

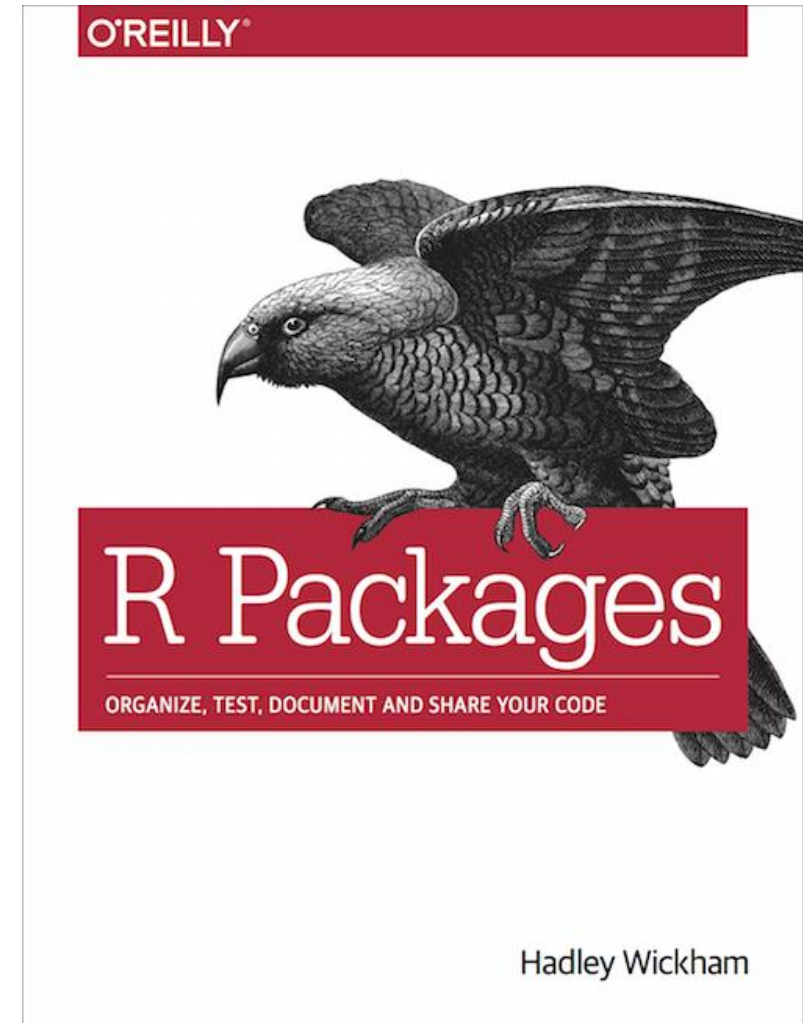
# Writing R Packages from scratch using RStudio & [roxygen2](#).

Susan Johnston, University of Edinburgh

EdinbR Meeting, 17<sup>th</sup> 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
 ASRepl.EstEffects.R	27/07/2015 18:23
 ASRepl.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
 IdPlot.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

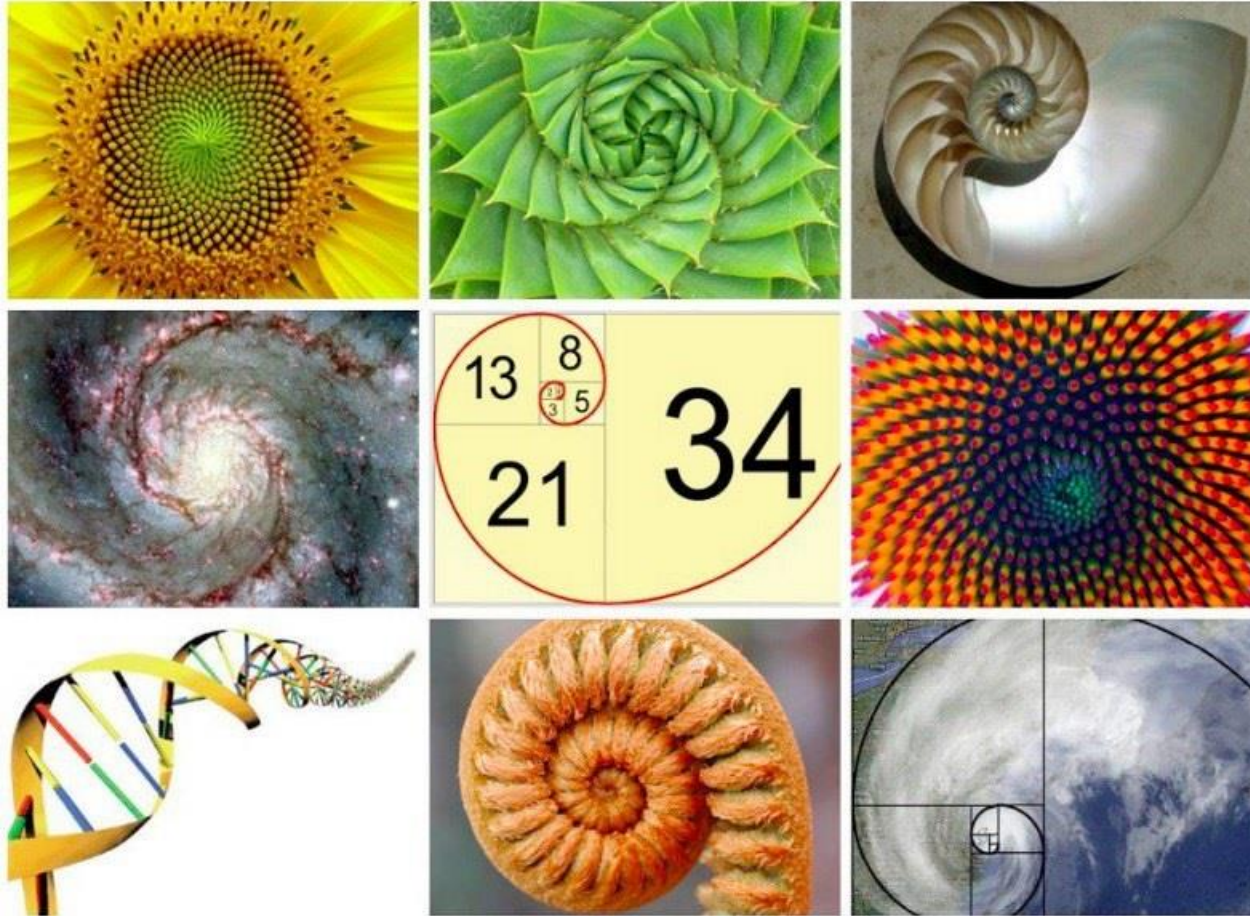
Create packages using .Rproj system.

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.

