



## Introduction to Building packages in R





BarcelonaR - Workshop





Who am I?

Name: Nicolas Attalides



- Coding in R since: 2005 (yes that's before RStudio!)
- **Profession:** Senior Data Scientist and trainer (6+ yrs.)
- Education: PhD in Statistical Science from UCL (2015)
- R Status: A never-ending evolving R dinosaur
- Hobbies: Tennis and coding (not at the same time)





## Workshop Setup:

#### Wi-Fi

- Network Name: N/A
- Password: N/A

#### Packages

- {tidyverse} (version 1.3.0) tdyverse
- {devtools} (version 2.3.2)
- {usethis} (version 2.0.1)

#### Resources

R (version 3.6.3)
 RStudio (version 1.4.1106)
 R Studio





## What is an R package?

# An R package is like a **collection** of **code**, **data** and **documentation** that follow some standard rules and formats.

This is the best way for an R user to **share** their **work** and enable others to use the functionality that is developed.







#### Comprehensive R Archive Network (CRAN)

The central repository of R packages is called the Comprehensive R Archive Network (**CRAN**). This contains an archive of R distributions and has more than 17,000 packages ready to be installed and used.

	Available CRAN Packages By Name		
	<u>A B C D E F G H I J K L M N O P Q R S T U V W X Y Z</u>		
<u>A3</u>	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models		
<u>aaSEA</u>	Amino Acid Substitution Effect Analyser		
AATtools	Reliability and Scoring Routines for the Approach-Avoidance Task		
ABACUS	Apps Based Activities for Communicating and Understanding Statistics		
<u>abbyyR</u>	Access to Abbyy Optical Character Recognition (OCR) API		
abc	Tools for Approximate Bayesian Computation (ABC)		
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)		
ADC DAD	Array Pasad CoG Pagion Analysis Dinalina		

Find out more about CRAN here: <u>https://cran.r-project.org/</u>





## Topics

Workshop aim:

Learn how to build an R package in an easy step-by-step approach.

#### Topics:

- Learn how to create an R package within **R** Studio •
- Understand the package structure and its various components •
- Learn how to write, document and test functions in (R) for a package •
- Learn how to check, build and install an R package •





#### Create an R package

File	Edit	Code	View	Plots	Session	Build	Deb	
	New File							
	New Project							
	Open F Open F Recent	File in N	ew Coli	umn	Ctrl+O			
	Open I	Project Project i t Project	n New	Session			•	
	Import Dataset					•		
	Save Save A	.s			Ctrl+S			
	Save A	dl –			Ctrl+Alt	+S		
	Publish	<b>)</b>						
	Print							
	Close				Ctrl+W			
	Close /	All			Ctrl+Shi	ft+W		
	Close /	All Exce	pt Curre	ent	Ctrl+Alt	+Shift+	W	
	Close F	Project						
	Quit Se	ession			Ctrl+Q			







### Create an R package

New	Project Wizard		
$\langle$	Back Project Type		
R	New Project	>	
R	R Package	>	Select "R Package"
R	Shiny Web Application	>	
	R Package using Rcpp	>	
	R Package using RcppArmadillo	>	
	R Package using RcppEigen	>	
	R Package using RcppParallel	>	
		Cancel	





#### Create an R package

New Project Wizard		
Back	Create R Package	
R	Type: Package name: Package Create package based on source files: Add Remove	Type in package name
	Create project as subdirectory of: C:/Users/Nicolas/Desktop Browse	 Select the location of the package
	Create a git repository Use renv with this project	
		Select "Create
Open in new sessio	n Create Project Cancel	Project"





#### Congratulations! You created an R package\*!

	▼	Name	Size	Modified
	t			
$\Box$		man		
$\Box$		R		
$\Box$	R	test.Rproj	376 B	Mar 30, 2021, 10:28 AM
$\Box$		.Rbuildignore	30 B	Mar 30, 2021, 10:28 AM
$\Box$		DESCRIPTION	377 B	Mar 30, 2021, 10:28 AM
$\Box$		NAMESPACE	32 B	Mar 30, 2021, 10:28 AM



\*An empty package with the basic structure...





