How to write an R package and publish it on GitHub

Aya Mitani

2021/01/07

What is R package?

- Collection of code, data, documentation developed by R community
- Addresses particular problem with specialized statistical technique, graphical device, etc.
- Core set of packages come with base R
- $>$ 15,000 Additional packages available from CRAN, Bioconductor, Omegahat, GitHub, etc.
- Popular R packages
  - `dplyr`
  - `ggplot2`

What is GitHub?

- Website that hosts software development and version control using Git
- Free basic services
- Truly open source
- “Facebook for programmers”

Why write R package?

- For yourself
    - Save time
        - Keep track of your functions
        - Have all your functions in one place
    - Document your work
    - For publishing papers
        - Increasingly, (bio)statistical journals ask for R package development of novel method
- For others
    - If package is useful to you, it is also useful to someone else
    - Readers of your paper can use proposed method
    - Advance science!

Why publish package on GitHub?

- Reproducibility
- Accessibility
- Collaboration
- Back-up method
- Version-control using git (more on this later)

What do you need to write R package?

- R Studio (https://rstudio.com/)
- `devtools` & `roxygen2` packages
- Git (install from https://git-scm.com/)
- GitHub account (https://github.com/)

Some other useful packages

- `here`
- `available`

Let's begin!

Major steps

1. Open R Studio
   i) New Project → New Directory → R Package → Enter info → Create Project
2. Write functions
   i) Each function should be saved in its own file
   ii) Write package description and document functions
3. Create new repo in GitHub
   i) Repo name = package name
4. Connect to GitHub
5. Pull + Commit + Push
6. Use/share package with `install_github()`!

There is more than one way!

Naming R package

Some tips:

- Make it simple & short
- Make it unique (use `available` package – see next slide)
- Must start with a letter & cannot end with a period
- Do not use special characters
- Trend towards using all lower case
- Hadley's book has more information

Naming R package

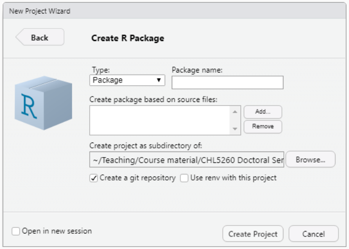
Check to see if name is unique, especially if you plan to submit to CRAN

```r
install.packages("available")
library(available)
```

```r
#available("ayapack", browse = FALSE)
```

In RStudio

- New Project → New Directory → R Package
  - Enter package name
  - Optional: Select R scripts that include your functions (if you leave blank, a default function is included)
  - Select subdirectory where you want to save the package (location is not too important since the final product will be saved on Github)
  - Check 'Create a git repository'



New Project Wizard

Create R Package

Type: Package

Package name:

Create package based on source files:

Create project as subdirectory of: ~/Teaching/Course material/CHL5260 Doctoral Ser

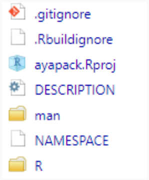
☑ Create a git repository ☐ Use renv with this project

☐ Open in new session

In RStudio

This will create the following files and folders

- packagename.Rproj: This indicates that the directory is a project
- DESCRIPTION: This is where all the meta-data about your package goes – you can edit this file manually
- NAMESPACE: This file indicates what needs to be exposed to users for your R package – we will recreate this file using `document()` (see next slide)
- R: This is where all your R code goes for your package
- man: This is where the manuals for your functions will be saved
- Don't worry too much about the rest (.gitignore, .Rbuildingore)


.gitignore
.Rbuildignore
ayapack.Rproj
DESCRIPTION
man
NAMESPACE
R

In RStudio

Next, load the `devtools` package

```r
library(devtools)
```

Then, delete the NAMESPACE file

Great resources

- Book by Hadley Wickham and Jenny Bryan
- devtools cheatsheet
- Blog post by MZES Social Science Data La

Turn this project into a package

```r
devtools::create(here::here())
```

This will create 3 additional files

- DESCRIPTION: This is where all the meta-data about your package goes. You can edit this file manually.
- NAMESPACE: This file indicates what needs to be exposed to users for your R package. Do not edit this file.
- R: This is where all your R code goes for your package.