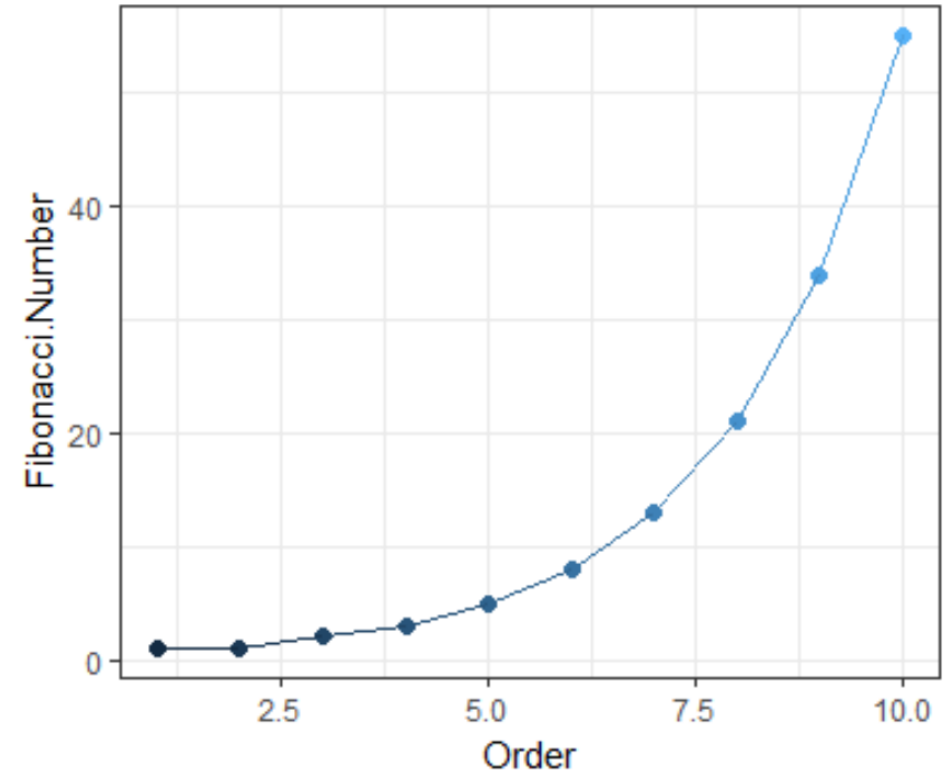
```
1
2 ▾ fibonacci <- function(len){
3
4   fibvec <- numeric(len)
5
6   fibvec[1] <- 1
7   fibvec[2] <- 1
8 ▾   for (i in 3:len) {
9     fibvec[i] <- fibvec[i-1]+fibvec[i-2]
10  }
11
12   fibvec
13 }
14
```

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

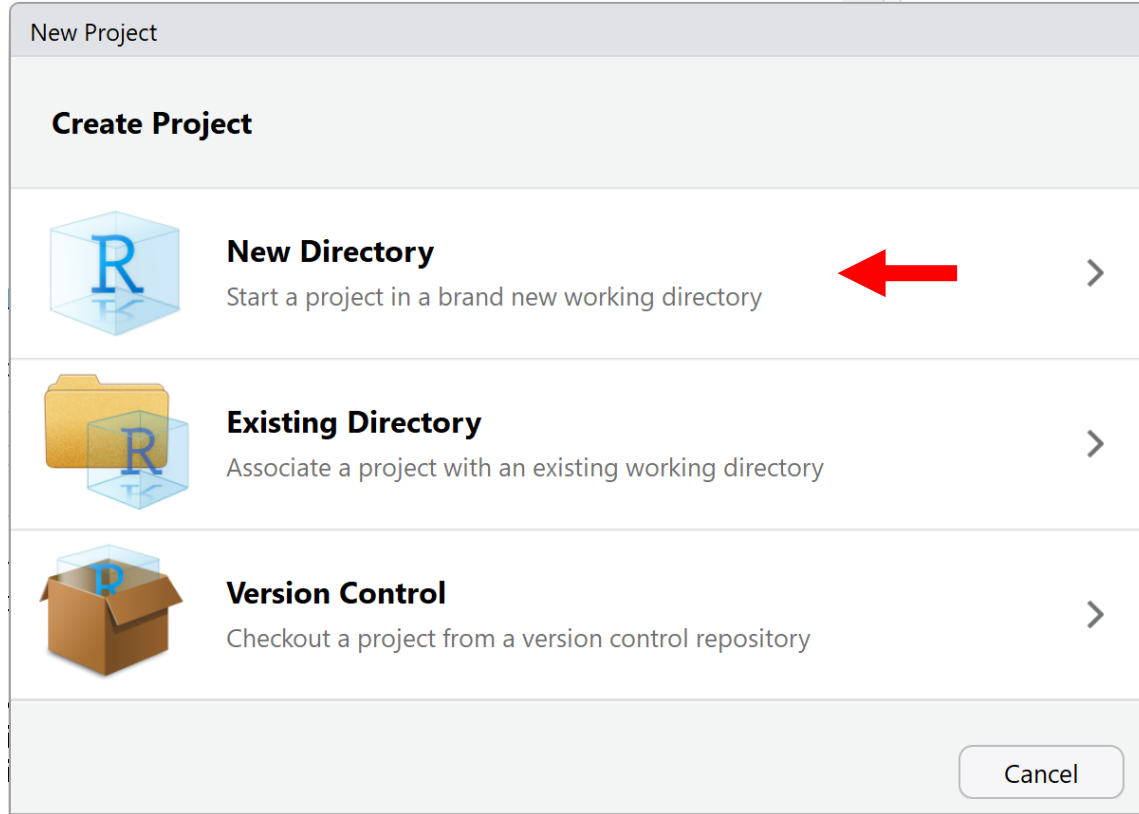


`plot_fibonacci.R` : Plots a Fibonacci sequence with colour gradient.

