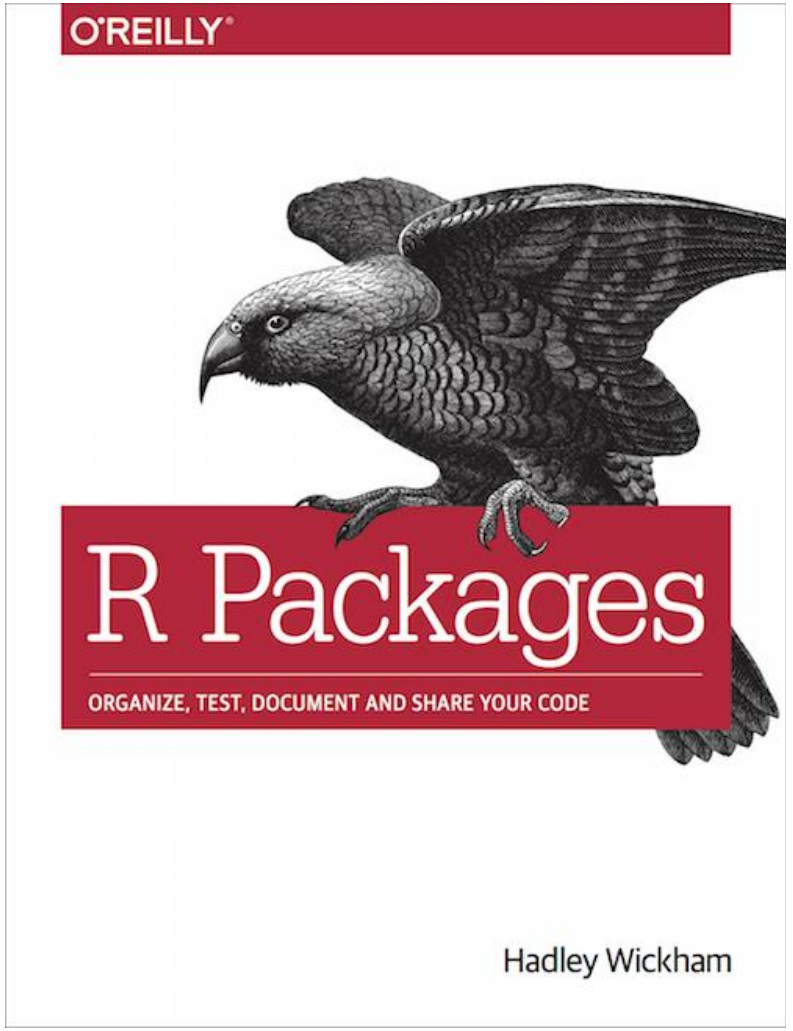
Writing R Packages from scratch using RStudio & roxygen2.

Susan Johnston, University of Edinburgh

EdinbR Meeting, 17th Jan 2018

Why write an R package?

- “Packages are the fundamental units of reproducible R code”
- Save yourself time with reusable R functions.
- Facilitate code sharing with others.
- **It does not have to go on CRAN.**



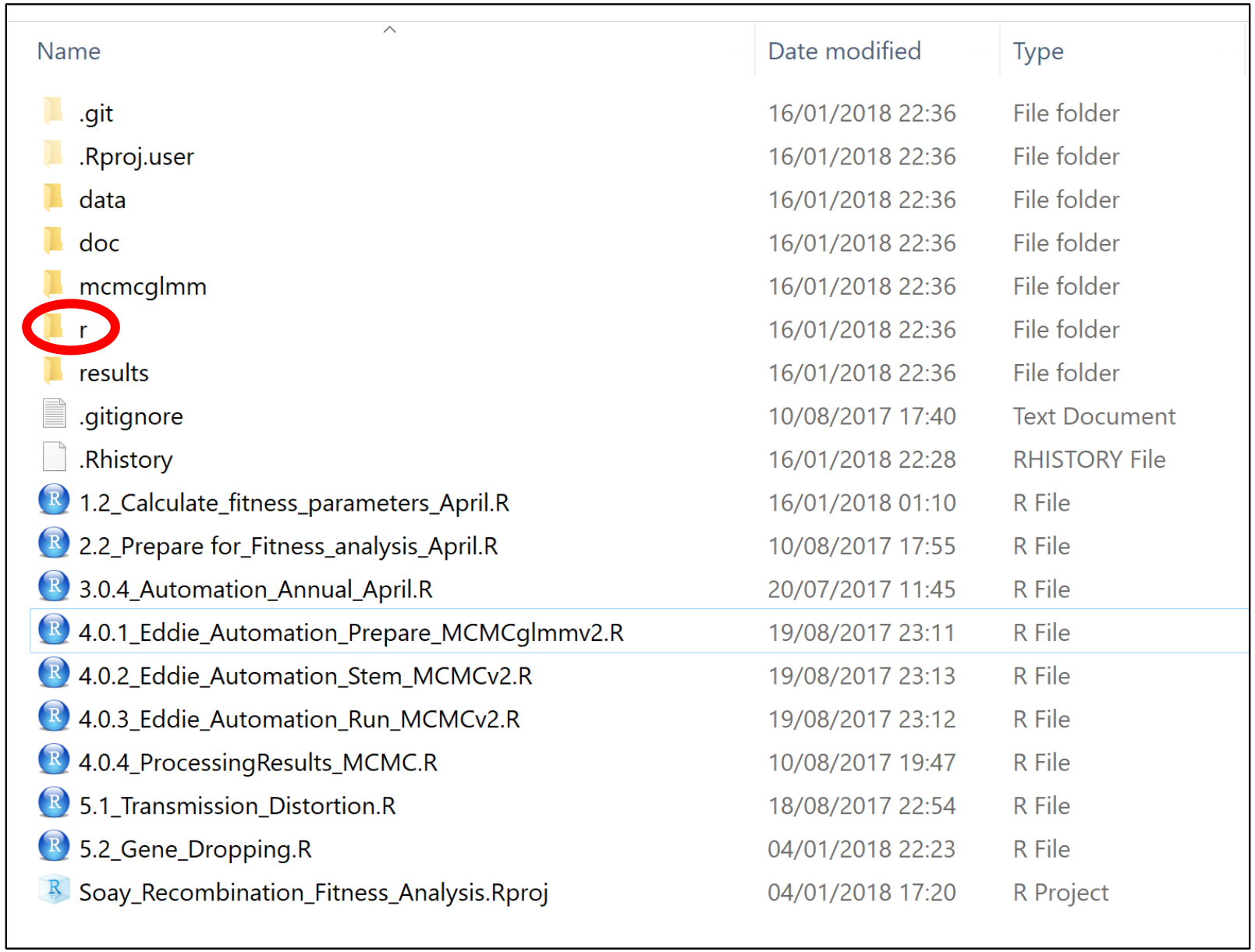
Who can write an R package?

- Can you open R or RStudio?
- Can you install a package?
- Have you ever written a function in R?
- Could you *learn* how to write a function in R?
- **You can write a package in R.**

Structuring an R Project.

http://nicercode.github.io/blog/2013-05-17-organising-my-project/

An example of an average R project. R functions are stored in the directory r/

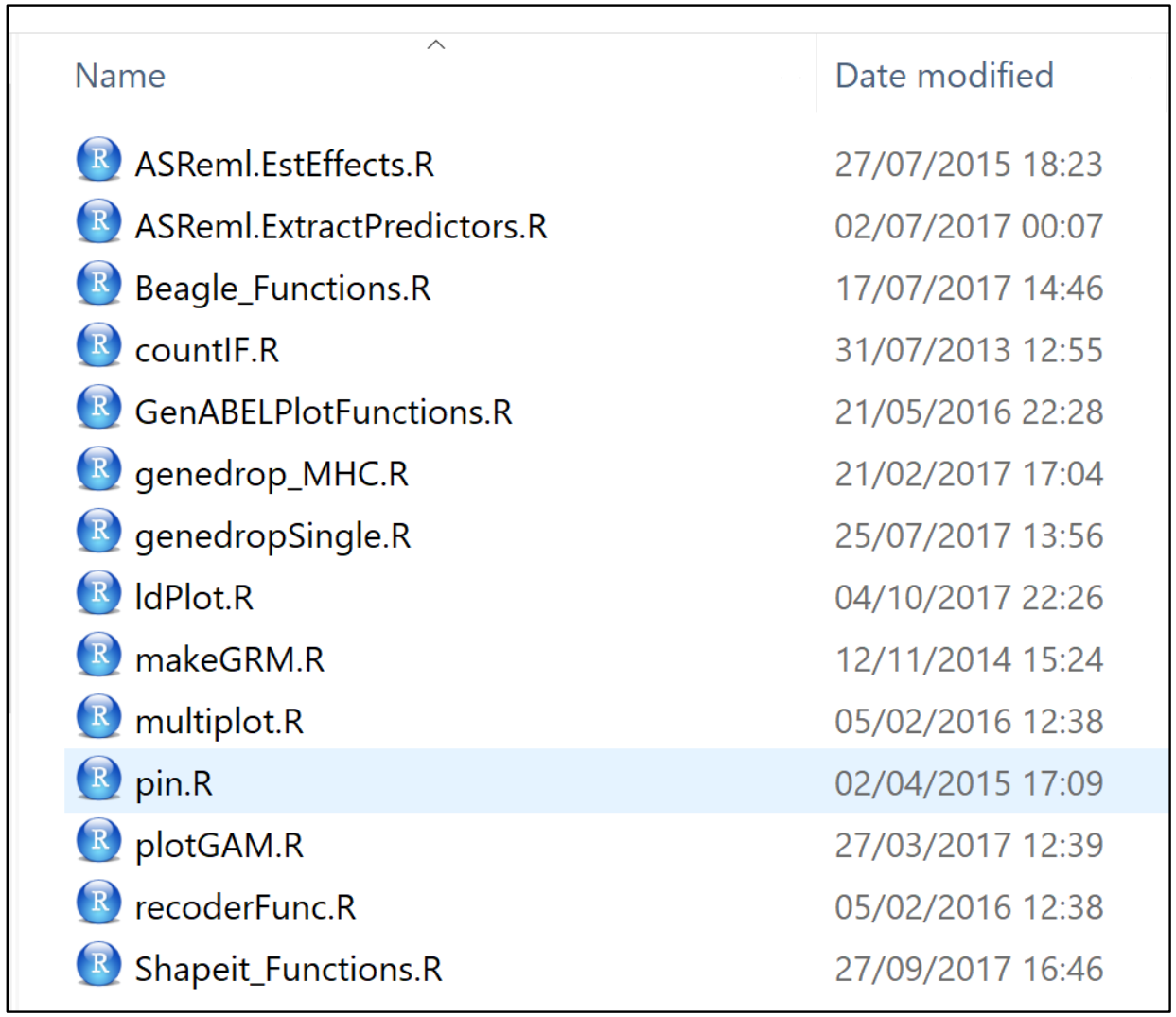


Name	Date modified	Type
.git	16/01/2018 22:36	File folder
.Rproj.user	16/01/2018 22:36	File folder
data	16/01/2018 22:36	File folder
doc	16/01/2018 22:36	File folder
mcmcglmm	16/01/2018 22:36	File folder
r	16/01/2018 22:36	File folder
results	16/01/2018 22:36	File folder
.gitignore	10/08/2017 17:40	Text Document
.Rhistory	16/01/2018 22:28	RHISTORY File
1.2_Calculate_fitness_parameters_April.R	16/01/2018 01:10	R File
2.2_Prepare for_Fitness_analysis_April.R	10/08/2017 17:55	R File
3.0.4_Automation_Annual_April.R	20/07/2017 11:45	R File
4.0.1_Eddie_Automation_Prepare_MCMCglmmv2.R	19/08/2017 23:11	R File
4.0.2_Eddie_Automation_Stem_MCMCv2.R	19/08/2017 23:13	R File
4.0.3_Eddie_Automation_Run_MCMCv2.R	19/08/2017 23:12	R File
4.0.4_ProcessingResults_MCMC.R	10/08/2017 19:47	R File
5.1_Transmission_Distortion.R	18/08/2017 22:54	R File
5.2_Gene_Dropping.R	04/01/2018 22:23	R File
Soay_Recombination_Fitness_Analysis.Rproj	04/01/2018 17:20	R Project

Structuring an R Project.

http://nicercode.github.io/blog/2013-05-17-organising-my-project/

Generally contains a bunch of functions that are read into R scripts using `source()`



Name	Date modified
ASReml.EstEffects.R	27/07/2015 18:23
ASReml.ExtractPredictors.R	02/07/2017 00:07
Beagle_Functions.R	17/07/2017 14:46
countIF.R	31/07/2013 12:55
GenABELPlotFunctions.R	21/05/2016 22:28
genedrop_MHC.R	21/02/2017 17:04
genedropSingle.R	25/07/2017 13:56
ldPlot.R	04/10/2017 22:26
makeGRM.R	12/11/2014 15:24
multiplot.R	05/02/2016 12:38
pin.R	02/04/2015 17:09
plotGAM.R	27/03/2017 12:39
recoderFunc.R	05/02/2016 12:38
Shapeit_Functions.R	27/09/2017 16:46

A better way: Functions can be bundled into R Packages.