Add your function

Open new R script and write your function

```r
myfunc <- function(x){
  y <- x + x
  return(y)
}
```

Add your function

Include `@export` tag above your function to indicate this function to be "exposed" to users.

```r
#' @export
myfunc <- function(x){
  y <- x + x
  return(y)
}
```

Add your function

Also, include documentation for your function when you go `?myfunc`.

```r
#' This is my function.
#'
#' This function returns a value from adding the parameters.
#' @param x
#' @return y
#' @export
myfunc <- function(x){
  y <- x + x
  return(y)
}
```

Add your function

Now run

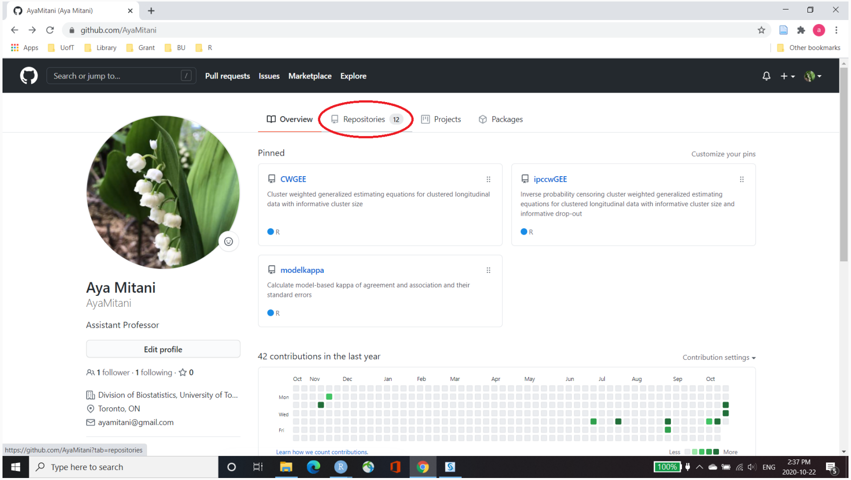
```
devtools::document()
```

- This will create `man` directory that includes read-only file `myfunc.Rd`
- Note that `NAMESPACE` file has been updated

Add data set

In GitHub

First, log in and go to Repositories



Aya Mitani
AyaMitani

Assistant Professor

Edit profile

1 follower · 1 following · 0

Division of Biostatistics, University of To...
Toronto, ON
ayamitani@gmail.com

Pinned

CWGEE
Cluster weighted generalized estimating equations for clustered longitudinal data with informative cluster size
● R

ipccwGEE
Inverse probability censoring cluster weighted generalized estimating equations for clustered longitudinal data with informative cluster size and informative drop-out
● R

modelkappa
Calculate model-based kappa of agreement and association and their standard errors
● R

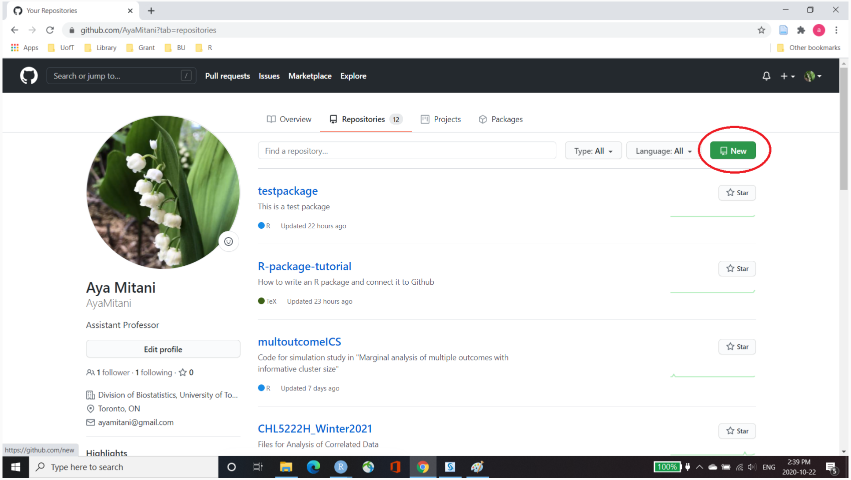
42 contributions in the last year

Learn how we count contributions.