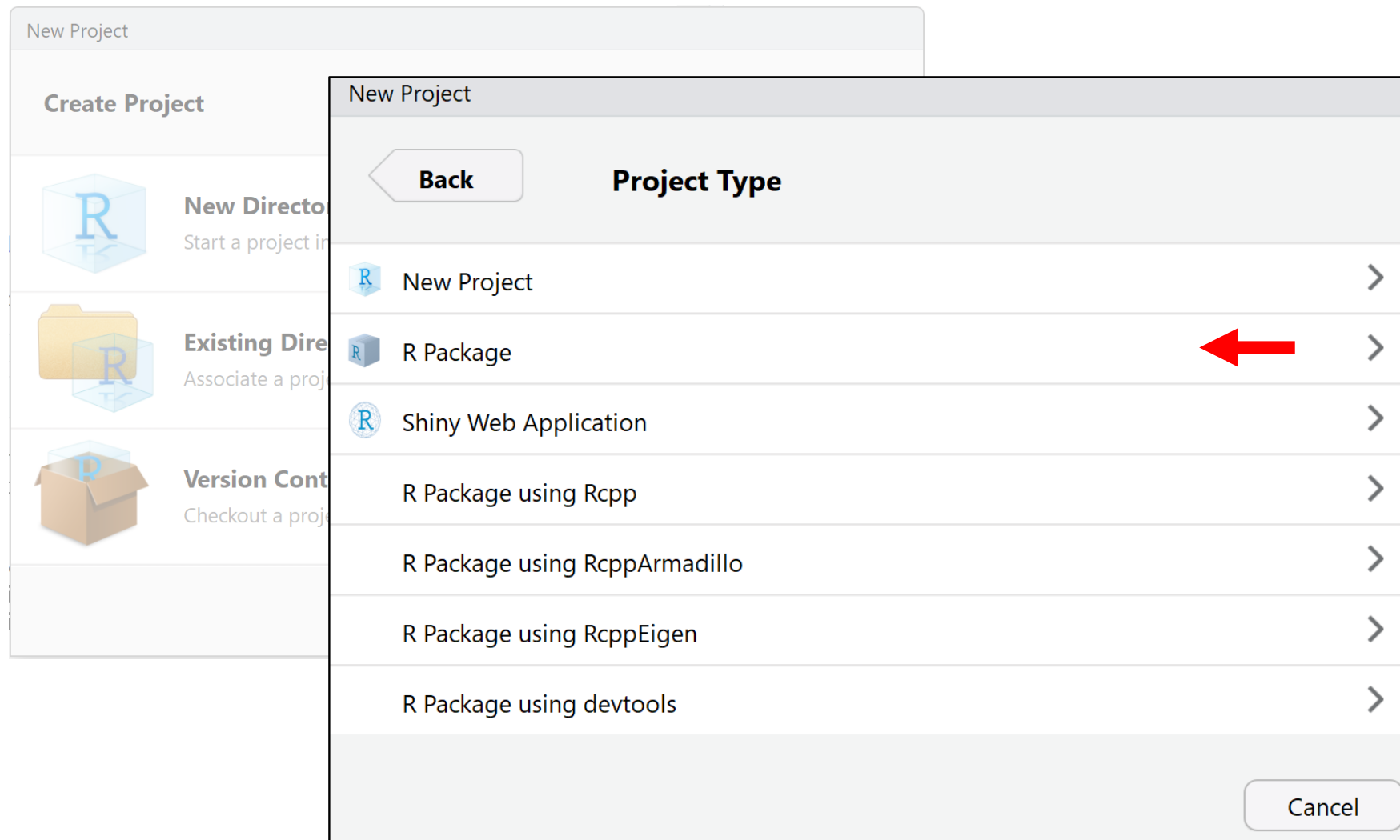
```
1
2 plot_fibonacci <- function(len, gradient = T){
3
4   require(ggplot2)
5
6   x <- data.frame(Order = 1:len,
7                   Fibonacci.Number = fibonacci(len))
8
9   if(gradient){
10     p <- ggplot(x, aes(Order, Fibonacci.Number, colour = Order))
11   } else {
12     p <- ggplot(x, aes(Order, Fibonacci.Number))
13   }
14
15   p +
16     geom_point(size = 2) +
17     geom_line() +
18     theme_bw() +
19     theme(legend.position = "none")
20
21 }
22
```



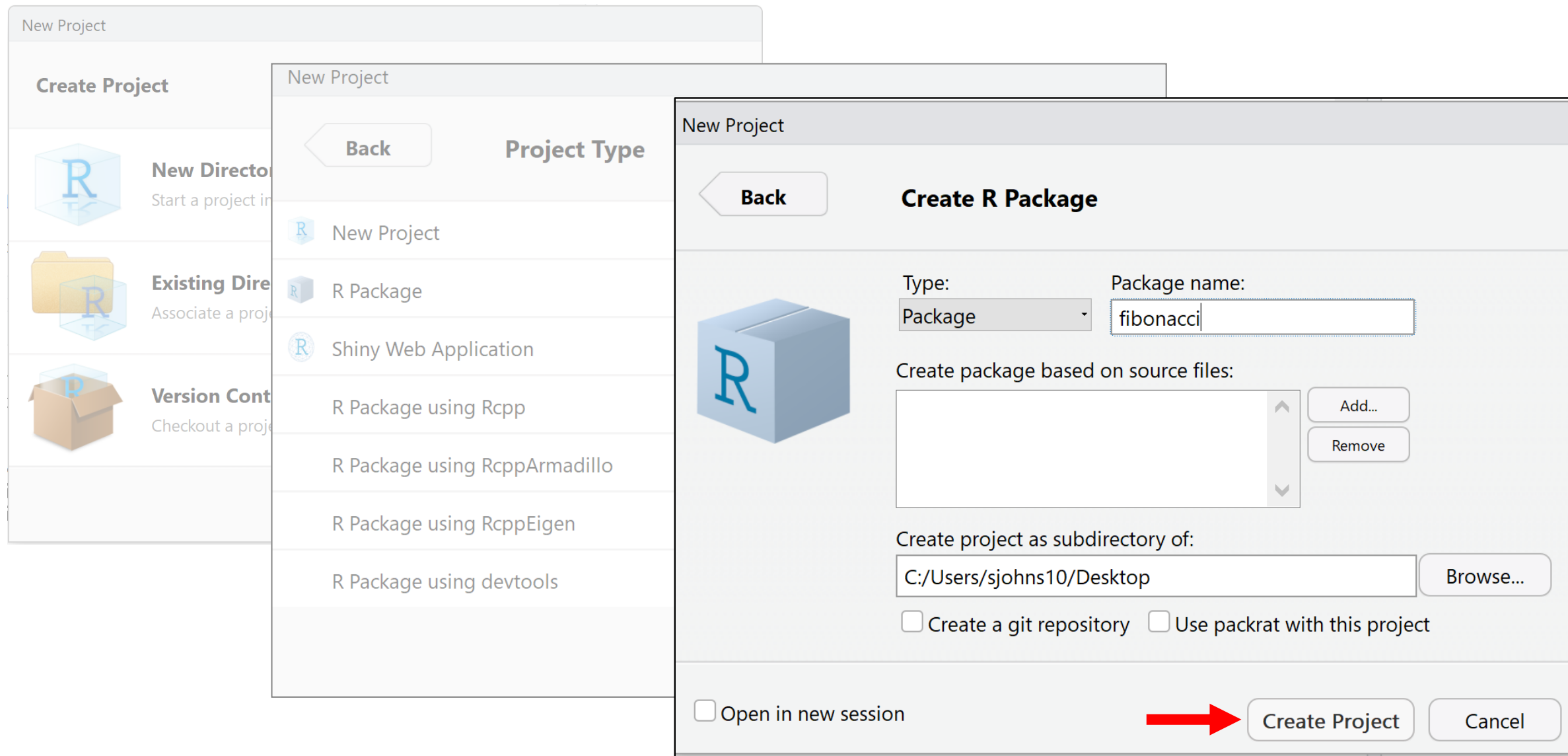
# Creating a new R Package in Rstudio.