Writing an R package the simple way

- Install RStudio

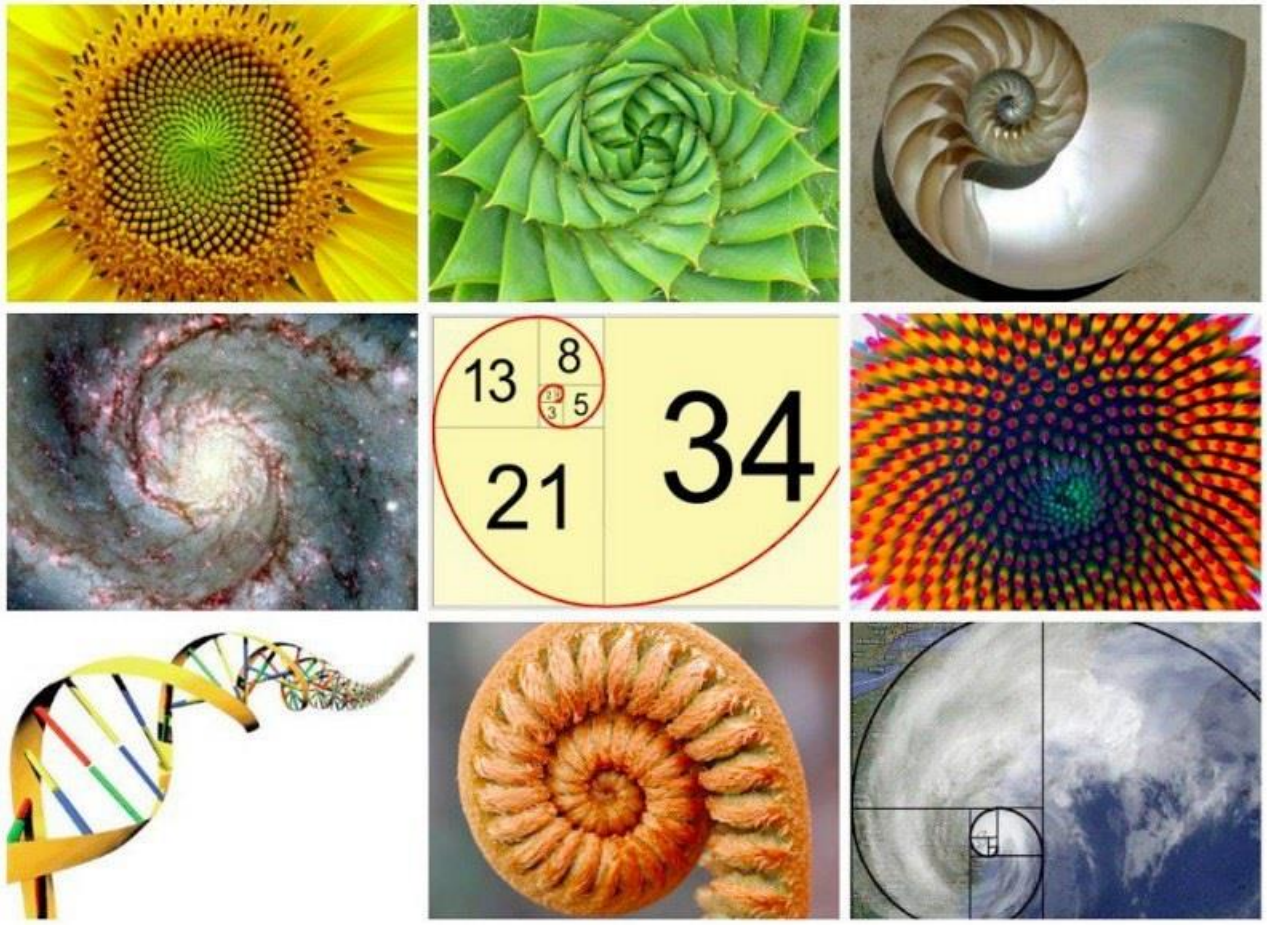
  Create packages using .Rproj system.

  Alternatively use `devtools::create()`

- Install the package `roxygen2`

  Used to create documentation for R packages

The `fibonacci` package.



The **Fibonacci sequence** is characterised by the fact that every number after the first two is the sum of the two preceding ones.

Common in biological settings, such as petal number, tree branching and shell shapes.

https://github.com/susjoh/fibonacci

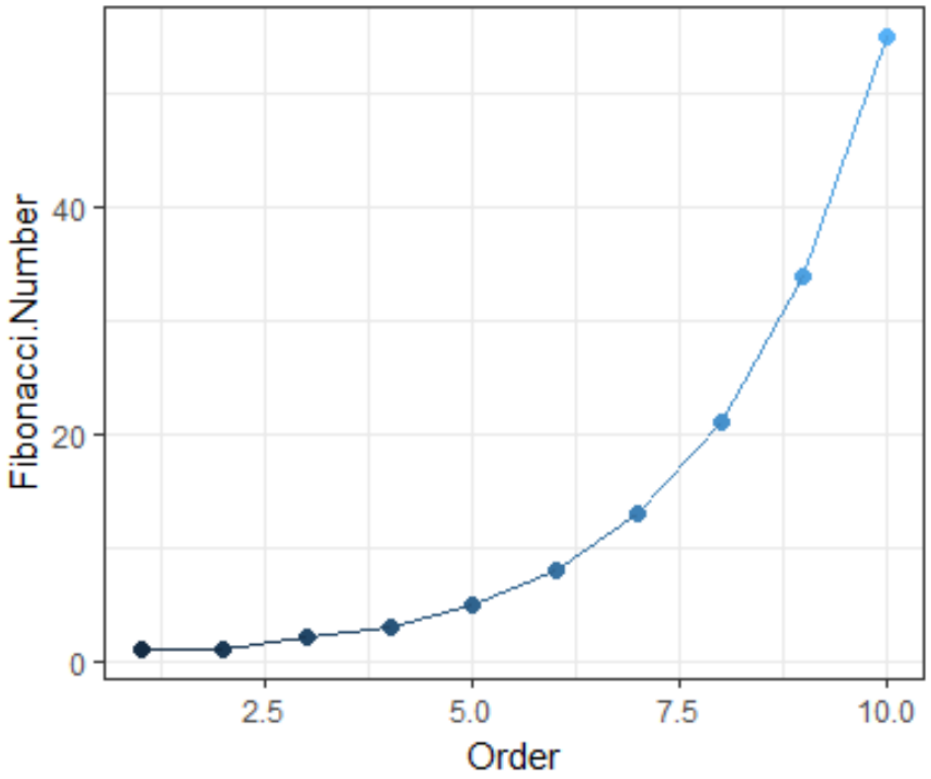
fibonacci.R : Generates a Fibonacci sequence.

```
1
2  fibonacci <- function(len){
3
4    fibvec <- numeric(len)
5
6    fibvec[1] <- 1
7    fibvec[2] <- 1
8    for (i in 3:len) {
9      fibvec[i] <- fibvec[i-1]+fibvec[i-2]
10   }
11
12   fibvec
13 }
14
```

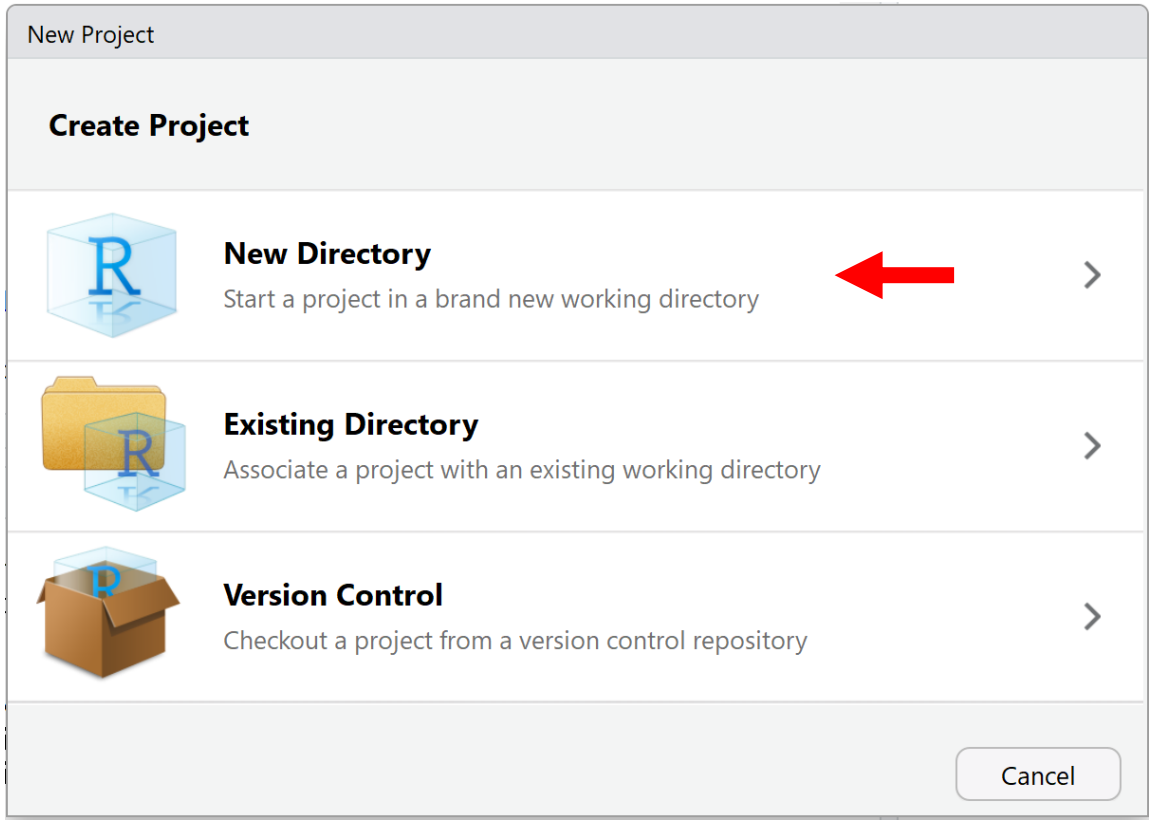
```
> source("R/fibonacci.R")
> fibonacci(10)
 [1]  1  1  2  3  5  8 13 21 34 55
```

plot_fibonacci.R : Plots a Fibonacci sequence with colour gradient.

