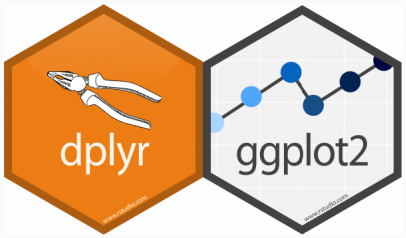
How to write an R package and publish it on GitHub

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

Why write R package?

- For yourself
- For others
  - Increasingly, (bio)statistical journals ask for R package development of novel method

Why publish package on GitHub?

- Reproducibility
- Accessibility
- Collaboration
- Back-up method

What do you need to write R package?

- R Studio (https://rstudio.com/)
- `devtools` package
- Git (https://git-scm.com/)
- GitHub account (https://github.com/)

Some other useful packages

- `here`
- `available`

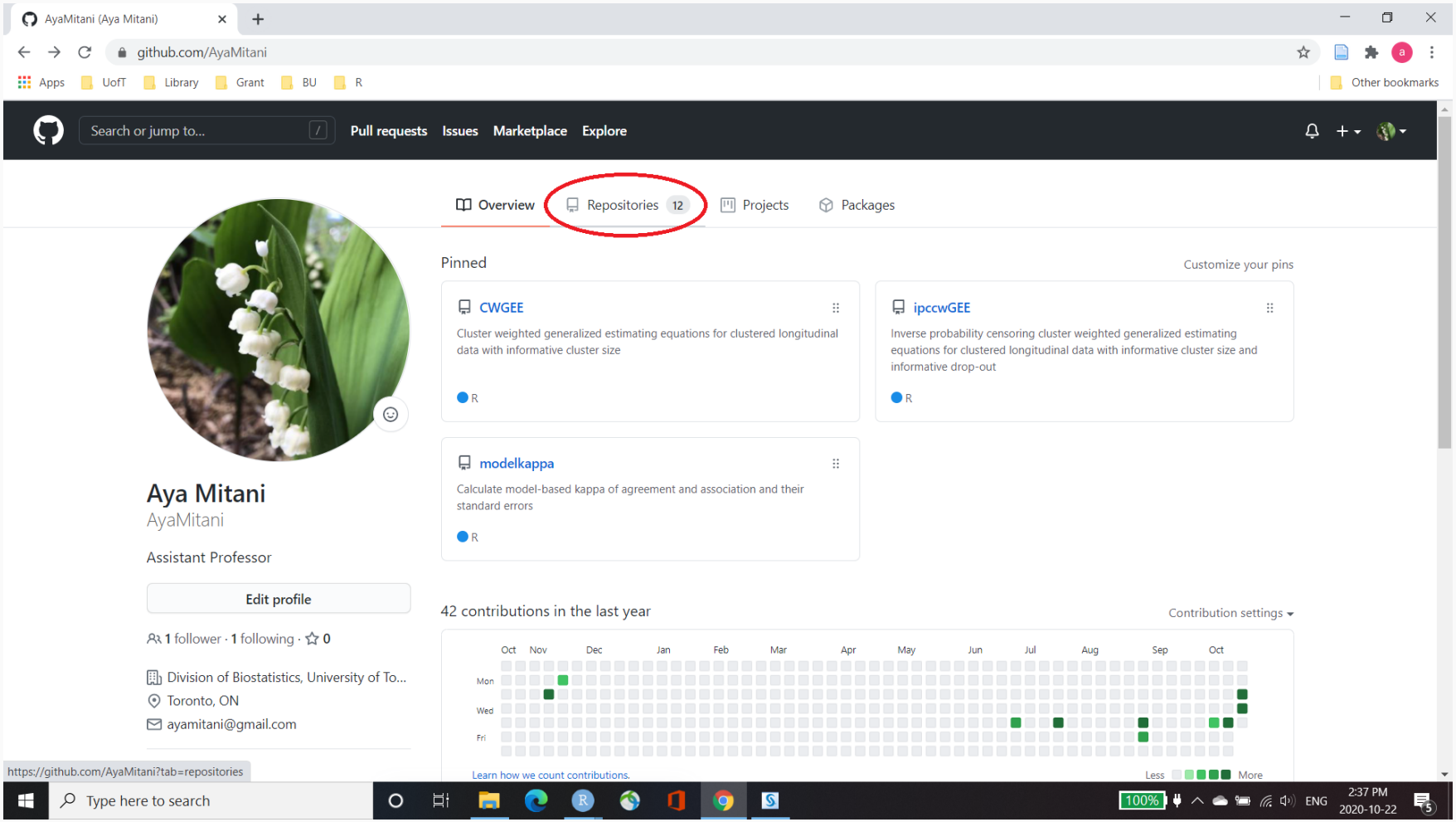
Let's begin!