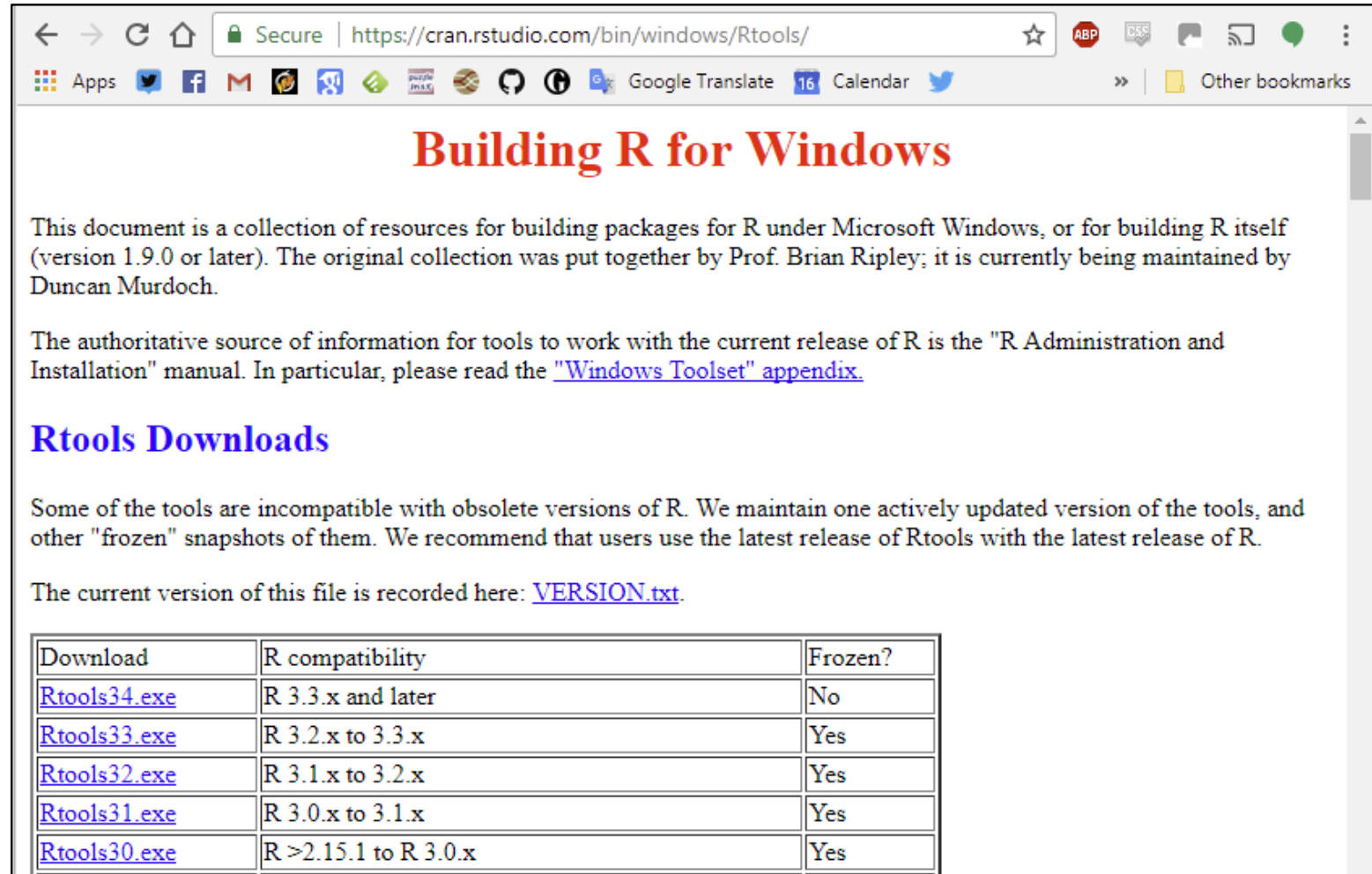
# Creating a new R Package in Rstudio.



# Creating a new R Package in Rstudio.



# Building packages in Windows requires Rtools



The screenshot shows a web browser window with the address bar displaying <https://cran.rstudio.com/bin/windows/Rtools/>. The page title is "Building R for Windows" in red. The main content explains that the document is a collection of resources for building R packages on Windows, originally by Prof. Brian Ripley and currently maintained by Duncan Murdoch. It references the "R Administration and Installation" manual and the "Windows Toolset" appendix. A section titled "Rtools Downloads" in blue text follows, stating that some tools are incompatible with older R versions and that the latest Rtools should be used with the latest R. It also mentions that the current version is recorded in [VERSION.txt](#). At the bottom, there is a table with three columns: "Download", "R compatibility", and "Frozen?". The table lists six versions of Rtools, from Rtools34.exe to Rtools30.exe, with their respective R version compatibility ranges and whether they are frozen.

## 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?
<a href="#">Rtools34.exe</a>	R 3.3.x and later	No
<a href="#">Rtools33.exe</a>	R 3.2.x to 3.3.x	Yes
<a href="#">Rtools32.exe</a>	R 3.1.x to 3.2.x	Yes
<a href="#">Rtools31.exe</a>	R 3.0.x to 3.1.x	Yes
<a href="#">Rtools30.exe</a>	R >2.15.1 to R 3.0.x	Yes

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.

&gt; |

The directory now  
contains the bare  
bones of an R package.

Environment

History

Connections

Build

Import Dataset

List

Global Environment

Environment is empty

Files

Plots

Packages

Help

Viewer

New Folder Delete Rename More

C: &gt; Users &gt; sjohns10 &gt; Desktop &gt; 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

# 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.



```
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 |
```

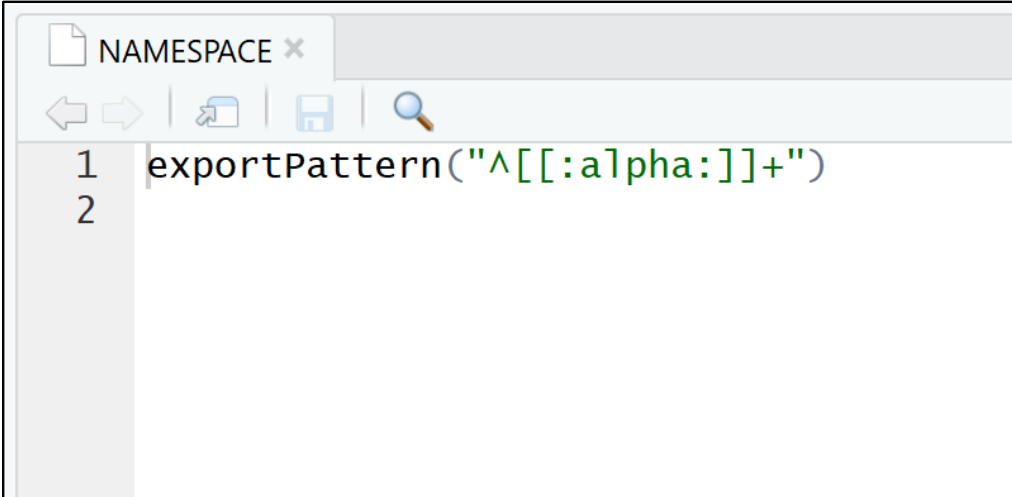
- Edit the title, author, description and imports.
- <http://r-pkgs.had.co.nz/description.html>