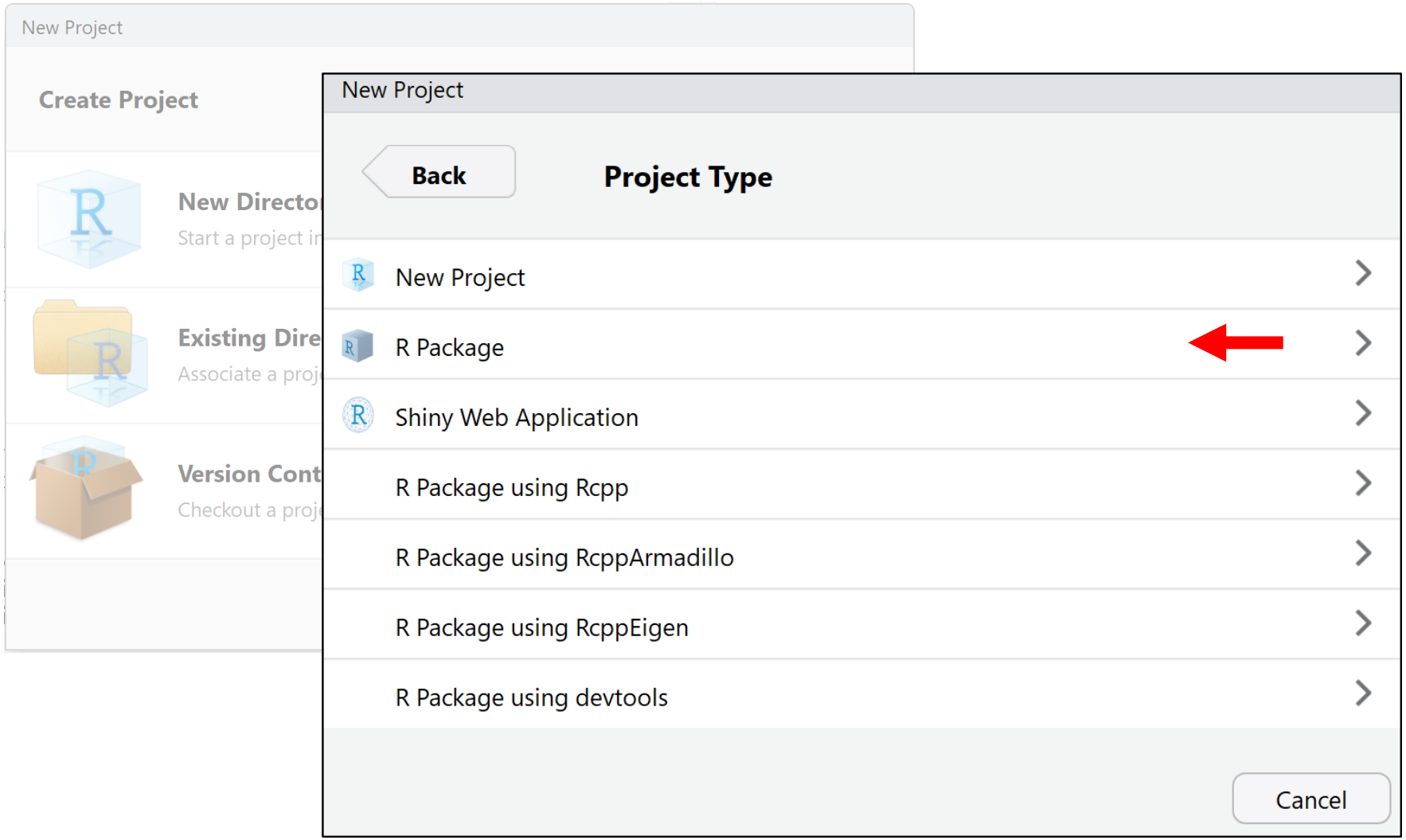
```r

plot_fibonacci <- function(len, gradient = T){

  require(ggplot2)

  x <- data.frame(Order = 1:len,
                  Fibonacci.Number = fibonacci(len))

  if(gradient){
    p <- ggplot(x, aes(Order, Fibonacci.Number, colour = Order))
  } else {
    p <- ggplot(x, aes(Order, Fibonacci.Number))
  }

  p +
    geom_point(size = 2) +
    geom_line() +
    theme_bw() +
    theme(legend.position = "none")

}

```



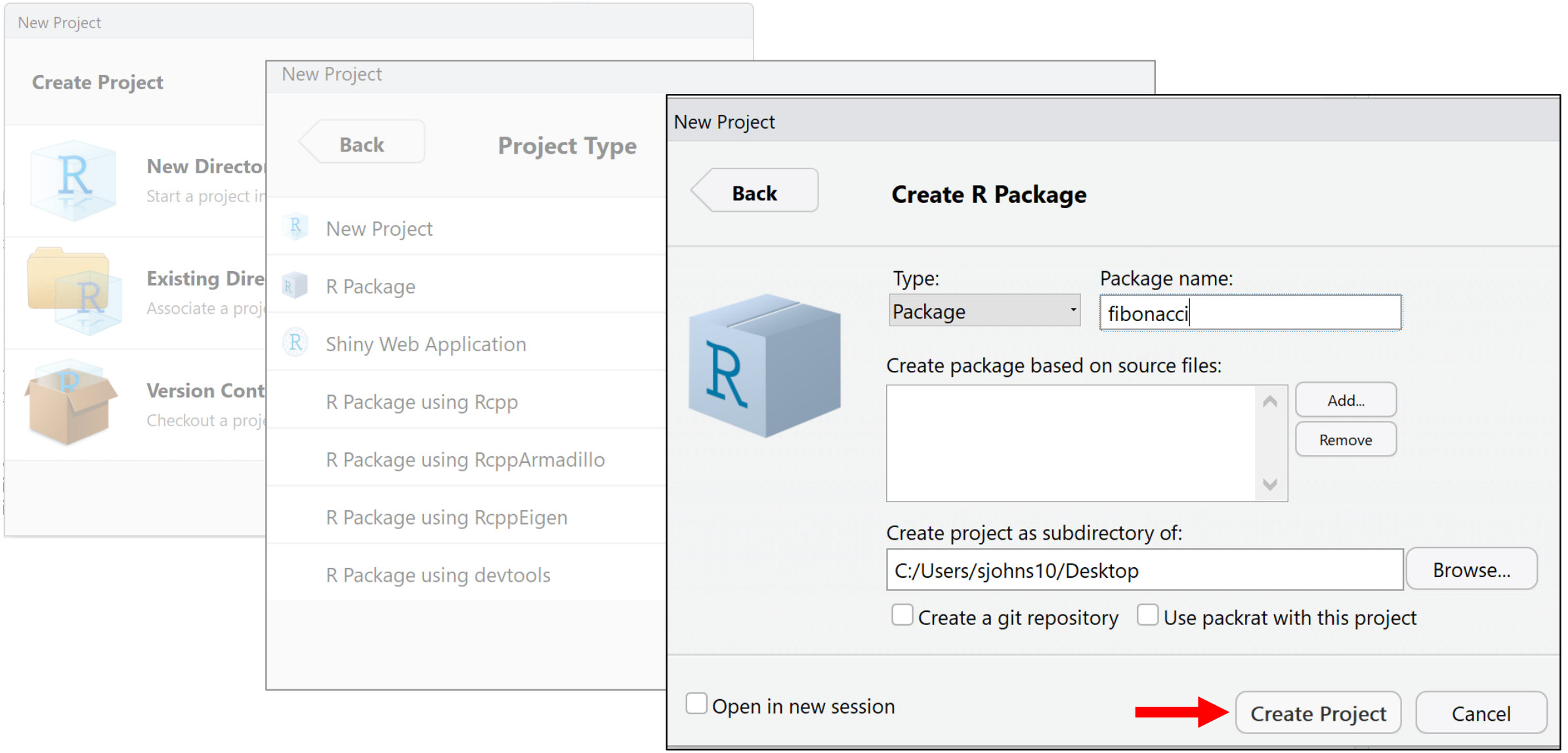
Creating a new R Package in Rstudio.



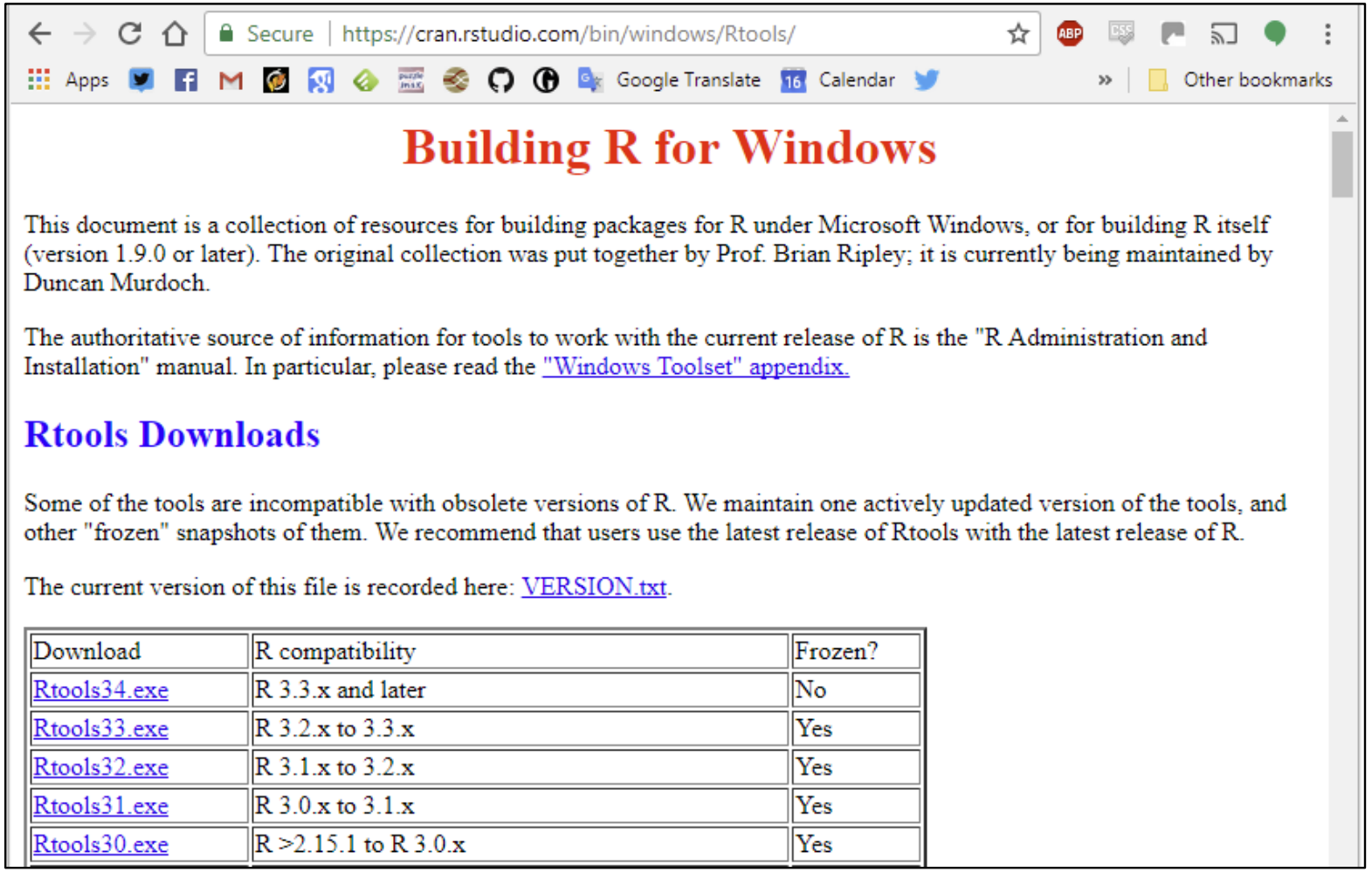
Creating a new R Package in Rstudio.



Creating a new R Package in Rstudio.



Building packages in Windows requires Rtools



Building R for Windows

This document is a collection of resources for building packages for R under Microsoft Windows, or for building R itself (version 1.9.0 or later). The original collection was put together by Prof. Brian Ripley; it is currently being maintained by Duncan Murdoch.