## Create an R package

An alternative way to create an R package with the same result is to use the **create\_package()** function from the **{usethis}** package.



This opens a new RStudio session with the new project loaded





#### Congratulations! You created an R package\*!

√ Creating 'C:/Users/Nicolas/Desktop/test/' √ Setting active project to 'C:/Users/Nicolas/Desktop/test' √ Creating 'R/' √ Writing 'DESCRIPTION' Package: test Title: What the Package Does (One Line, Title Case) Version: 0.0.0.9000 Authors@R (parsed): \* First Last <first.last@example.com> [aut, cre] (<https://orcid.org/YOUR-ORCID-ID>) Description: What the package does (one paragraph). License: `use\_mit\_license()`, `use\_gpl3\_license()` or friends to pick a license Encoding: UTF-8 LazyData: true Roxygen: list(markdown = TRUE) RoxygenNote: 7.1.1 √ Writing 'NAMESPACE' √ Writing 'test.Rproj' Adding 'Atest\\.Rproj\$' to '.Rbuildignore' √ Adding '.Rproj.user' to '.gitignore' ✓ Adding '^\\.Rproj\\.user\$' to '.Rbuildignore' √ Opening 'C:/Users/Nicolas/Desktop/test/' in new RStudio session Setting active project to '<no active project>'



\*An empty package with the basic structure...





## Live Coding Example 1 </>



- 1. Load the **{devtools}** package
- Create an empty R package called "myRpackage"

#### <u>Note:</u>

- The functionality of this package
   will be kept simple
- The main goal is to help the learning process and practice the package development cycle!





## File and Folder structure of an R package





The **man** folder will be missing when you use the **create\_package()** function but it will be automatically created with the first documentation step





### Document an R package

The function **document()** from **{devtools}** is used to build all the documentation for a package.

document() # CTRL + SHIFT + D

For an empty package this function will simply create the "man" folder if it does not exist.

> document()
Updating myRpackage documentation
Loading myRpackage





## Check an R package

The function **check()** from **{devtools}** automatically builds and checks a package. It runs through a number of checks and will return a summary of the check results.





Live Coding Example 2 </>



- 1. Load the **{devtools}** package
- 2. Document the package

#### {myRpackage}

- 3. Check the package
  - {myRpackage}





## **DESCRIPTION file**

#### The **DESCRIPTION** file is used to

store important **metadata** about the

package. For example:

- What is the package title
- What is the package version
- Who to contact
- Who can use it (the license)
- What other packages are needed for it to work

```
Package: test
Title: What the Package Does (One Line, Title Case)
Version: 0.0.0.9000
Authors@R:
    person(given = "First",
           family = "Last"
           role = c("aut", "cre"),
           email = "first.last@example.com",
           comment = c(ORCID = "YOUR-ORCID-ID"))
Description: What the package does (one paragraph).
License: `use_mit_license()`, `use_gpl3_license()` or friends to
    pick a license
Encoding: UTF-8
LazyData: true
Roxygen: list(markdown = TRUE)
RoxygenNote: 7.1.1
```





## DESCRIPTION fields

Below is some important guidelines to follow for the DESCRIPTION file fields.

Field	Description
Title	This is typically a one line description of the package. It should be <b>plain text</b> , not more than <b>65 characters</b> long, <b>capitalised like a title</b> , and <b>NOT end in a period</b> !
Description	This is a more detailed text about your package. It can be multiple sentences but it is <b>limited to one paragraph</b> , each line must be no more than 80 characters wide and new lines must be indented with 4 spaces!
Imports	The packages that are listed in this field <b>must</b> be installed on your computer for your package to work because they are being used by your package. If any package is missing, it <b>will</b> be <b>automatically installed</b> when your package is installed.
Suggests	The packages that are listed in this field can be used by your package but they are <b>not required</b> . If any package is missing, it <b>will not</b> be <b>automatically installed</b> when your package is installed.