# 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.



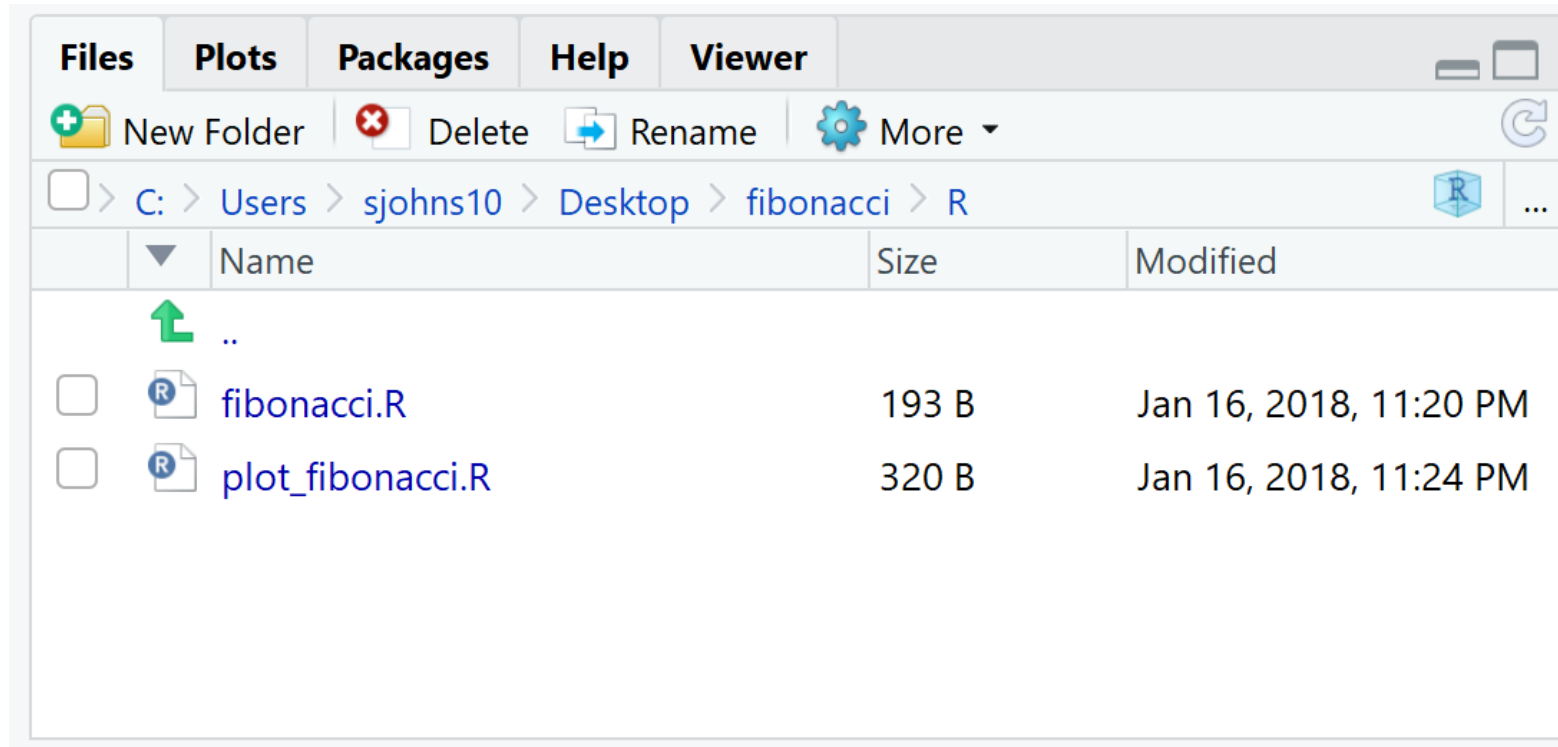
A screenshot of the RStudio interface showing the NAMESPACE file. The window title is "NAMESPACE x". The toolbar includes navigation arrows, a copy icon, a save icon, and a search icon. The code editor shows two line numbers, 1 and 2, on the left. Line 1 contains the text `exportPattern("^[:alpha:]+")`, where the pattern string is highlighted in green.

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

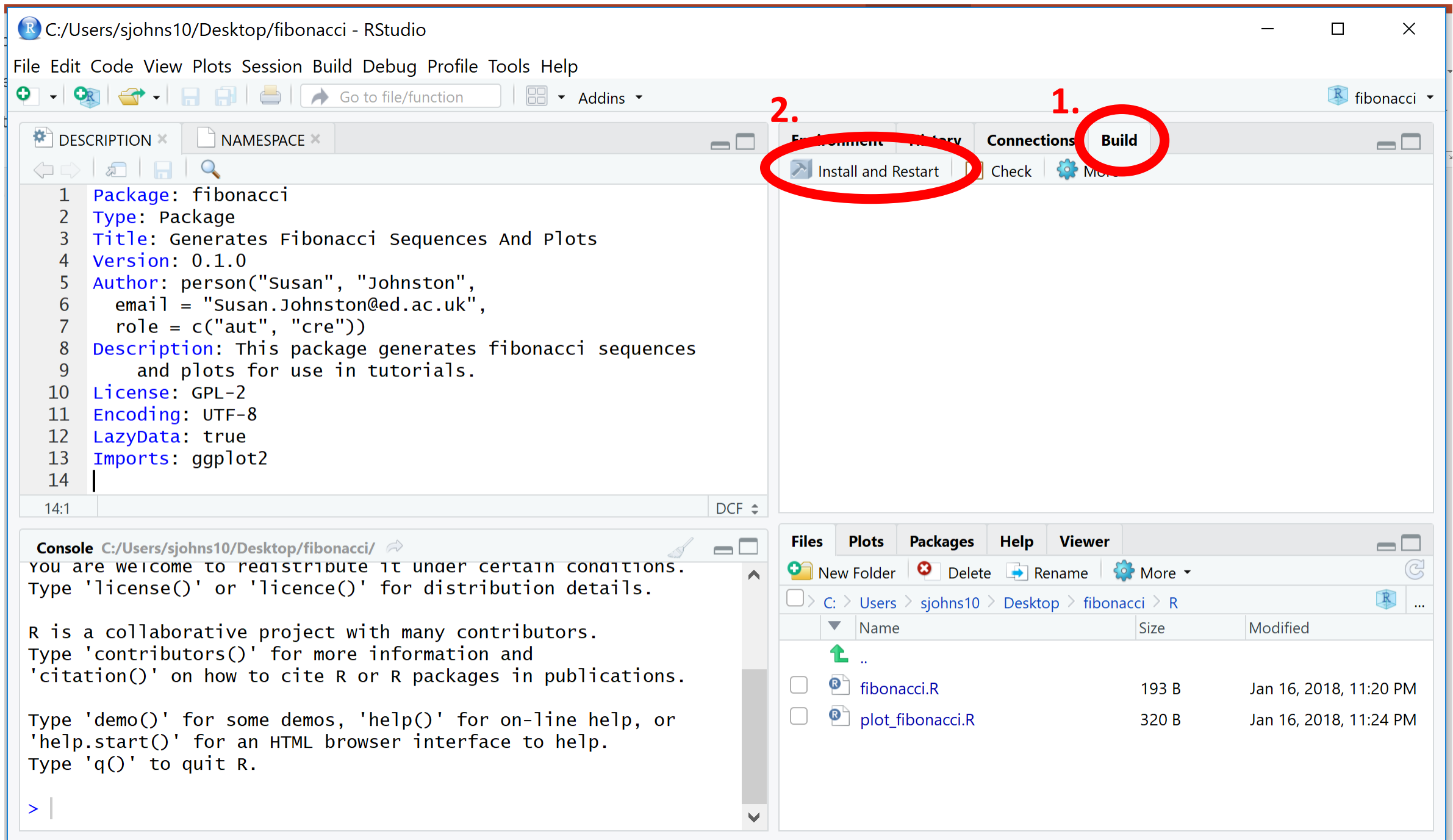
# Package components: bare bones

- 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.