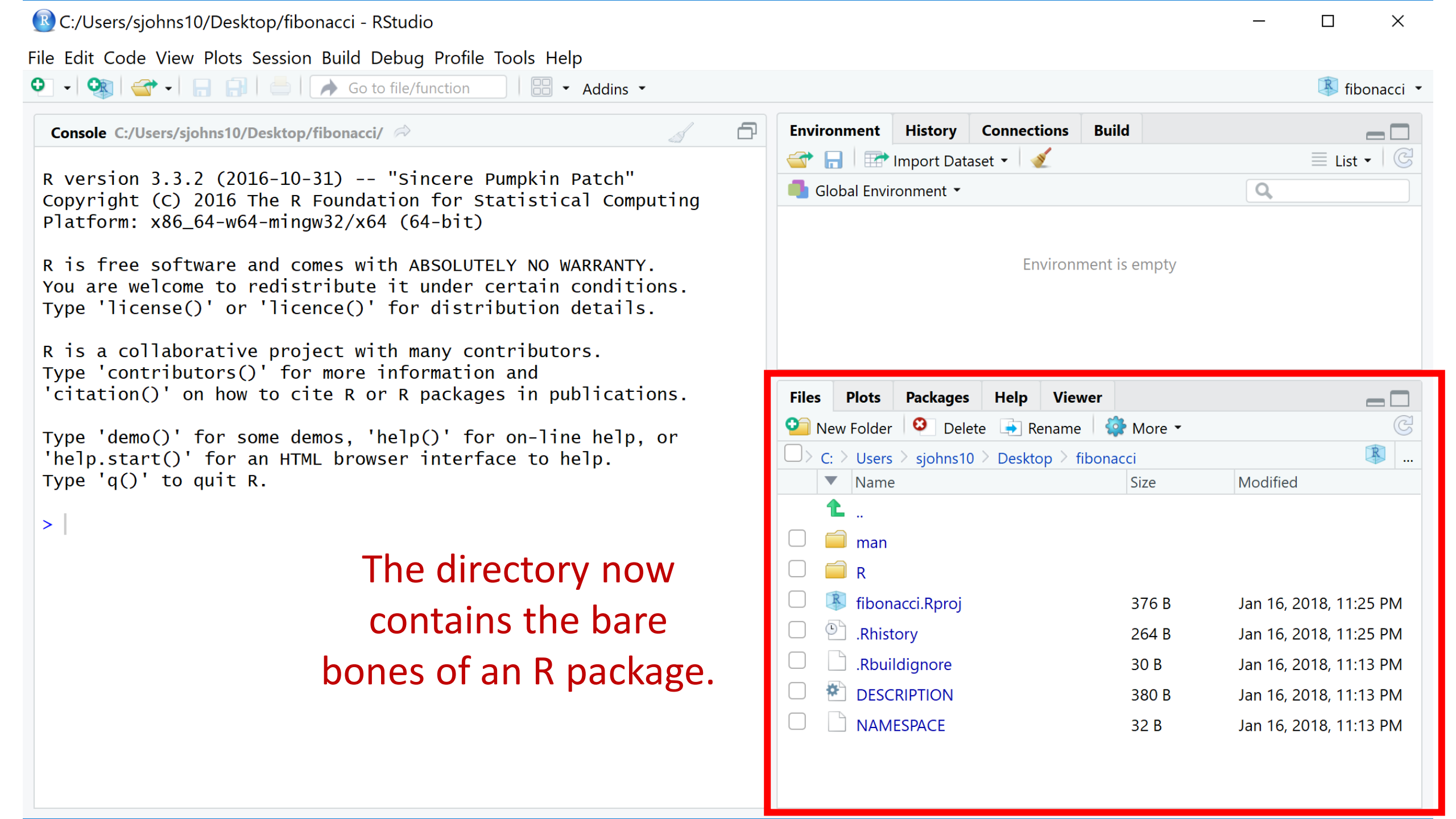
The authoritative source of information for tools to work with the current release of R is the "R Administration and Installation" manual. In particular, please read the "Windows Toolset" appendix.

Rtools Downloads

Some of the tools are incompatible with obsolete versions of R. We maintain one actively updated version of the tools, and other "frozen" snapshots of them. We recommend that users use the latest release of Rtools with the latest release of R.

The current version of this file is recorded here: VERSION.txt.

Download	R compatibility	Frozen?
Rtools34.exe	R 3.3.x and later	No
Rtools33.exe	R 3.2.x to 3.3.x	Yes
Rtools32.exe	R 3.1.x to 3.2.x	Yes
Rtools31.exe	R 3.0.x to 3.1.x	Yes
Rtools30.exe	R >2.15.1 to R 3.0.x	Yes

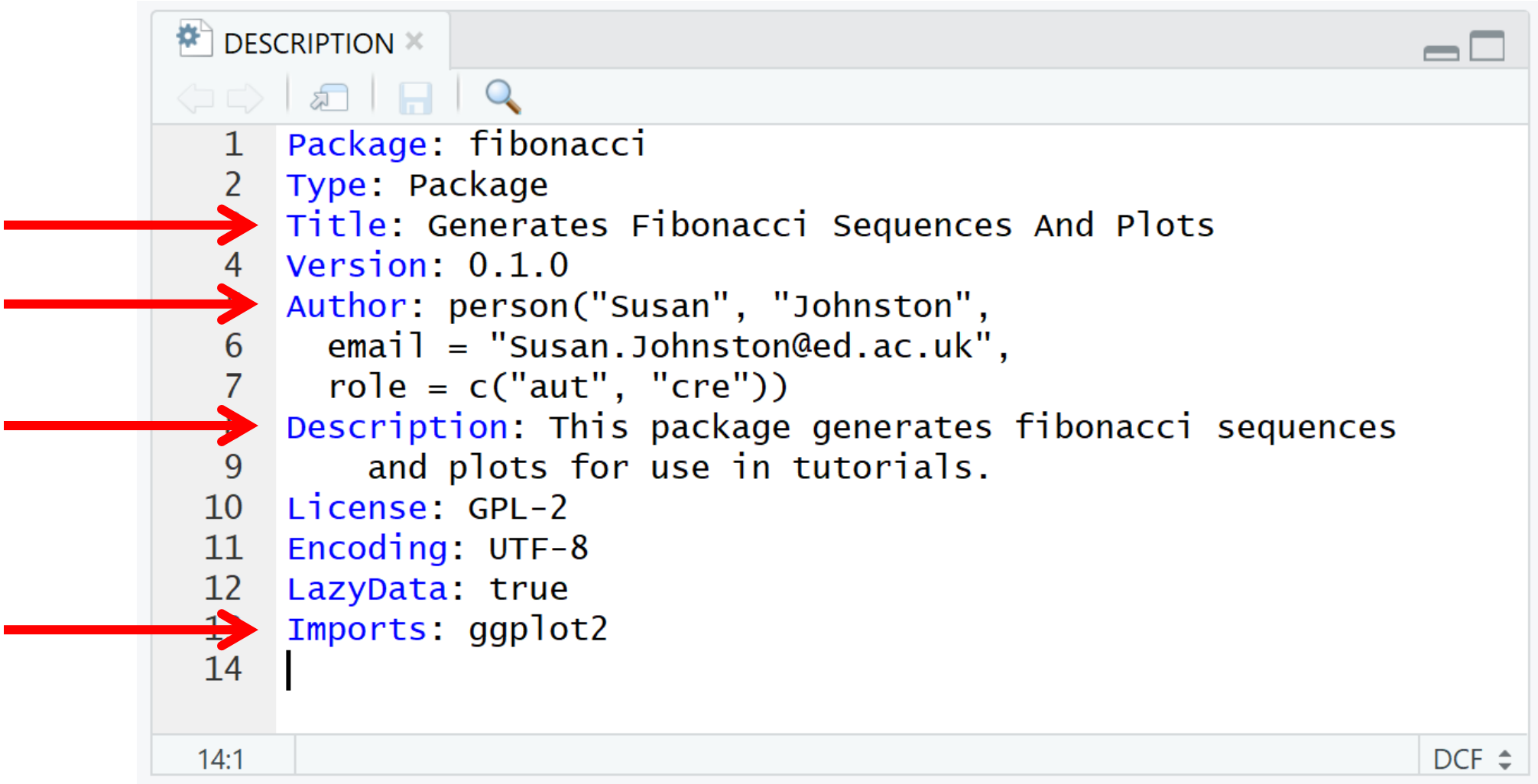

C:/Users/sjohns10/Desktop/fibonacci - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
fibonacci
Console C:/Users/sjohns10/Desktop/fibonacci/
R version 3.3.2 (2016-10-31) -- "Sincere Pumpkin Patch"
Copyright (C) 2016 The R Foundation for Statistical 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.
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.
>
Environment History Connections Build
Import Dataset
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
New Folder
Delete
Rename
More
C: Users sjohns10 Desktop fibonacci
Name Size Modified
..
man
R
fibonacci.Rproj 376 B Jan 16, 2018, 11:25 PM
.Rhistory 264 B Jan 16, 2018, 11:25 PM
.Rbuildignore 30 B Jan 16, 2018, 11:13 PM
DESCRIPTION 380 B Jan 16, 2018, 11:13 PM
NAMESPACE 32 B Jan 16, 2018, 11:13 PM
The directory now contains the bare bones of an R package.

Package components: bare bones

- DESCRIPTION — Metadata (authors, dependencies, version etc.)
- NAMESPACE — Namespaces (don't worry about this file yet).
- R/ — Directory containing code.
- man/ — Directory containing object documentation.

Package components: bare bones

- **DESCRIPTION** **Metadata (authors, dependencies, version etc.)**
- NAMESPACE Namespaces (don't worry about this file yet).
- R/ Directory containing code.
- man/ Directory containing object documentation.



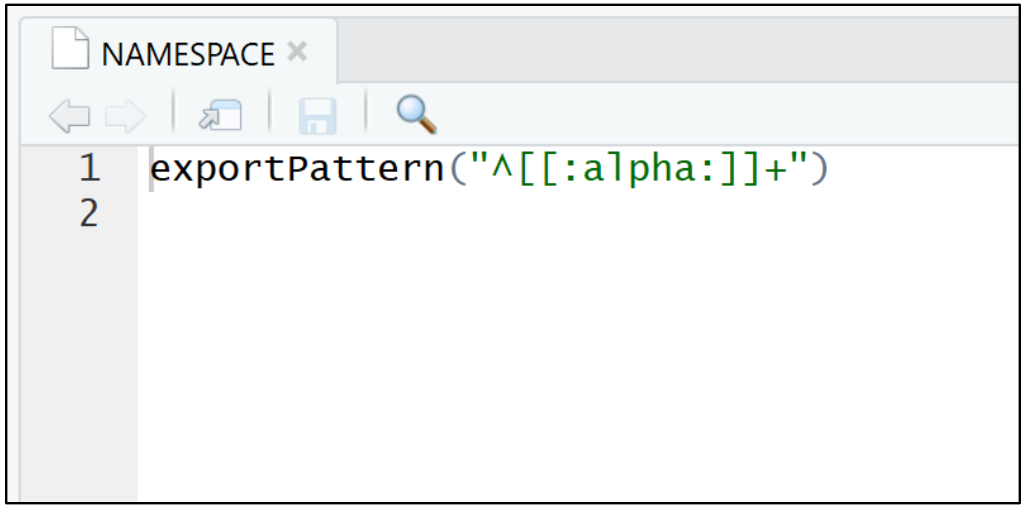
- Edit the title, author, description and imports.
- http://r-pkgs.had.co.nz/description.html

Package components: bare bones

- DESCRIPTION — Metadata (authors, dependencies, version etc.)
- **NAMESPACE** — **Namespaces (don't worry about this file yet).**
- R/ — Directory containing code.
- man/ — Directory containing object documentation.

- It is only important to understand NAMESPACE if you are submitting your package to CRAN.

- You shouldn’t edit it by hand.