## Package version

The **Version** field of the DESCRIPTION file shows the package's version number.

 Typically a version number is made up of three numbers:

```
<major>.<minor>.<patch>
```

For example: 1.3.0

A package that is in-development usually has ends with "9000"

For example: 1.3.0.9000



```
Package: devtools
Title: Tools to Make Developing R Packages Easier
Version: 2.3.2.9000
Authors@R:
    c(person(given = "Hadley",
             family = "Wickham",
             role = "aut"),
      person(given = "Jim",
             family = "Hester",
             role = c("aut", "cre"),
             email = "jim.hester@rstudio.com"),
      person(given = "Winston",
             family = "Chang",
             role = "aut"),
      person(given = "RStudio",
             role = "cph"),
      person(given = "R Core team",
             role = "ctb",
             comment = "Some namespace and vignette code extracted from base R"))
Description: Collection of package development tools.
License: GPL (>= 2)
URL: https://devtools.r-lib.org/, https://github.com/r-lib/devtools
BugReports: https://github.com/r-lib/devtools/issues
Depends:
    R (>= 3.0.2),
    usethis (>= 2.0.1)
```





Live Coding Example 3 </>



- Edit the DESCRIPTION file with the metadata about the package (*Title* and *Description*) and your details (*Author*)
- 2. Save the changes
- Document the package
   {myRpackage}
- 4. Check the package

#### {myRpackage}





## NAMESPACE file

The **NAMESPACE** file can be quite confusing and is considered an advanced topic (hence the warning!). In a simplified way, it is the file that controls the **communication between** packages and their functions. For example, it manages the **functions** to export (from your package) and functions to **import** (from other packages).



1	2	This	docum	nent	is	read	only.
---	---	------	-------	------	----	------	-------

2

3

# Generated by roxygen2: do not edit by hand





## What is an R function

An R function is an **R object** that contains code to be executed.

- In a simplified way, an R function takes inputs and generates outputs
- R functions are useful when we have code that is **repeated** in a script and therefore help to avoid "copy-pasting" code
- An R function ideally self-contains a complex piece of code and is dedicated to solving one task





## Components of an R function

An R function is created just like any other object in R and follows a specific structure.

		Arguments
Component	Description	
Name	The name of the function	<pre>add_two_numbers &lt;- function(a, b) {</pre>
Arguments	The values passed to the function (inputs)	total <- a + b Body
Body	R code that the function executes	return(total)
Return Value	The value/s the function is required to return (outputs)	<u>}</u>
		Return Value



The return value is the last expression evaluated by the function. If the **return()** expression is used then the function will output the contents of **return()** 



return(total)



## Create a function

```
The package {usethis} makes it easy for us to add R functions to a
package. The function use_r() takes as input the function name
and creates the ".R" file for that function inside the "R" folder.
# Create a function file in the R folder
use_r("add_two_numbers")
# Place the function code inside add_two_numbers.R
add_two_numbers <- function(a, b) {</pre>
  total <-a + b
```

}





Live Coding Example 4 </>



- Create the following functions for the package {myRpackage}
- add\_two\_numbers()
- subtract\_two\_numbers()





## Live Coding Example 4 </>

```
# Create a function file in the R folder
use_r("subtract_two_numbers")
```

```
# Place the function code inside add_two_numbers.R
subtract_two_numbers <- function(a, b) {</pre>
```

```
total <- a - b
```

```
return(total)
```





## Try out the function

Before we invest more time to properly document and test the function, it

is a good idea to first **check if the function works**!

There are three ways to do this:

- (Messy way): Create the function arguments as objects and then run the code inside the function line by line
- (Script way): Source the ".R" file containing the function and call the function

(Dev way): Use load\_all() and call the function – more on this later





## Live Coding Example 5 </>



Try out the function **add\_two\_numbers()** using the "Messy" and "Script" ways.





## Document a function

Function documentation can be somewhat fiddly to work with...

Typically each ".R" file in the "R" folder (containing a function) has its corresponding ".Rd" file in the "man" folder with the same name (containing the documentation).

