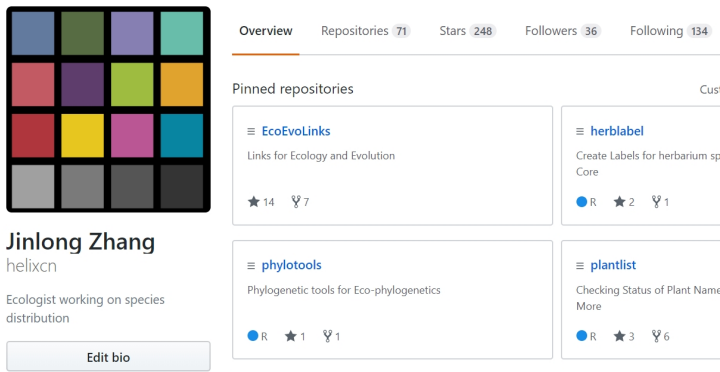


图 8: <https://github.com/helixcn>

github 上分享 R 包源代码

helixcn / herblabel

Unwatch 1 Unstar 2 Fork 1

Code Issues 0 Pull requests 1 Projects 0 Wiki Insights Settings

Create Labels for herbarium specimens based on Darwin Core Edit

Add topics

234 commits 1 branch 0 releases 1 contributor

Branch: master New pull request Create new file Upload files Find file Clone or download

helixcn Add explanation to manual	Latest commit 367b723 on May 10
R	Add the function repeat_rows a month ago
build	build folder added to store Sexpr in Rd files a year ago
data	fill_dwc now automatically update the fields a year ago
inst	citation info added a year ago
man	Add explanation to manual a month ago
DESCRIPTION	Add the function repeat_rows a month ago
NAMESPACE	remove openxlsx as depends 2 years ago
README.md	citation info added a year ago

图 9: herblabel 的源代码

github 上的 R 程序包

- 分享在 github 上的程序包，可以通过 `devtools::install_github()` 安装。
- github 上的版本一般为正在开发的最新版本，用户可以尝试一些新的功能等。
- 程序包可能有很多作者，分别负责一部分。在 github 上，每个人可 fork 不同分支，然后分别开发，之后将不同分支合并 (merge) 到主分支 master 上。
- 通过 git 即可完成上述操作。

R-Forge

R-Forge 是 R 程序包版本管理的网站。

在提交到 CRAN 之前，用户可以在 R-Forge 创建项目，批准后即可上传源代码。

R-Forge 用 SVN 系统管理代码。在 Windows 下可以使用 Tortoise svn 管理。



图 10: R-Forge

R-Forge 上的代码及其安装

packfor

Forward Selection with permutation (Canoco p.46)

Forward Selection as in CANOCO based on permutation procedure using residuals from the regression can be univariate (multiple regression) or multivariate (redundancy analysis).

Version: **0.0-8** | Last change: **2016-06-17 09:22:55+02** | Rev.: **136**

Download:  ([.tar.gz](#)) |  ([.zip](#)) | Build status: **Current**

R install command: `install.packages("packfor", repos="http://R-Forge.R-project.org")`

Show/Hide extra Info

sedarJombart

Thibaut Jombarts contribution to SEDAR package

Various spatial tools originally developed in the adegenet package.

Version: **1.0-0** | Last change: **2015-08-17 15:12:57+02** | Rev.: **131**

Download:  ([.tar.gz](#)) |  ([.zip](#)) | Build status: **Current**

R install command: `install.packages("sedarJombart", repos="http://R-Forge.R-project.org")`

Show/Hide extra Info

图 11: Tortoise SVN



TortoiseSVN

About TortoiseSVN

TortoiseSVN is a really easy to use **Revision control** / version control / so for Windows. It is based on **Apache™ Subversion (SVN)®**; TortoiseSVN pro user interface for Subversion.

It is developed under the **GPL**. Which means it is completely free for any or a commercial environment, without any restriction. The source code is all you can even develop your own version if you wish to.

图 12: R forge 上的程序包安装页

Tortoise SVN 的主要命令

- `svn checkout` 指定要工作的远程库，并获取相应文件
- `svn add` 添加文件到远程库
- `svn commit` 为修改添加相应标签
- `svn update` 将本地文件更新到最新版本
- `svn status` 查看修改状态
- `svn delete` 删除文件
- `svn log` 查看日志
- `svn diff` 比较不同版本
- `svn merge` 合并不同分支
- `svn help` 查看帮助

讲了什么？

- <https://www.bio-one.org.cn/App/VideoClass/play.html?guid=R-in-brief>
- <https://space.bilibili.com/1979221372>

课程材料和源代码:

<https://gitee.com/Hello-BD/R-in-brief>

微信公众号:

ecoinformatics、HelloBD、生物多样性与生态安全大数据

致谢

感谢提出意见的各位老师、同学（以下排名不分次序），特别是：
赖江山、赵宇莹、刘璐妹、余光创、向小果、胡文浩、安瑞志、杨拓、
朱慧玲、乔慧捷、文香英、葛学军、陈建平、曾思金、刘景欣、杨雪飞、
胡小丽、李勤、崔煜文、周通、赵锦等

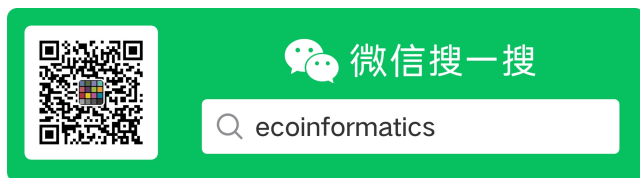


图 13: 敬请关注 `ecoinformatics`

- 如有任何问题、意见或者建议，欢迎联系本人
- Email: `jinlongzhang01@gmail.com`