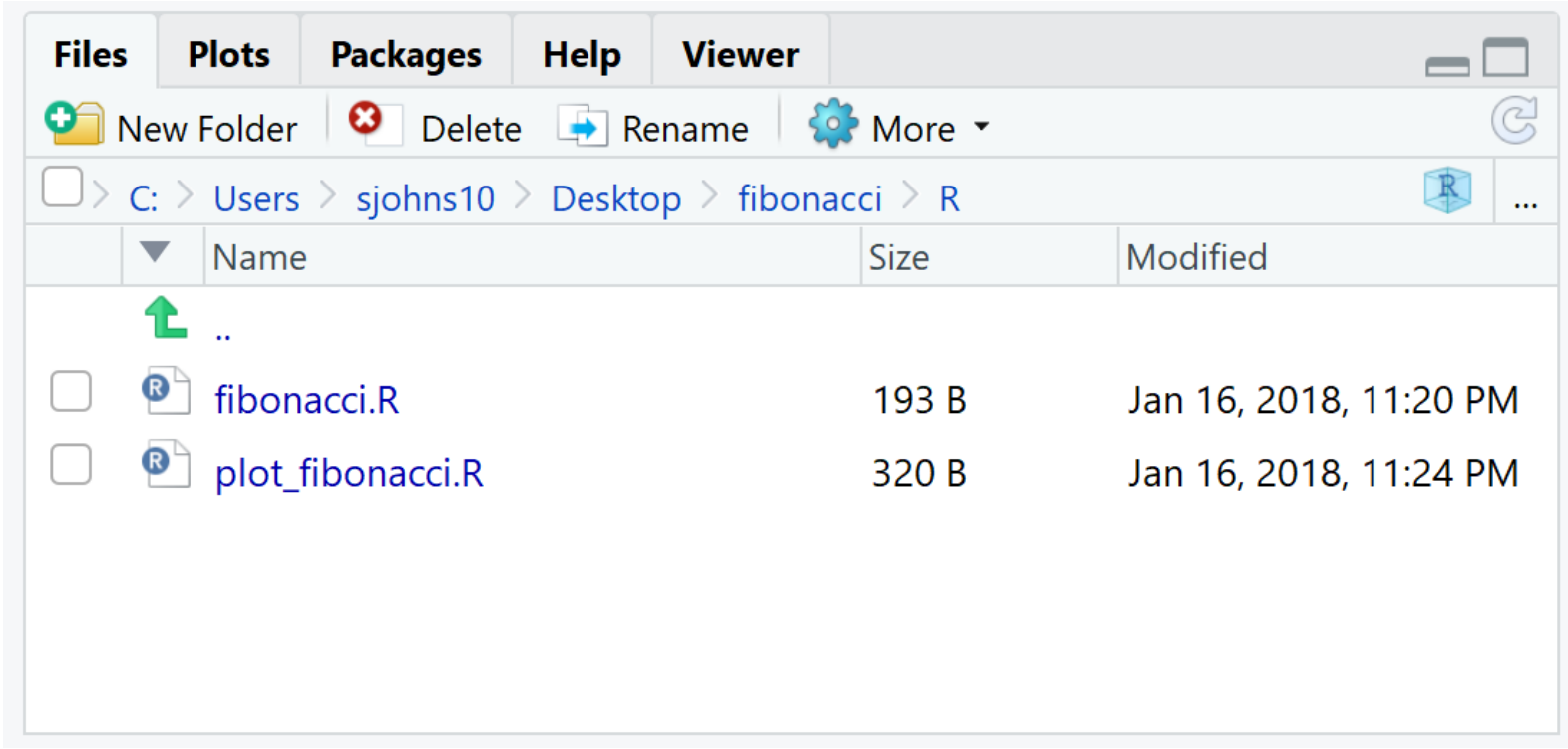


- (I will contradict myself on this in a few slides’ time.)

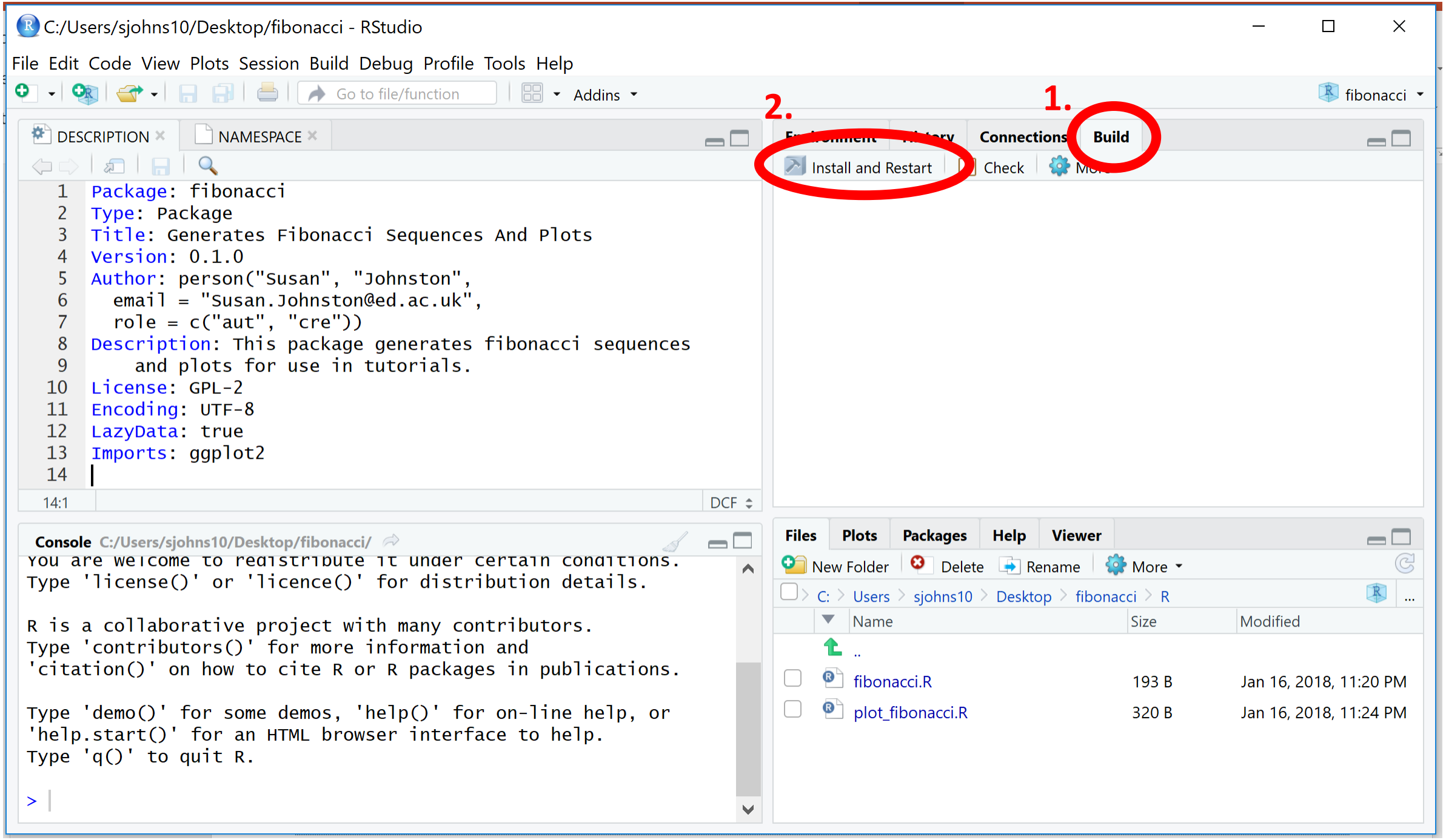
Package components: bare bones

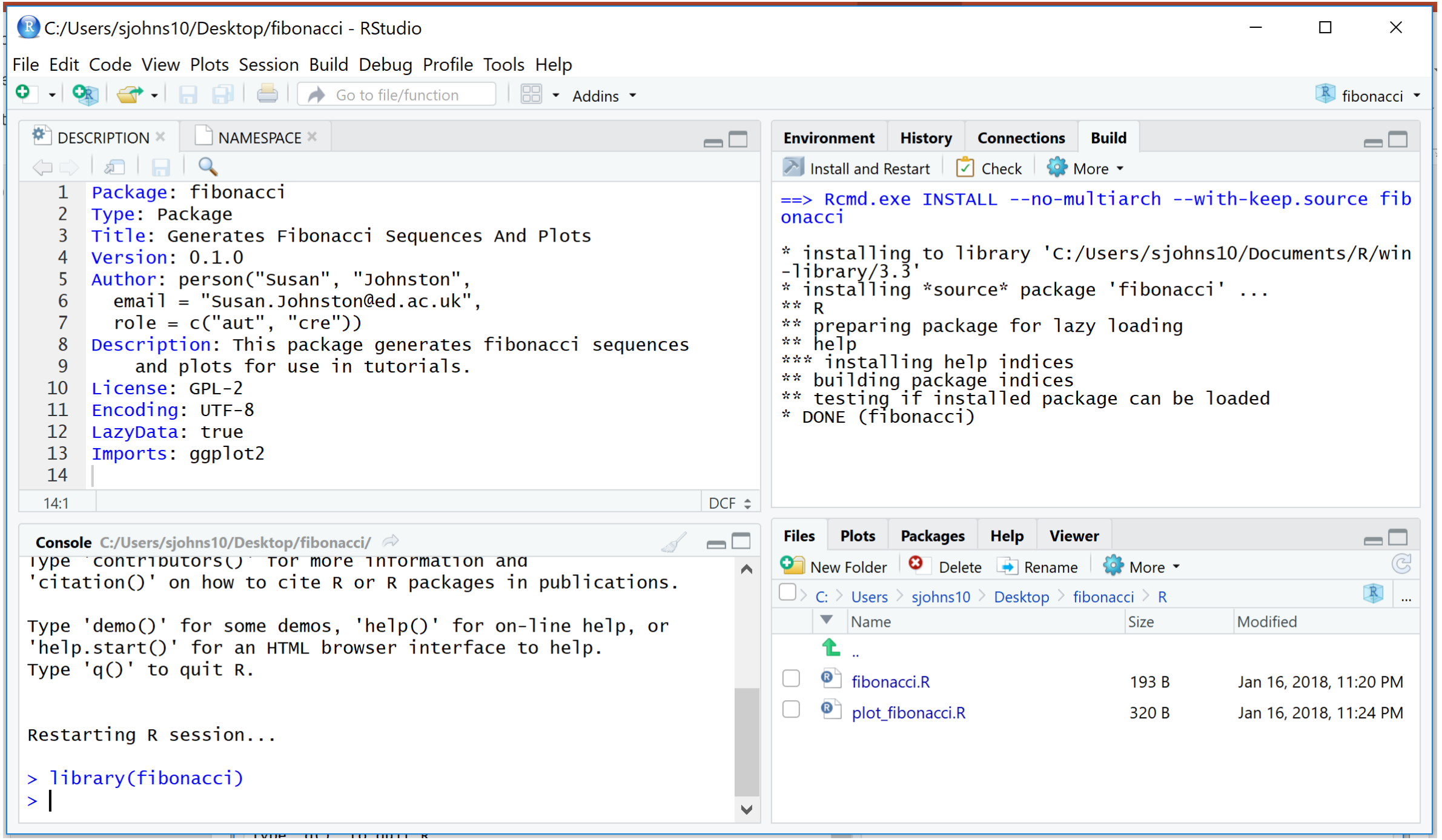
- DESCRIPTION — Metadata (authors, dependencies, version etc.)
- NAMESPACE — Namespaces (don't worry about this file yet).
- **R/** — **Directory containing code.**
- man/ — Directory containing object documentation.