https://github.com/AyaMitani?tab=repositories

In GitHub

Then, create new repository



Aya Mitani
AyaMitani

Assistant Professor

1 follower · 1 following · 0

Division of Biostatistics, University of To...
Toronto, ON
ayamitani@gmail.com

testpackage
This is a test package
R Updated 22 hours ago

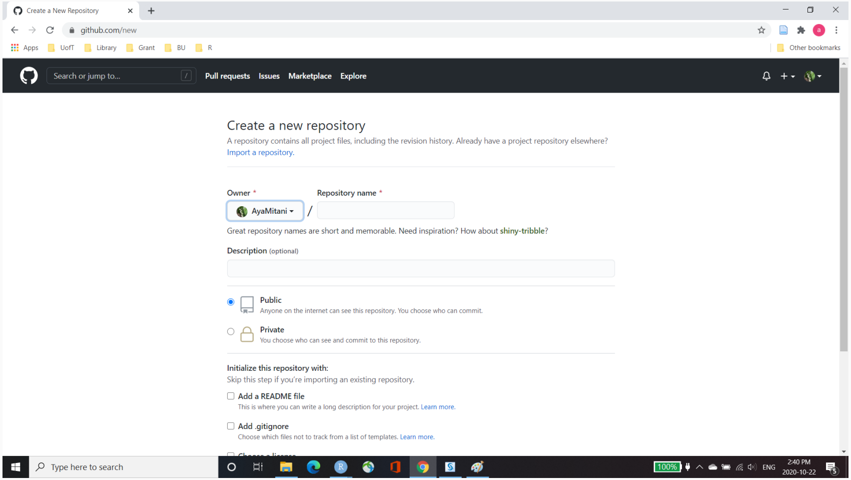
R-package-tutorial
How to write an R package and connect it to Github
TeX Updated 23 hours ago

multoutcomeICS
Code for simulation study in "Marginal analysis of multiple outcomes with informative cluster size"
R Updated 7 days ago

CHL5222H_Winter2021
Files for Analysis of Correlated Data

In GitHub

Repo name should be same as package name



Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? Import a repository.

Owner * AyaMitani / Repository name *

Great repository names are short and memorable. Need inspiration? How about shiny-tribble?

Description (optional)

● Public
Anyone on the internet can see this repository. You choose who can commit.

○ Private
You choose who can see and commit to this repository.

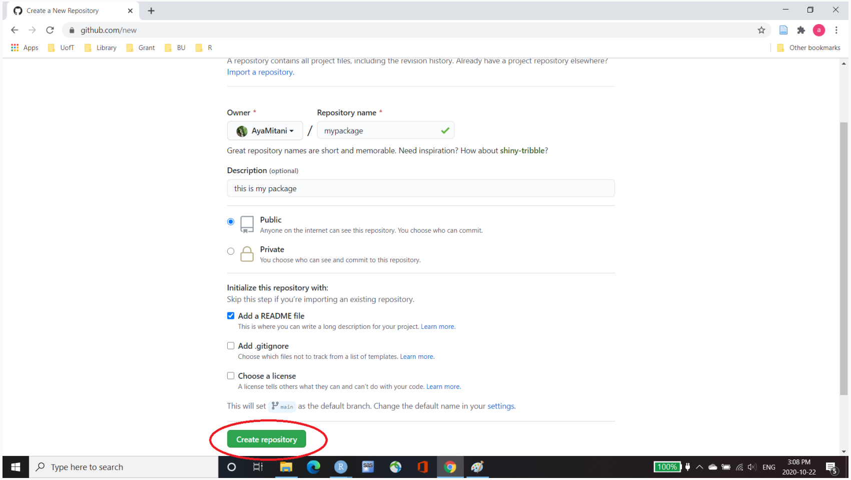
Initialize this repository with:
Skip this step if you're importing an existing repository.