The ".Rd" file that contains the **documentation** in an R-specific markup language ... luckily we do not have to worry about that!





## Document a function

What we need to do in order to document a function is write specially formatted comments above the function. These are called **roxygen** comments and the package **{roxygen2}** creates and edits the ".Rd" files for us!







## Roxygen comments

A **roxygen block** of comments above a function will determine the

help provided to the user about the function.

#'	#' Title #' #' Description	Comment	Description
# #'		Title	The first sentence and represents the title of the documentation
#' #'	anaran	Description	The second paragraph and describes what the package does
#'	#' @param #'	@param	Argument name followed by a description and what it does
#' #'	@return	@return	Describes the output of the function
#'	@export	@export	Specifies that the function is usable outside of package

**Step 1:** Open the ".R" file of the function

**Step 2:** Place cursor somewhere in the function code





Live Coding Example 6 </>



- Documenting the function add\_two\_numbers()
- Document the package
   {myRpackage}
- 3. Inspect help information of the function add\_two\_numbers()
- 4. Check the package

{myRpackage}





## Live Coding Example 6 </>

```
Add two numbers together
#'
#'
#'
   The purpose of this function is to take two numbers as inputs and add them
   together. The numbers can be positive or negative but not NA.
#'
#'
   @param a (numeric) A positive or negative number
#'
   @param b (numeric) A positive or negative number
#'
#'
   @return The total sum of the two numbers
#'
#'
#'
   @export
#' @examples
\#' add_two_numbers(a = 1, b = 1)
add_two_numbers <- function(a, b) {</pre>
  total <-a + b
  return(total)
}
```





## Live Coding Example 6 </>

add\_two\_numbers {demoConvertR}

R Documentation

#### Add two numbers together

Description

The purpose of this function is to take two numbers as inputs and add them together. The numbers can be positive or negative but not NA.

#### Usage

```
add_two_numbers(a, b)
```

#### Arguments

- a (numeric) A positive or negative number
- b (numeric) A positive or negative number

#### Value

The total sum of the two numbers

#### Examples

 $add_two_numbers(a = 1, b = 1)$ 







## Add tests to a package

The next step is to add **tests**. This is a **formal** way to test the functionality of your package and that your functions work as you expect them to! First we need to initialise testing for the package.

use\_testthat()

The function **use\_testthat()** from the **{usethis}** package adds *"Suggests: testthat"* to the DESCRIPTION file. It also creates the folders *"tests/testthat/" and adds a generic script "testthat.R" in the "tests" folder.* 



Spend some time to write unit tests for your functions!




## Test a function

The use\_test() function from {usethis} takes as input the function name and creates the "test-[name].R" file for that function inside the "testthat" folder with a generic test ready to edit.

**NOTE:** It is the developer's responsibility to write the unit tests.

#### Objects

expect\_equal() expect\_identical()

expect\_type() expect\_s3\_class()
expect\_s4\_class()

#### Vectors

expect\_length()

expect\_lt() expect\_lte() expect\_gt()
expect\_gte()

expect\_named()

expect\_setequal() expect\_mapequal()

expect\_true() expect\_false()

expect\_vector()



Check out more information here: <u>https://testthat.r-lib.org/reference/index.html</u>





### Test an R package

The function **test()** from the package **{devtools}** is used to run all of the tests of a package. It also prints out a test report about test **failures**, **warnings**, skipped tests and of course **passes**!

test() # CTRL + SHIFT + T







# Live Coding Example 7 </>



- Add testing to the package
   {myRpackage}
- 2. Write some unit tests for the function **add\_two\_numbers()**
- 3. Test the package {myRpackage}
- 4. Document the package
  - {myRpackage}
- 5. Check the package {myRpackage}