DESCRIPTION x

NAMESPACE x



```
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/ 

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)
> |
```

Environment

History

Connections

Build

 Install and Restart  Check  More

```
==> Rcmd.exe INSTALL --no-multiarch --with-keep.source fib
onacci
```

```
* 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)
```


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

DESCRIPTION x

NAMESPACE x



```
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 fib
onacci

* 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  
'citation()' on how to cite R or R packages.

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)
> |
```

Restarting R session...

```
> library(fibonacci)
> fibonacci(10)
[1] 1 1 2 3 5 8 13 21 34 55
> |
```

Modified

Jan 16, 2018, 11:20 PM

Jan 16, 2018, 11:24 PM



# Package components: bare bones

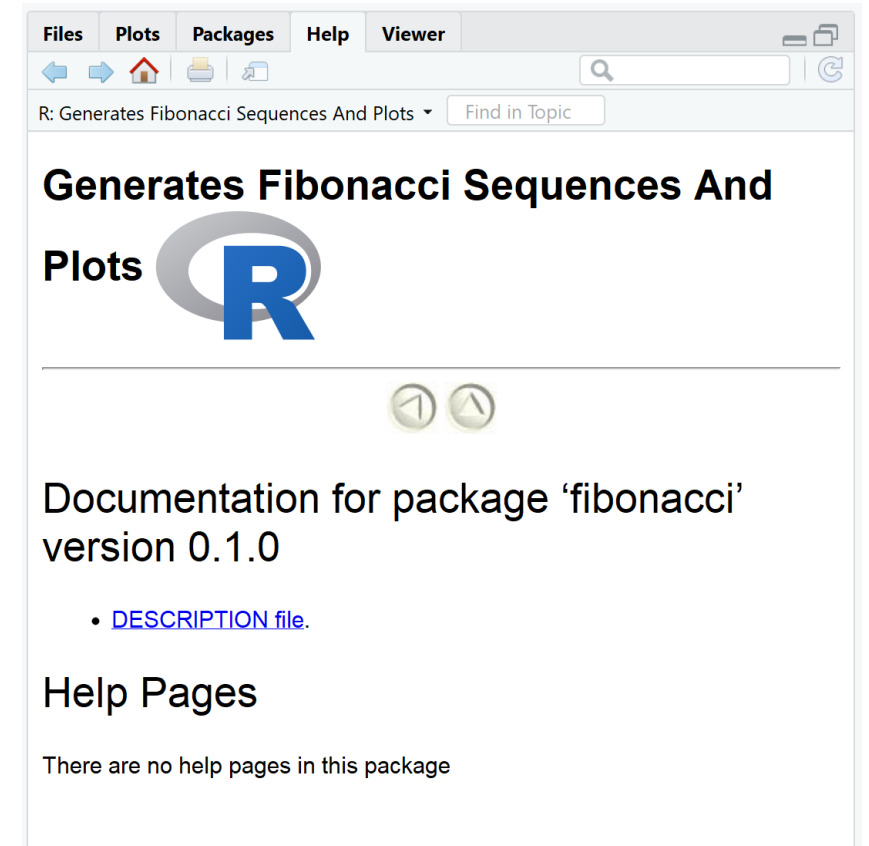
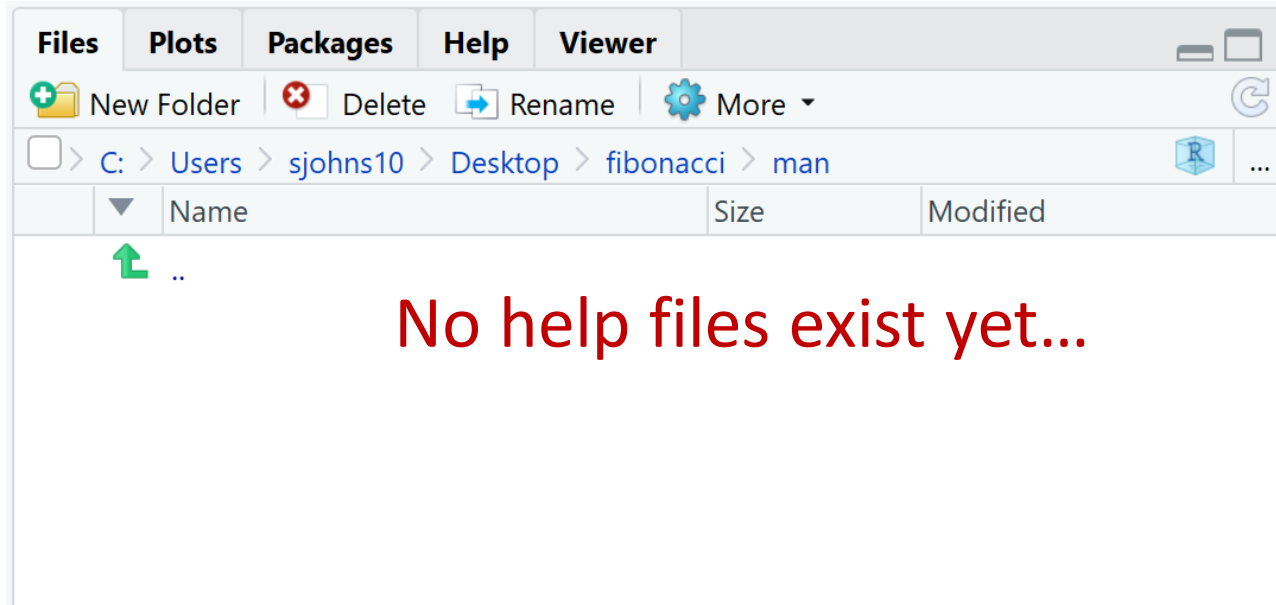
- DESCRIPTION                      Metadata (authors, dependencies, version etc.)
- NAMESPACE                        Namespaces (don't worry about this file yet).
- R/                                    Directory containing code.
- **man/**                              **Directory containing object documentation.**

```
> ?fibonacci
```

```
No documentation for 'fibonacci' in specified packages and libra  
ries:
```

```
you could try '??fibonacci'
```

```
> |
```