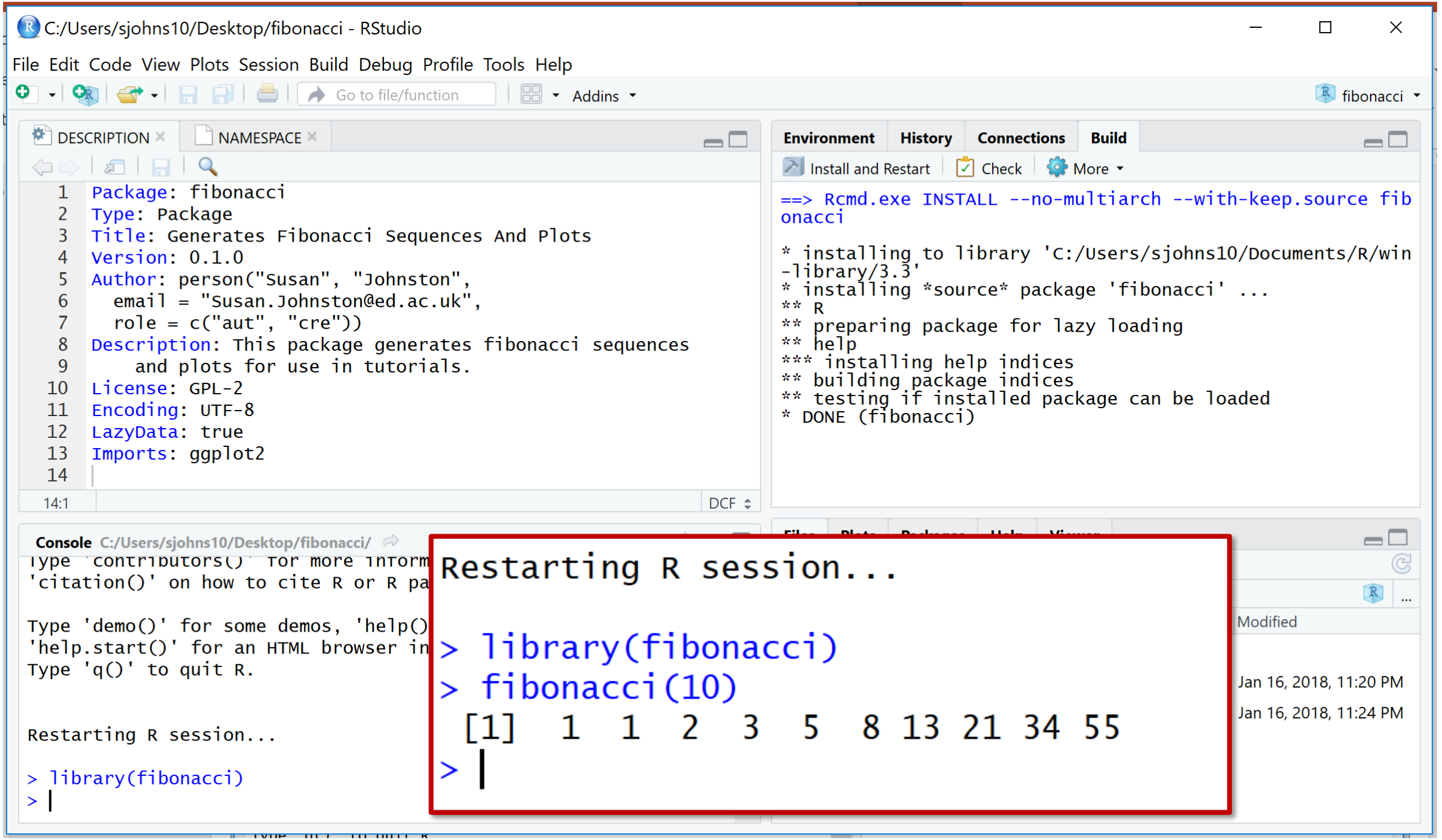
- Copy and paste the function files into the R directory.



- Now we are ready to run the first package build.


C:/Users/sjohns10/Desktop/fibonacci - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
fibonacci
2.
1.
DESCRIPTION
NAMESPACE
Connections
Build
Install and Restart
Check
1 Package: fibonacci
2 Type: Package
3 Title: Generates Fibonacci Sequences And Plots
4 Version: 0.1.0
5 Author: person("Susan", "Johnston",
6 email = "Susan.Johnston@ed.ac.uk",
7 role = c("aut", "cre"))
8 Description: This package generates fibonacci sequences
9 and plots for use in tutorials.
10 License: GPL-2
11 Encoding: UTF-8
12 LazyData: true
13 Imports: ggplot2
14
14:1
DCF
Console C:/Users/sjohns10/Desktop/fibonacci/
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
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.
>
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
C: Users sjohns10 Desktop fibonacci R
Name
Size
Modified
..
fibonacci.R
193 B
Jan 16, 2018, 11:20 PM
plot_fibonacci.R
320 B
Jan 16, 2018, 11:24 PM

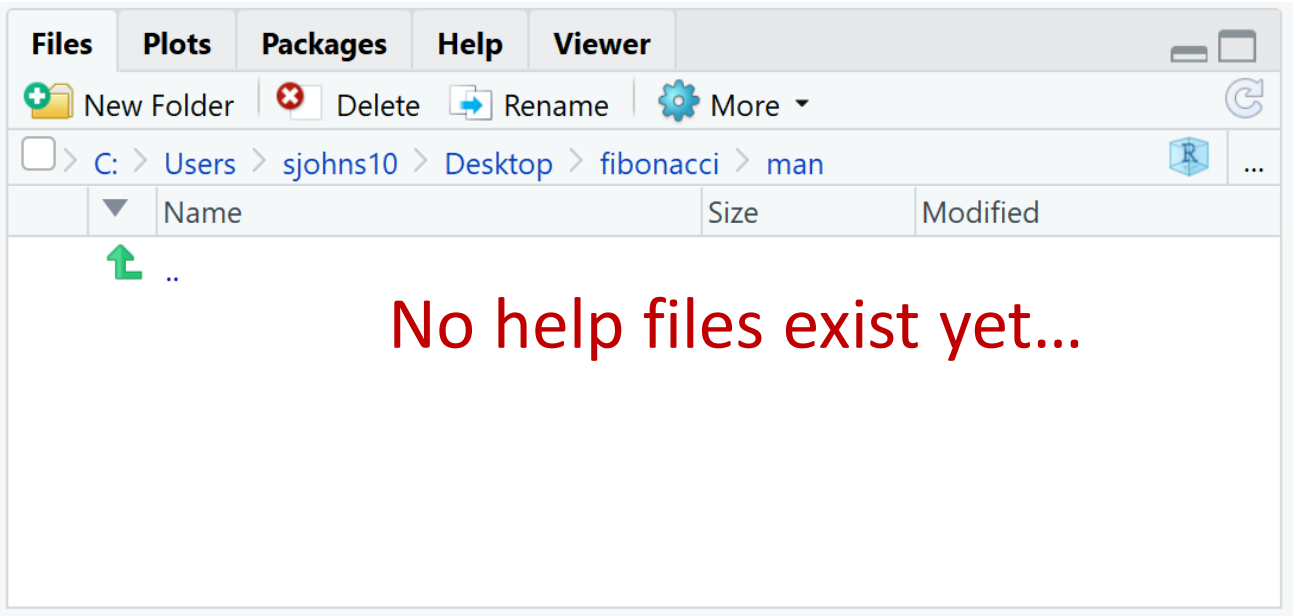

C:/Users/sjohns10/Desktop/fibonacci - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
fibonacci
DESCRIPTION
NAMESPACE
1 Package: fibonacci
2 Type: Package
3 Title: Generates Fibonacci Sequences And Plots
4 Version: 0.1.0
5 Author: person("Susan", "Johnston",
6 email = "Susan.Johnston@ed.ac.uk",
7 role = c("aut", "cre"))
8 Description: This package generates fibonacci sequences
9 and plots for use in tutorials.
10 License: GPL-2
11 Encoding: UTF-8
12 LazyData: true
13 Imports: ggplot2
14
14:1
DCF
Environment
History
Connections
Build
Install and Restart
Check
More
==> Rcmd.exe INSTALL --no-multiarch --with-keep.source fibonacci
* installing to library 'C:/Users/sjohns10/Documents/R/win-library/3.3'
* installing *source* package 'fibonacci' ...
** R
** preparing package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded
* DONE (fibonacci)
Console C:/Users/sjohns10/Desktop/fibonacci/
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.
Restarting R session...
> library(fibonacci)
>
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
C: Users sjohns10 Desktop fibonacci R
Name
Size
Modified
..
fibonacci.R
193 B
Jan 16, 2018, 11:20 PM
plot_fibonacci.R
320 B
Jan 16, 2018, 11:24 PM


C:/Users/sjohns10/Desktop/fibonacci - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
fibonacci
DESCRIPTION
NAMESPACE
1 Package: fibonacci
2 Type: Package
3 Title: Generates Fibonacci Sequences And Plots
4 Version: 0.1.0
5 Author: person("Susan", "Johnston",
6 email = "Susan.Johnston@ed.ac.uk",
7 role = c("aut", "cre"))
8 Description: This package generates fibonacci sequences
9 and plots for use in tutorials.
10 License: GPL-2
11 Encoding: UTF-8
12 LazyData: true
13 Imports: ggplot2
14
14:1
DCF
Environment
History
Connections
Build
Install and Restart
Check
More
==> Rcmd.exe INSTALL --no-multiarch --with-keep.source fibonacci
* installing to library 'C:/Users/sjohns10/Documents/R/win-library/3.3'
* installing *source* package 'fibonacci' ...
** R
** preparing package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded
* DONE (fibonacci)
Console C:/Users/sjohns10/Desktop/fibonacci/
'citation()' on how to cite R or R pa
Type 'demo()' for some demos, 'help()
'help.start()' for an HTML browser in
Type 'q()' to quit R.
Restarting R session...
> library(fibonacci)
>
Modified
Jan 16, 2018, 11:20 PM
Jan 16, 2018, 11:24 PM
Restarting R session...
> library(fibonacci)
> fibonacci(10)
[1] 1 1 2 3 5 8 13 21 34 55
>

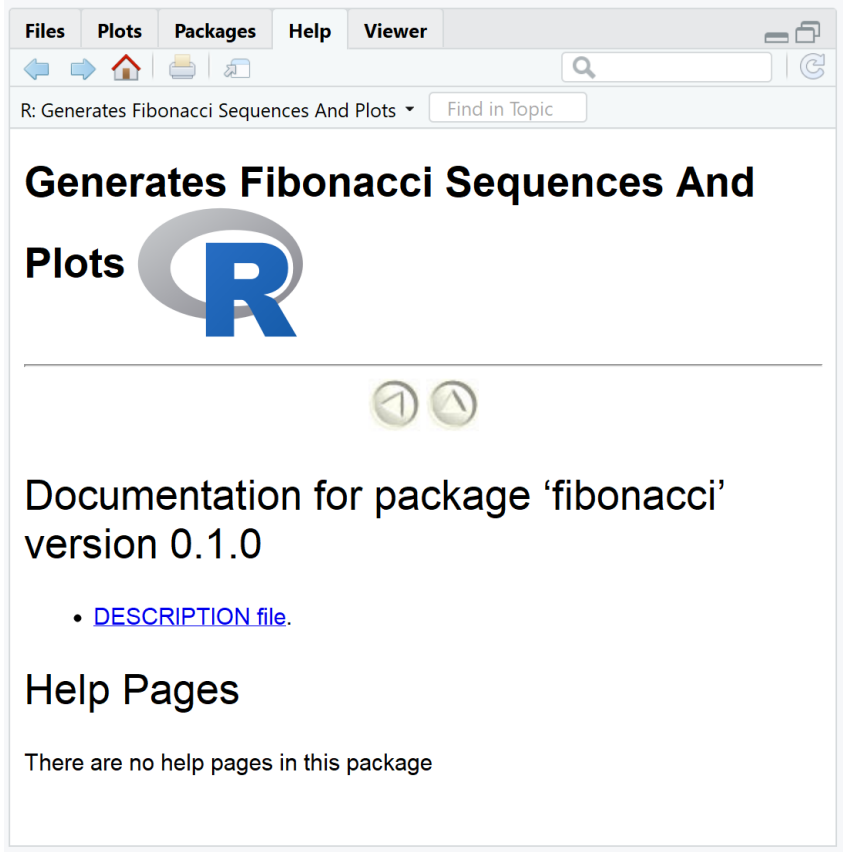
Package components: bare bones

- DESCRIPTION — Metadata (authors, dependencies, version etc.)
- NAMESPACE — Namespaces (don't worry about this file yet).
- R/ — Directory containing code.
- **man/** — **Directory containing object documentation.**

```
> ?fibonacci
No documentation for ‘fibonacci’ in specified packages and libra
ries:
you could try ‘??fibonacci’
> 
```



No help files exist yet…



Documentation can be built using the package roxygen2.

Adding documentation with roxygen2

```
#’ Brief description of the function
#’
#’ Detailed description
#' @param arg1 description of 1st argument.
#' @param arg2 description of 2nd argument.
#' @import ggplot2
#' @import dplyr
#' @export
```

Critical line: tells roxygen2 to add function/imports to NAMESPACE

```
x <- function (arg1, arg2){ … }
```

plot_fibonacci.R

Source on Save | Run | Source

```r
#' A function to plot a Fibonacci sequences.
#'
#' This function creates a plot for a Fibonacci sequence of a specified length.
#'
#' @param len numeric. Specify the length of the sequence.
#' @param gradient logical. Default TRUE. Indicates whether points on the plot
#'   should be coloured by gradient.
#' @import ggplot2
#' @export


plot_fibonacci <- function(len, gradient = T){

  require(ggplot2)

  x <- data.frame(Order = 1:len,
                  Fibonacci.Number = fibonacci(len))

  if(gradient){
    p <- ggplot(x, aes(Order, Fibonacci.Number, colour = Order))
  } else {
    p <- ggplot(x, aes(Order, Fibonacci.Number))
  }
```

This information should be added to the beginning of the R function file.