☐ Add a README file
This is where you can write a long description for your project. Learn more.

☐ Add .gitignore
Choose which files not to track from a list of templates. Learn more.

In GitHub

Finish creating new repo


A repository contains all project files, including the revision history. Already have a project repository elsewhere?
Import a repository.

Owner * AyaMitani / Repository name * mypackage

Great repository names are short and memorable. Need inspiration? How about shiny-tribble?

Description (optional)
this is my package

(●) Public
Anyone on the internet can see this repository. You choose who can commit.

( ) Private
You choose who can see and commit to this repository.

Initialize this repository with:
Skip this step if you're importing an existing repository.

[✓] Add a README file
This is where you can write a long description for your project. Learn more.

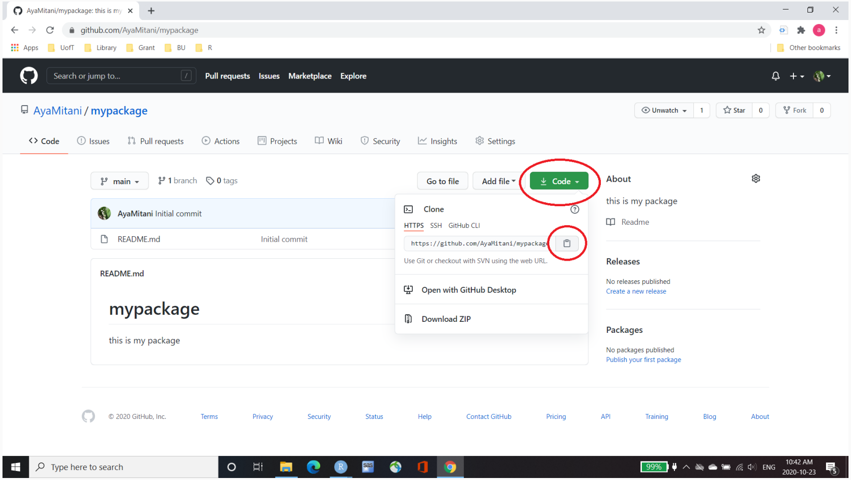
[ ] Add .gitignore
Choose which files not to track from a list of templates. Learn more.

[ ] Choose a license
A license tells others what they can and can't do with your code. Learn more.

This will set main as the default branch. Change the default name in your settings.

In GitHub

Finally, copy URL


AyaMitani/mypackage: this is my
github.com/AyaMitani/mypackage
Apps
UofT
Library
Grant
BU
R
Other bookmarks
Search or jump to...
Pull requests
Issues
Marketplace
Explore
AyaMitani / mypackage
Unwatch 1
Star 0
Fork 0
Code
Issues
Pull requests
Actions
Projects
Wiki
Security
Insights
Settings
main
1 branch
0 tags
Go to file
Add file
Code
AyaMitani Initial commit
README.md
Initial commit
Clone
HTTPS
SSH
GitHub CLI
https://github.com/AyaMitani/mypackage
Use Git or checkout with SVN using the web URL.
Open with GitHub Desktop
Download ZIP
About
this is my package
Readme
Releases
No releases published
Create a new release
Packages
No packages published
Publish your first package
README.md
mypackage
this is my package
© 2020 GitHub, Inc.
Terms
Privacy
Security
Status
Help
Contact GitHub
Pricing
API
Training
Blog
About
Type here to search
99%
ENG
10:42 AM
2020-10-23