Major steps

1. Create new repo in GitHub
    i) Copy URL
2. Open R Studio
    i) New Project → Version Control → Git → Enter info → Create Project
    ii) Install `devtools`
    iii) Build package
3. Pull + Commit + Push
4. Use/share package with `install_github()`!

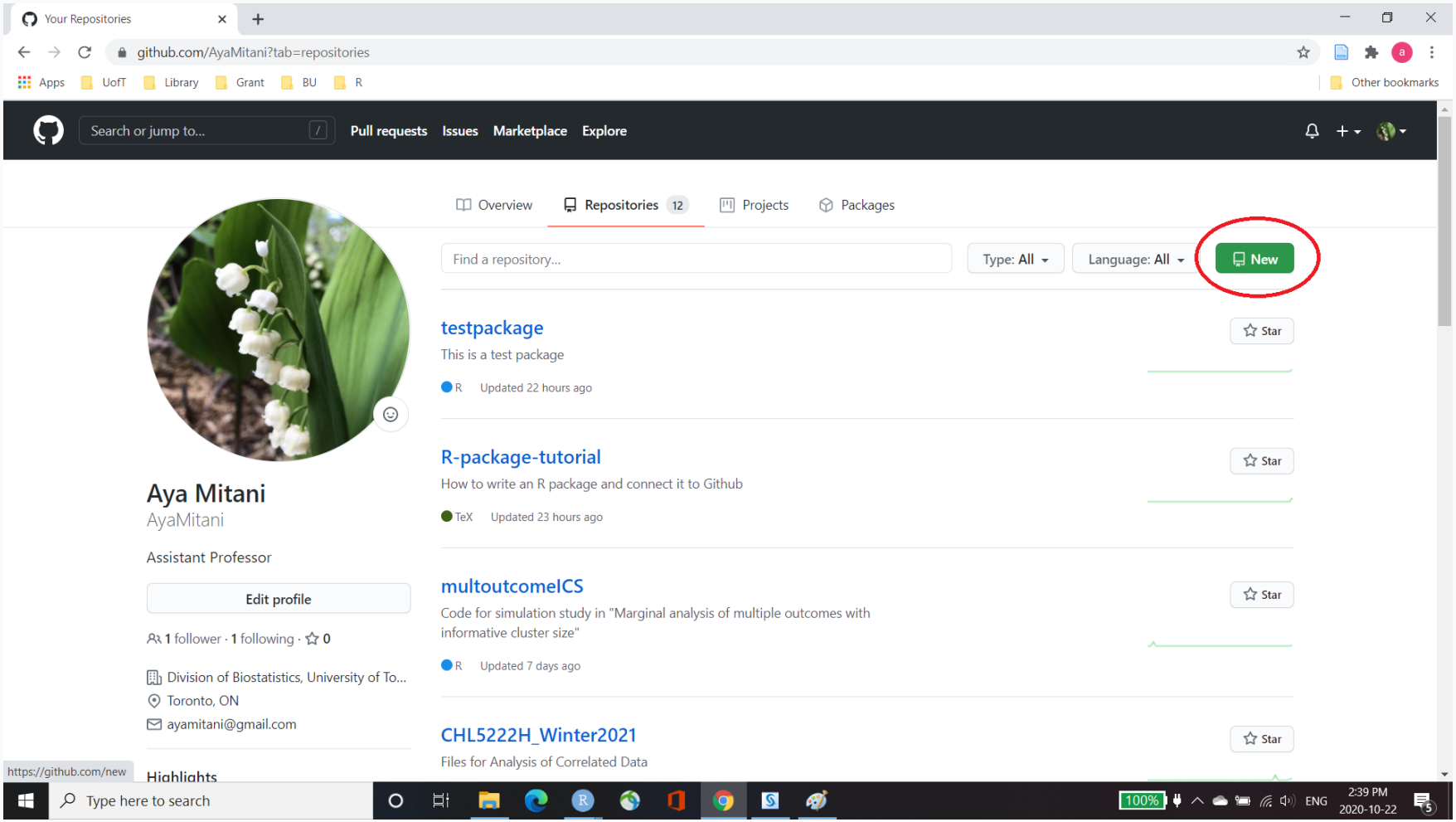
In GitHub

First, log in and go to Repositories



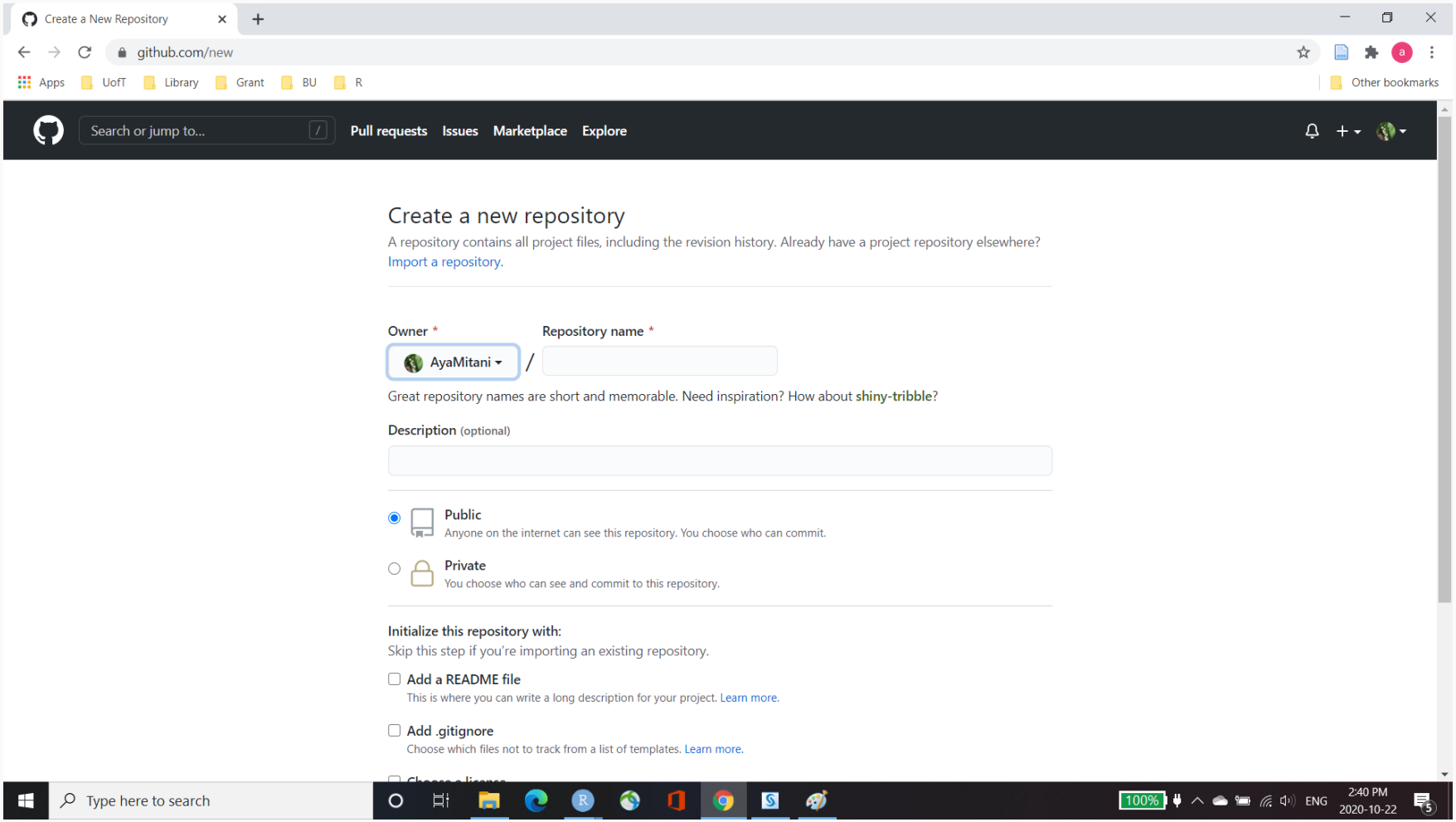
In GitHub

Then, create new repository



In GitHub

Repo name should be same as package name



Aside: Naming R package

Some tips:

- Make it short
- If one word, use all lowercase
- If multiple words, mix lowercase and uppercase
- Avoid using characters (., _) between words

Aside: Naming R package

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

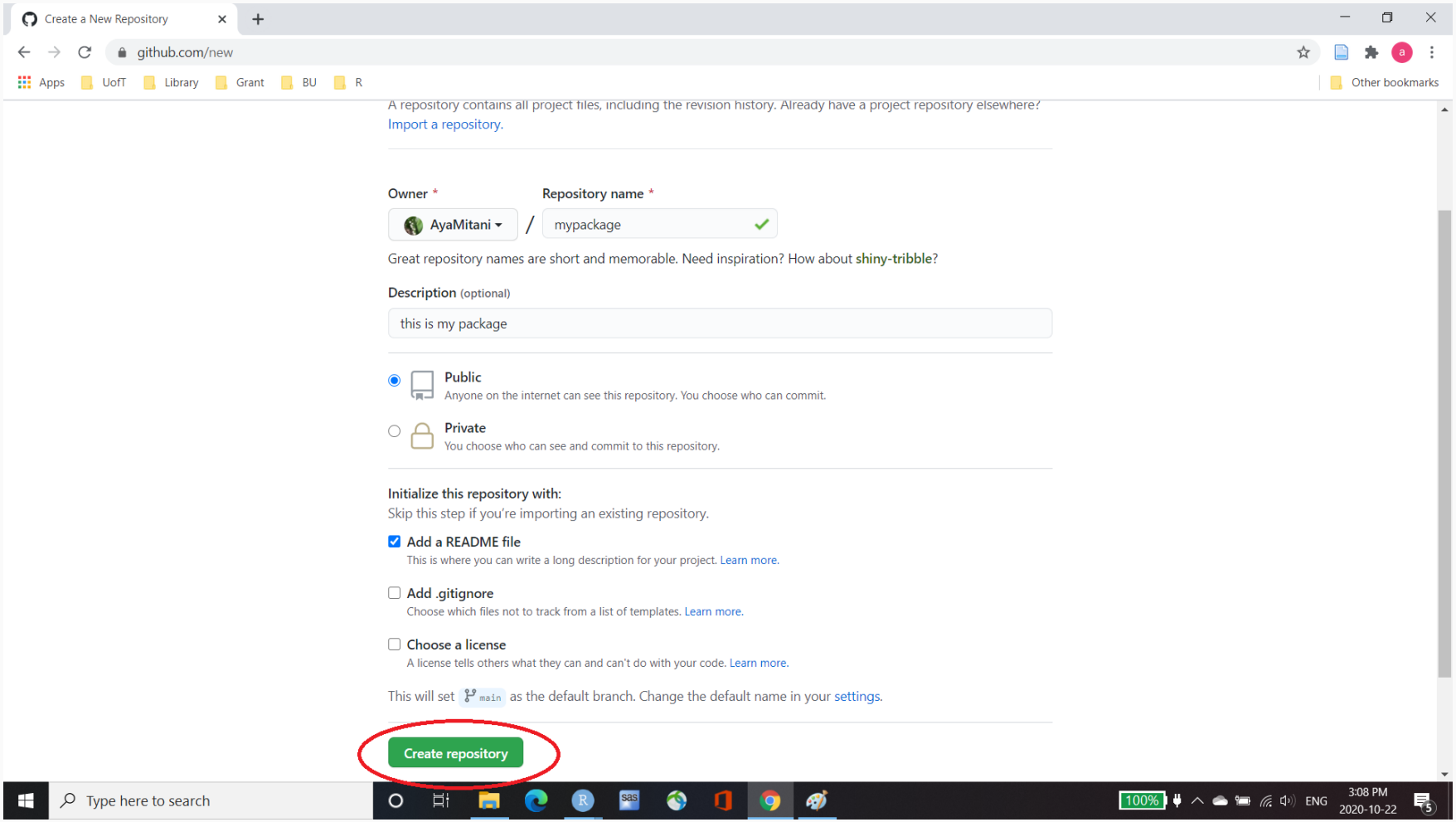
```
install.packages("available")
library(available)
```

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

```
## -- ayapack ---------------------------------------------------
## Name valid: <U+2714>
## Available on CRAN: <U+2714>
## Available on Bioconductor: <U+2714>
## Available on GitHub:  <U+2714>
## Abbreviations: http://www.abbreviations.com/ayapack
## Wikipedia: https://en.wikipedia.org/wiki/ayapack
## Wiktionary: https://en.wiktionary.org/wiki/ayapack
## Urban Dictionary:
##   Not found.
## Sentiment:???
```