# Live Coding Example 7 </>

```
test_that("add_two_numbers returns the correct value and type", {
    expect_identical(object = add_two_numbers(a = 1, b = 2), expected = 3)
```

```
expect_identical(object = add_two_numbers(a = 1, b = -1), expected = 0)
```

```
expect_type(object = add_two_numbers(a = 1, b = 2), type = "double")
})
```

```
test_that("add_two_numbers returns NA if one of the arguments is NA", {
    expect_identical(object = add_two_numbers(a = NA, b = 2), expected = as.numeric(NA))
})
```

```
test_that("add_two_numbers returns Inf (+/-) if one of the arguments is infinite", {
    expect_identical(object = add_two_numbers(a = Inf, b = 2), expected = Inf)
```

```
expect_identical(object = add_two_numbers(a = -Inf, b = 2), expected = -Inf)
})
```





## Live Coding Example 7 </>

Add testing components

```
    Adding 'testthat' to Suggests field in DESCRIPTION
    Setting Config/testthat/edition field in DESCRIPTION to '3'
    Creating 'tests/testthat/'
    Writing 'tests/testthat.R'
```

Add test for function

```
√ Writing 'tests/testthat/test-add_two_numbers.R'
* Modify 'tests/testthat/test-add_two_numbers.R'
```

✤ Check DESCRIPTION file



#### Run tests



[ FAIL 0 | WARN 0 | SKIP 0 | PASS 6 ]







## Test drive a package

It is a good idea to **regularly** test drive the functionality that we develop in our package. This means **fewer bugs** to worry about! An important function to use during the development cycle is the **load\_all()** of the **{devtools}** package.







### Test drive a package

The function **load\_all()** is very useful because it allows you to **interact** with your package and its functions.

You can think of it as a way to **simulate** what happens when a package is installed and loaded with **library()**.

```
library(devtools)
```

# Working directory is set at the top level of package

```
load_all() # CTRL + SHIFT + L
```

Buil	d	Debug	Profile	Tools	Help
	Load All			Ctrl	+Shift+L
	Install and Restart			Ctrl	+Shift+B
	Clean and Rebuild				
	Test Package			Ctrl	+Shift+T
	Check Package		Ctrl	Ctrl+Shift+E	
	Build Source Package		e		
	Build Binary Package		e		
	D	ocument		Ctrl	+Shift+D
	St	op Build			
	Configure Build Tools			ls	





Live Coding Example 8 </>



1. Restart the R session

Go to → Session → Restart R

(or CTRL + SHIFT + F10)

 Test drive the {myRpackage} package using the load\_all() function





## Install a package

The function **install()** from the **{devtools}** package installs the package from the source state. More specifically the "R CMD INSTALL" command is executed behind the scenes.

Buil	d	Debug	Profile	Tools	Help		
	Load All			Ctrl	Ctrl+Shift+L		
		stall and l ean and F		Ctrl	+Shift+B		
	Te	est Packag	je	Ctrl	+Shift+T		
	Cł	neck Pack	age	Ctrl	Ctrl+Shift+E		
		uild Sourc uild Binan	_			ł	
	D	ocument		Ctrl	+Shift+D		
	St	op Build					
	Co	onfigure E	Build Too	ls			

#### \_ \_ \_ . . . \_ \_ .

#### install()

Alternatively the option "Install and Restart" (CTRL + SHIFT + B) installs the package, restarts R and loads it.





### Uninstall a package

An installed package is stored on your computer in a **library directory**.

Typically we update, however it is a good idea to know how to uninstall them.

The function **remove.packages()** from the **{utils}** package removes a package from the library directory.

remove.packages()

Ι.		

You can also navigate to the "Packages" tab, search for the package and click on the cross to uninstall the package