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 x
Source on Save
Run Source
1 #' A function to plot a Fibonacci sequences.
2 #'
3 #' This function creates a plot for a Fibonacci sequence of a specified length.
4 #'
5 #' @param len numeric. Specify the length of the sequence.
6 #' @param gradient logical. Default TRUE. Indicates whether points on the plot
7 #'   should be coloured by gradient.
8 #' @import ggplot2
9 #' @export
10
11
12 plot_fibonacci <- function(len, gradient = T){
13
14   require(ggplot2)
15
16   x <- data.frame(Order = 1:len,
17                   Fibonacci.Number = fibonacci(len))
18
19   if(gradient){
20     p <- ggplot(x, aes(Order, Fibonacci.Number, colour = Order))
21   } else {
22     p <- ggplot(x, aes(Order, Fibonacci.Number))
23   }
24 }
```



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 x  |  NAMESPACE x  |  DESCRIPTION x  |  flower_match.Rd x  |
1 # Generated by roxygen2: do not edit by hand
2
3 export(fibonacci)
4 export(plot_fibonacci)
5 import(ggplot2)
6
```

NAMESPACE

4:1 | NAMESPACE ▾

Files	Plots	Packages	Help	Viewer
New Folder Delete Rename More ▾				
C: > Users > sjohns10 > Desktop > fibonacci > man				
Name		Size	Modified	
..				
<input type="checkbox"/>	 fibonacci.Rd	373 B	Jan 17, 2018, 12:13 AM	
<input type="checkbox"/>	 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

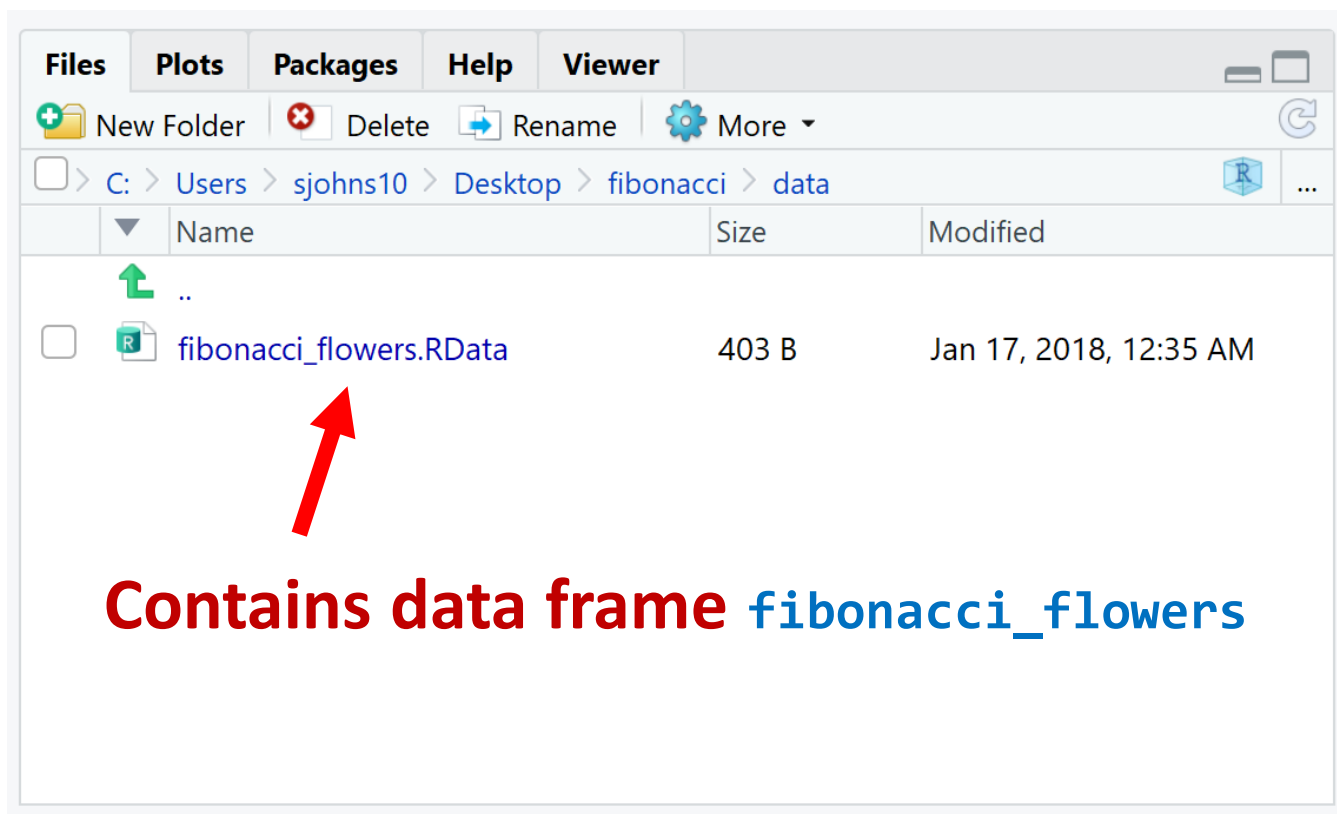
logical. Default TRUE. Indicates whether points on the plot should be coloured by gradient.