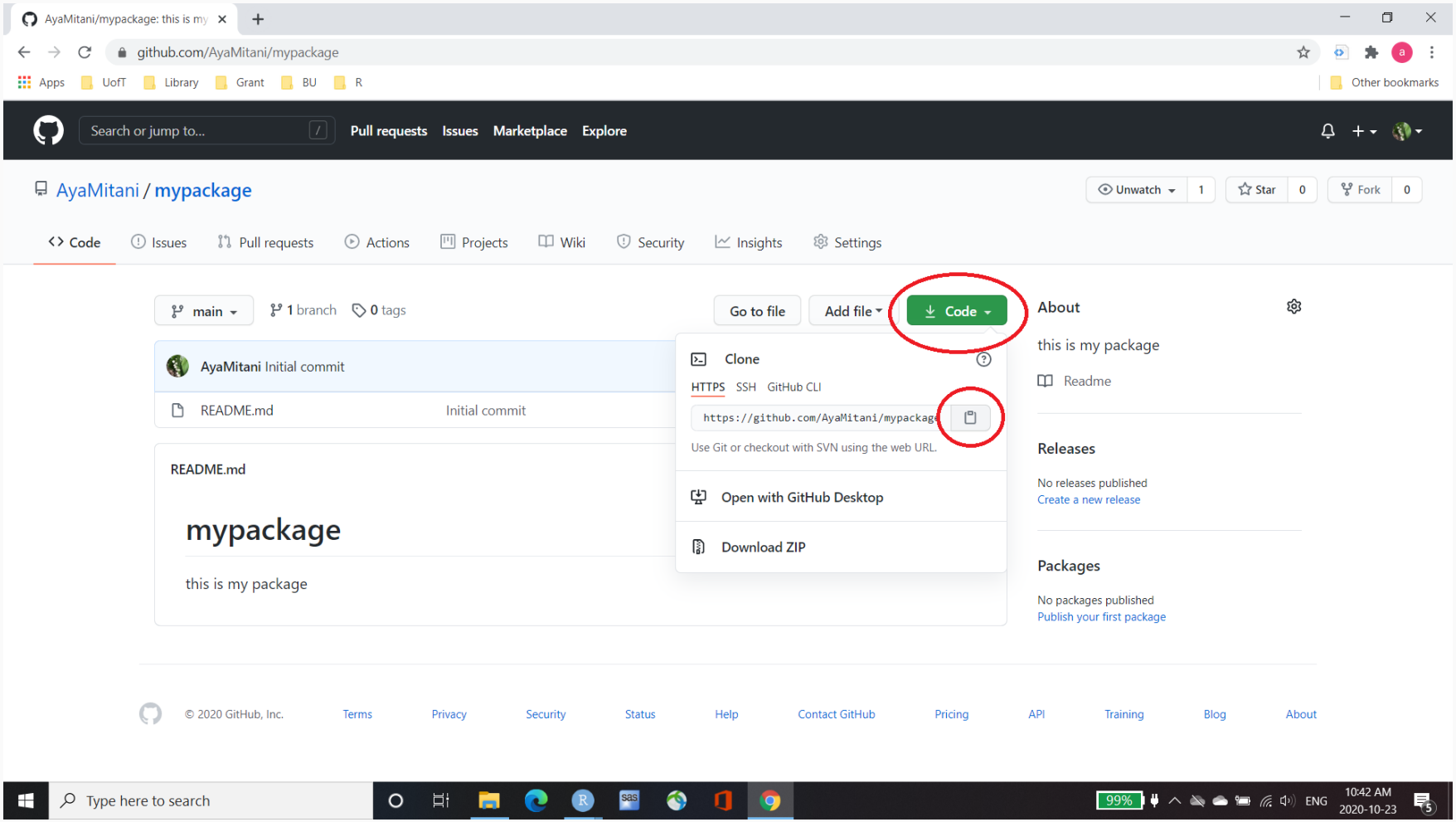
Back to GitHub

Finish creating new repo



In GitHub

Finally, copy URL



In RStudio

Load libraries

```
library(devtools)
library(here)
```

Turn this project into a package

```
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

```
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.

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

Add your function

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

```
#' 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