Files	Plots	Packages	Help	Viewer				- [
0 In	stall 🤇	Update				۹.		
N	Name		Desc	ription		Version		
User I	Library							
	additivity	/Tests		itivity Tes nbers	ts in the Two Way Anova with Single Sub-class	1.1-4		0
	askpass		Safe	Password	d Entry for R, Git, and SSH	1.1		$\otimes$
	asserttha	t	Easy	Pre and I	Post Assertions	0.2.1		$\otimes$
	backport	S	Rein	nplement	ations of Functions Introduced Since R-3.0.0	1.2.1		$\otimes$
	base64ei	nc	Тоо	ls for base	e64 encoding	0.1-3		$\otimes$
	bayesplo	t	Plot	ting for Ba	ayesian Models	1.7.2		$\otimes$
	BH		Boo	st C++ He	eader Files	1.72.0-3		$\otimes$
	biclust		BiCI	uster Algo	prithms	2.0.2		$\otimes$
	bitops		Bitw	ise Opera	tions	1.0-6		$\otimes$
	blob			mple S3 C OBS')	lass for Representing Vectors of Binary Data	1.2.1		8
	bmp		Rea	d Window	s Bitmap (BMP) Images	0.3		$\otimes$
	boundin	gbox	Crea	ate a Bour	nding Box in an Image	1.0.1		$\otimes$
	brew		Tem	plating Fr	amework for Report Generation	1.0-6		$\otimes$
	brio		Basi	c R Input	Output	1.1.0		$\otimes$
	broom		Con	vert Statis	tical Objects into Tidy Tibbles	0.7.3		$\otimes$
	C50		C5.0	Decision	Trees and Rule-Based Models	0.1.3.1		$\otimes$
	callr		Call	R from R		3.5.1		$\otimes$
	caTools		Тоо	ls: Moving	Window Statistics, GIF, Base64, ROC AUC, etc	1.18.0		$\otimes$
	colleanac		Tran	alata Care	addheat Call Danges to Dows and Columns	110	æ	







Live Coding Example 9 </>



- Install the **{myRpackage}** package
- 2. Try out the package
- Uninstall the {myRpackage}
   package





## Install a package from GitHub

We can also install an R package that is available on a **public GitHub repo**. The **{devtools}** package offers the function **install\_github()** which installs a package directly from GitHub (GitHub **username** and repository **name** are needed for this function).

For example:

install\_github("tidyverse/dplyr")



This is useful when a package is not available to install from CRAN or you want to install the latest in-development version of a package.





### Add the pipe operator to your package

The pipe operator (%>%) from the **{magrittr}** package is extremely useful when writing code and especially for data transformations. To use the pipe operator within a package that you are developing call the function **use\_pipe()** from {usethis} which carries out the necessary setup.

#### Description

Does setup necessary to use magrittr's pipe operator, %>% in your package. This function requires the use roxygen.

- Adds magrittr to "Imports" in DESCRIPTION.
- Imports the pipe operator specifically, which is necessary for internal use.
- Exports the pipe operator, if export = TRUE, which is necessary to make %>% available to the users of your package.





### Use other packages in your package

It is very likely that you will want to use functionality from another package within your package. To do this you need to **add** the "external" package to the "**Imports**" field of the **DESCRIPTION** file.

An easy way to do this is by using the function **use\_package()** from **{usethis}**. For example:

use\_package("dplyr")

This adds the **{dplyr}** package to the **"Imports**" field of the **DESCRIPTION** file. It is recommended to use the **::** operator when using functions from other packages. For example: **dplyr::filter()** 





## Keyboard Shortcuts

Below is the collection of keyboard shortcuts seen in this course that

helps speed up the development cycle of an R package.

Shortcut	Description
CTRL + SHIFT + D	Runs document() - build all the documentation for a package
CTRL + SHIFT + E	Runs <b>check()</b> – builds and checks a package
CTRL + ALT + SHIFT + R	Go to "Code" then "Insert Roxygen Skeleton"
CTRL + SHIFT + T	Runs <b>test()</b> – run all the tests of a package
CTRL + SHIFT + L	Runs load_all() - enables test driving a package
CTRL + SHIFT + F10	Go to "Session" then "Restart R" to Restart R session
CTRL + SHIFT + B	Installs the package, restarts R and loads it





### Package Development cheat sheet



https://github.com/rstudio/cheatsheets/raw/master/package-development.pdf





### Next online R event!

Build Interactive {shiny} Apps to Share Your Work With Anyone!



https://www.meetup.com/Oslo-useR-Group/events/277702734/





# Thank you to our sponsors and partners!