Adding documentation with roxygen2

- Delete the NAMESPACE (just on the first run!)

```
> library(roxygen2)
> roxygenise()
First time using roxygen2. Upgrading automatically...
Updating roxygen version in C:\Users\sjohns10\Desktop\fibonacci/
DESCRIPTION
Writing NAMESPACE
Writing fibonacci.Rd
Writing plot_fibonacci.Rd
> 
```

- Build > Install and restart.

flower_match.R | NAMESPACE | DESCRIPTION | flower_match.Rd

```
1  # Generated by roxygen2: do not edit by hand
2
3  export(fibonacci)
4  export(plot_fibonacci)
5  import(ggplot2)
6
```

4:1 NAMESPACE

NAMESPACE

Files | Plots | Packages | Help | Viewer

New Folder | Delete | Rename | More

C: > Users > sjohns10 > Desktop > fibonacci > man

Name	Size	Modified
..		
fibonacci.Rd	373 B	Jan 17, 2018, 12:13 AM
plot_fibonacci.Rd	398 B	Jan 17, 2018, 12:13 AM

man/

?plot_fibonacci

Files | Plots | Packages | Help | Viewer

R: A function to plot a Fibonacci sequences. | Find in Topic

plot_fibonacci {fibonacci} R Documentation

A function to plot a Fibonacci sequences.

Description

This function creates a plot for a Fibonacci sequence of a specified length.

Usage

```
plot_fibonacci(len, gradient = T)
```

Arguments

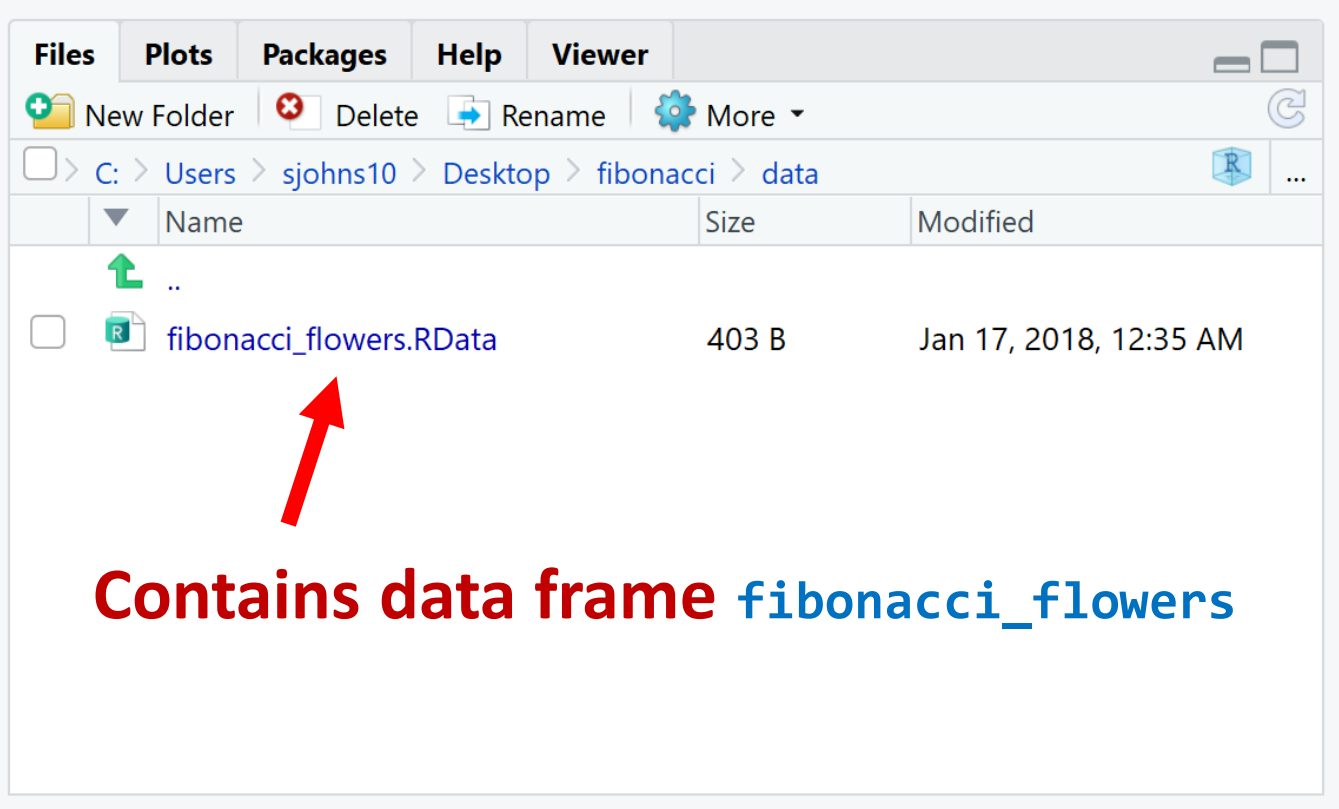
`len`	numeric. Specify the length of the sequence.
`gradient`	logical. Default TRUE. Indicates whether points on the plot should be coloured by gradient.

[Package *fibonacci* version 0.1.0 Index]

Package components: bare bones

- DESCRIPTION — Metadata (authors, dependencies, version etc.)
- NAMESPACE — Namespaces (don't worry about this file yet).
- R/ — Directory containing code.
- man/ — Directory containing object documentation.
- **data/** — **Datasets for the package.**

.RData files can be added to data/



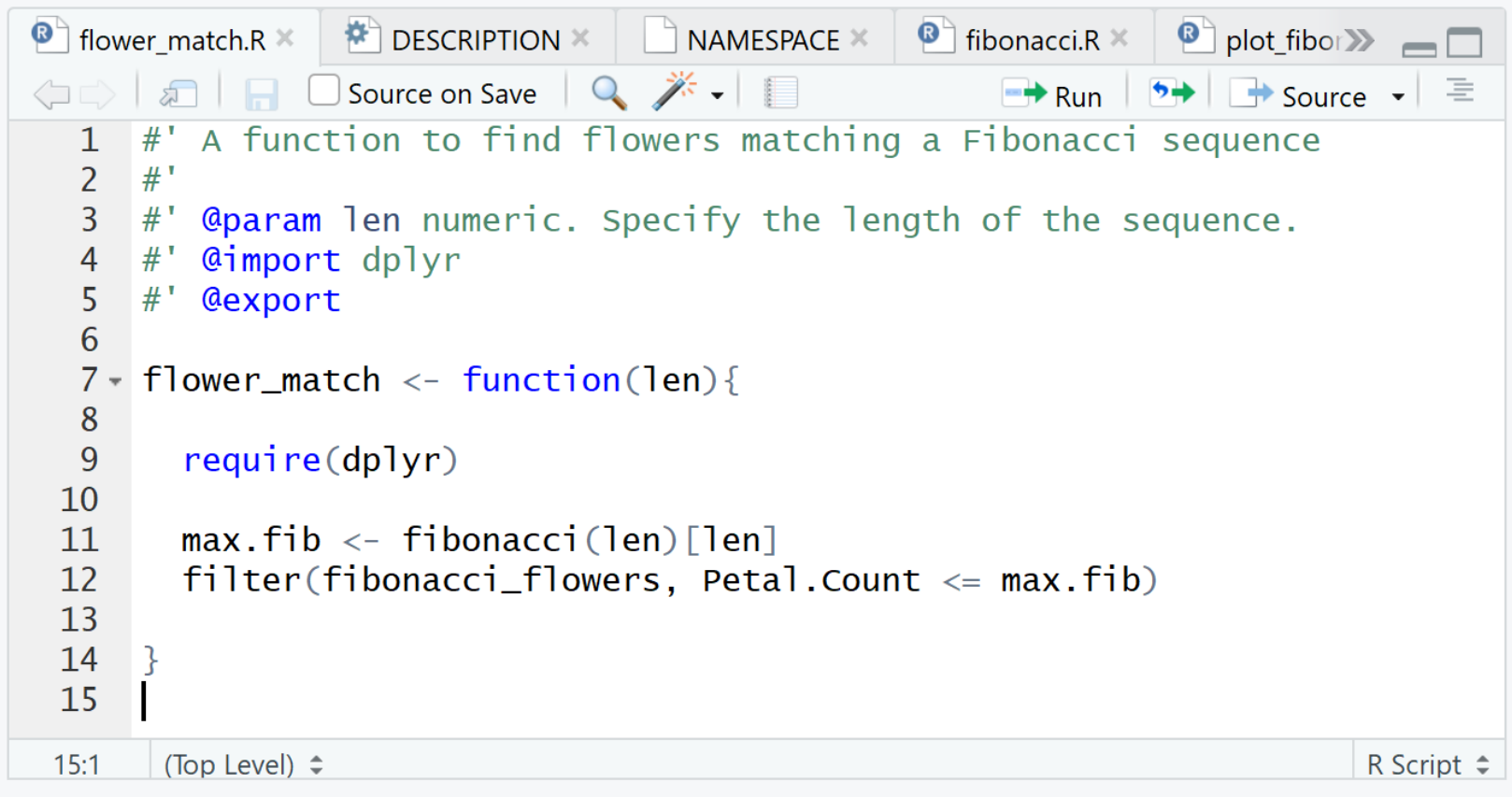
Contains data frame `fibonacci_flowers`

Build > Install and restart.

Console C:/Users/sjohns10/Desktop/fibonacci/

```
> library(fibonacci)
> fibonacci_flowers
   Petal.Count            Flower
1            3              lily
2            3              iris
3            5         buttercup
4            5         wild rose
5            5          larkspur
6            5         columbine
7            8        delphinium
8           13           ragwort
9           13    corn marigold
10          13         cineraria
11          21             aster
12          21 black-eyed susan
13          21           chicory
14          34          plantain
15          34         pytethrum
16          55  michelmas daisy
> |
```

Changing and adding new functions:

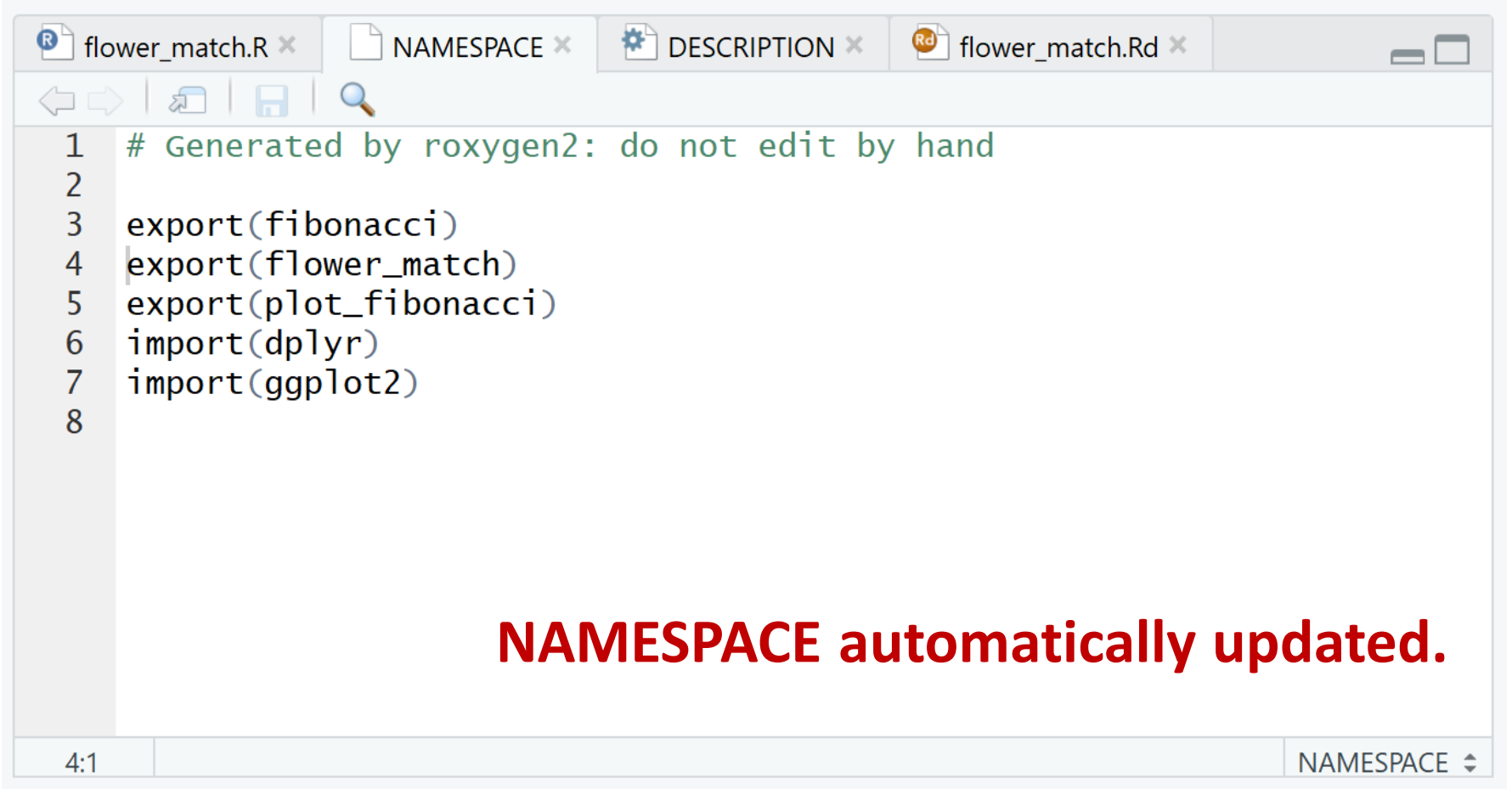


Write function and documentation.

`roxygenise()`

Build > Install and restart.

Changing and adding new functions:



Sharing your package

- Build > More > Build source package

  `fibonacci_0.1.0.tar.gz`

- Create git repo and upload to github (not shown in this talk!).
  Package can be downloaded using:

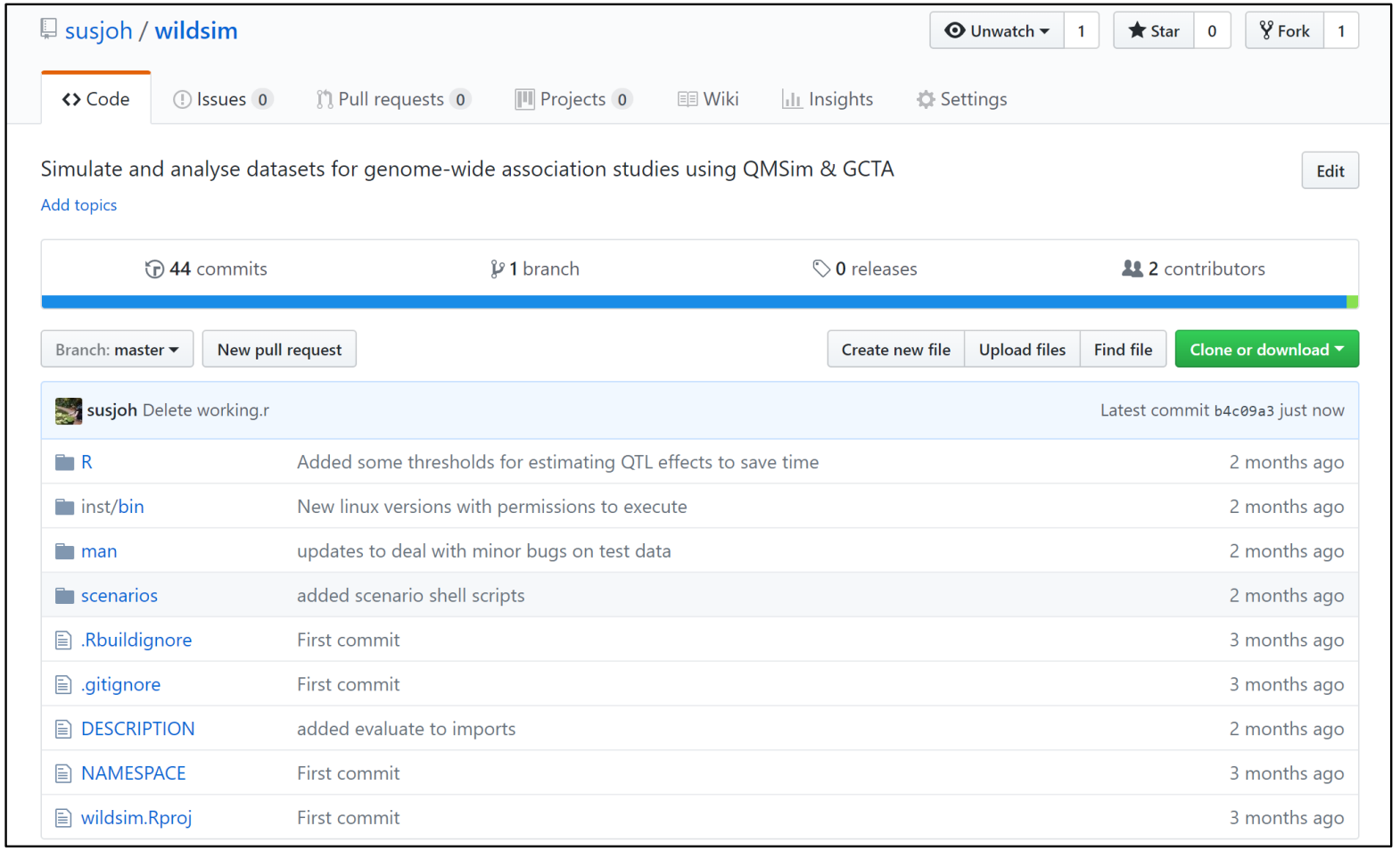
```
library(devtools)
install_github("susjoh/fibonacci")
```

- Submit to CRAN

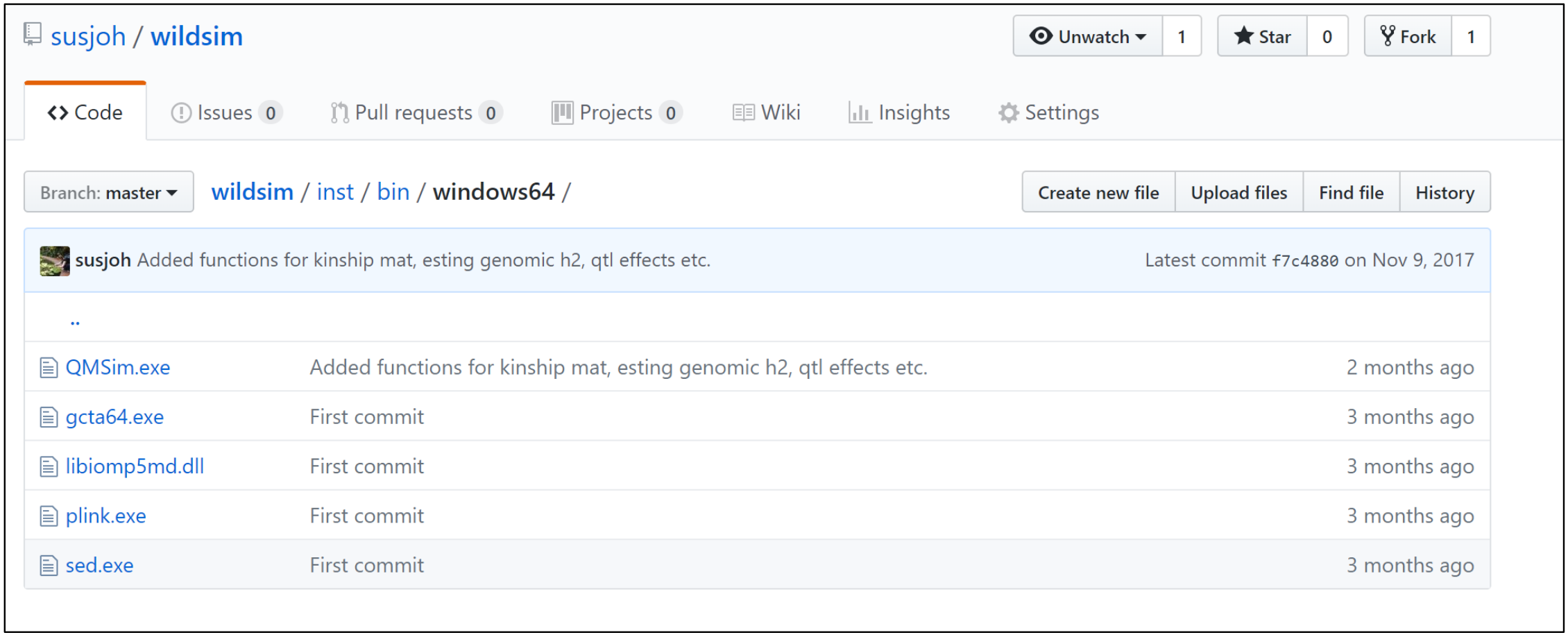
  `Good luck ☺`

Bundling executables into R

- DESCRIPTION — Metadata (authors, dependencies, version etc.)
- NAMESPACE — Namespaces (don't worry about this file yet).
- R/ — Directory containing code.
- man/ — Directory containing object documentation.
- data/ — Datasets for the package.
- **inst/bin** — **Add exectubles here (with caution)**



- It is possible to include compiled executables in R in **inst/bin**
- **This is strictly forbidden for packages on CRAN! e.g.** `library(glmmADMB)`



Executable paths can be called as follows and run using `system()`

Windows: `paste0(.libPaths()[1], "/wildsim/bin/windows64/QMSim.exe")`

Mac/Linux: `paste0(.libPaths()[length(.libPaths())], "/wildsim/bin/linux/QMSim")`

```r
#' RunQMSim: Run QMSim
#' @param input.string Command to be passed to QMSim Do not include "gcta64" at
#'   the beginning of the beginning of the string e.g. RunGCTA("--help")
#' @export
#'


RunQMSim <- function(input.string){

  if(Sys.info()["sysname"] == "Windows") {

    qmsim.path <- paste0(.libPaths()[1], "/wildsim/bin/windows64/QMSim.exe")

  } else {

    if(Sys.info()["sysname"] == "Linux"){
      qmsim.path <- paste0(.libPaths()[length(.libPaths())], "/wildsim/bin/linux/QMSim")
    } else {
      qmsim.path <- paste0(.libPaths()[length(.libPaths())], "/wildsim/bin/macosx/QMSim")
    }


  }

  system(paste(qmsim.path, input.string))

}
```

Building a basic R Package

1. Create a new package in RStudio.
2. Update `DESCRIPTION`.
3. Add functions to `R/` > Build & Reload.
4. Add function descriptions, `roxygen2::roxygenise()` > Build & Reload.
5. Create source file to share.
6. Advanced steps: add data, upload to github, bundle executables, vignettes, etc.

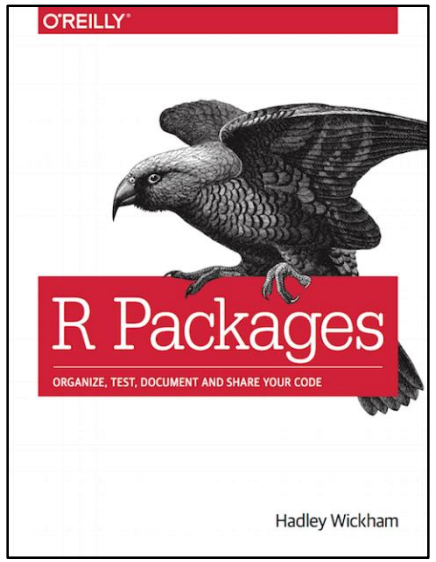
Resources/Inspiration

Hilary Parker

[Writing an R package from scratch]

Karl Broman

[R Package Primer: A minimal tutorial]



Hadley Wickham

[R Packages]