[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/



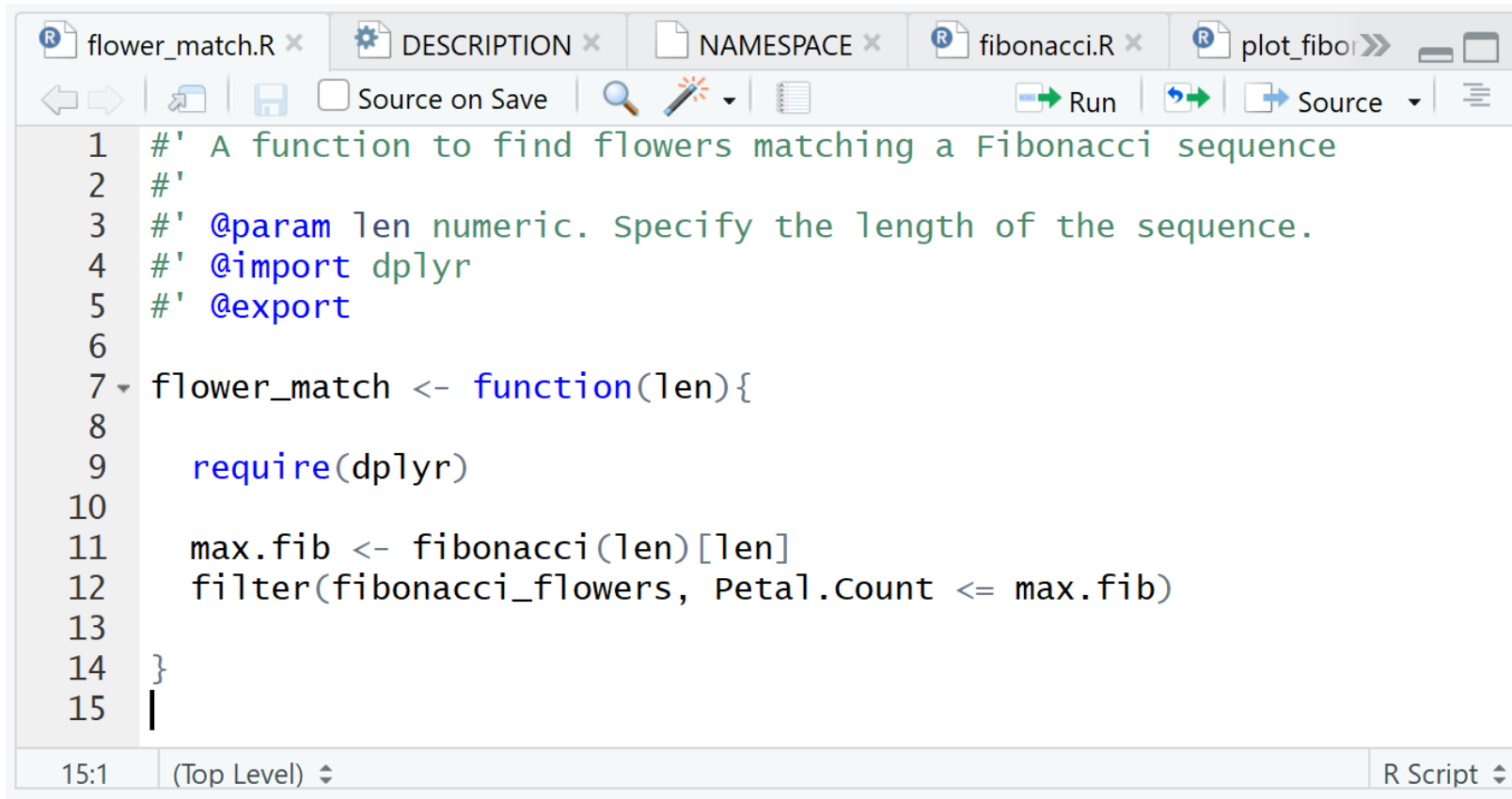
**Build > Install and restart.**

```
Console C:/Users/sjohns10/Desktop/fibonacci/
Restarting R session...

> 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:



```
1 #' A function to find flowers matching a Fibonacci sequence
2 #'
3 #' @param len numeric. Specify the length of the sequence.
4 #' @import dplyr
5 #' @export
6
7 flower_match <- function(len){
8   require(dplyr)
9
10   max.fib <- fibonacci(len)[len]
11   filter(fibonacci_flowers, Petal.Count <= max.fib)
12
13 }
14
15 |
```

**Write function and documentation.**

**roxygenise()**

**Build > Install and restart.**

# Changing and adding new functions:

```
1 # Generated by roxygen2: do not edit by hand
2
3 export(fibonacci)
4 export(flower_match)
5 export(plot_fibonacci)
6 import(dplyr)
7 import(ggplot2)
8
```

**NAMESPACE automatically updated.**

4:1 NAMESPACE

# 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 executables here (with caution)**

The screenshot shows the GitHub repository page for 'susjoh / wildsim'. At the top, there are navigation links for 'Code', 'Issues', 'Pull requests', 'Projects', 'Wiki', 'Insights', and 'Settings'. Below these, a description reads: 'Simulate and analyse datasets for genome-wide association studies using QMSim & GCTA'. A bar indicates repository statistics: 44 commits, 1 branch, 0 releases, and 2 contributors. Below this is a section for file management with buttons for 'Branch: master', 'New pull request', 'Create new file', 'Upload files', 'Find file', and 'Clone or download'. The main content area shows a list of files and folders with their commit history:

File/Folder	Commit Message	Time
<a href="#">R</a>	Added some thresholds for estimating QTL effects to save time	2 months ago
<a href="#">inst/bin</a>	New linux versions with permissions to execute	2 months ago
<a href="#">man</a>	updates to deal with minor bugs on test data	2 months ago
<a href="#">scenarios</a>	added scenario shell scripts	2 months ago
<a href="#">.Rbuildignore</a>	First commit	3 months ago
<a href="#">.gitignore</a>	First commit	3 months ago
<a href="#">DESCRIPTION</a>	added evaluate to imports	2 months ago
<a href="#">NAMESPACE</a>	First commit	3 months ago
<a href="#">wildsim.Rproj</a>	First commit	3 months ago

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

susjoh / wildsim

Unwatch 1 Star 0 Fork 1

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master wildsim / inst / bin / windows64 /

Create new file Upload files Find file History

susjoh Added functions for kinship mat, esting genomic h2, qtl effects etc. Latest commit f7c4880 on Nov 9, 2017

..

QMSim.exe	Added functions for kinship mat, esting genomic h2, qtl effects etc.	2 months ago
gcta64.exe	First commit	3 months ago
libiomp5md.dll	First commit	3 months ago
plink.exe	First commit	3 months ago
sed.exe	First commit	3 months ago

**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")`

```
1  #' RunQMSim: Run QMSim
2  #' @param input.string Command to be passed to QMSim Do not include "gcta64" at
3  #'   the beginning of the beginning of the string e.g. RunGCTA("--help")
4  #' @export
5  #'
6
7
8  RunQMSim <- function(input.string){
9
10     if(Sys.info()["sysname"] == "Windows") {
11
12         qmsim.path <- paste0(.libPaths()[1], "/wildsim/bin/windows64/QMSim.exe")
13
14     } else {
15
16         if(Sys.info()["sysname"] == "Linux"){
17             qmsim.path <- paste0(.libPaths()[length(.libPaths())], "/wildsim/bin/linux/QMSim")
18         } else {
19             qmsim.path <- paste0(.libPaths()[length(.libPaths())], "/wildsim/bin/macosx/QMSim")
20         }
21
22
23     }
24
25     system(paste(qmsim.path, input.string))
26
27 }
```

---

# 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.



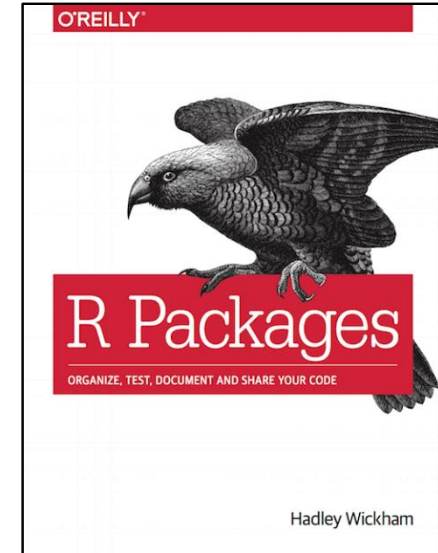
# Resources/Inspiration



Hilary Parker  
[Writing an R  
package from  
scratch](#)



Karl Broman  
[R Package Primer:  
A minimal tutorial](#)



Hadley Wickham  
[R Packages](#)