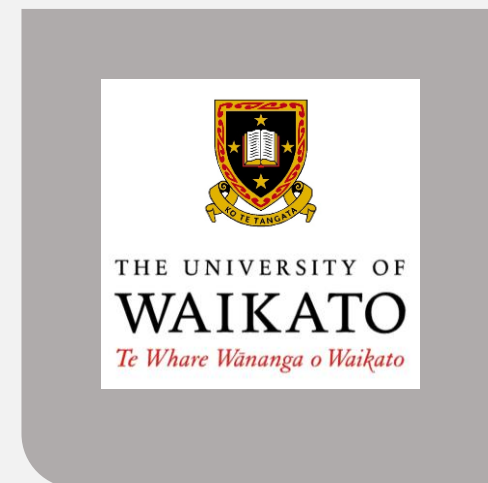
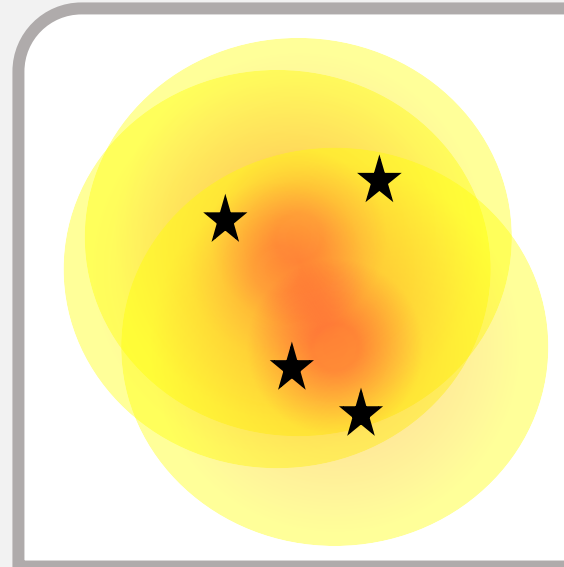


Writing R packages

How to write R packages that make life easier for you yourself and for others

NSCR Workshop 9 June 2022

Sophie Curtis-Ham



OVERVIEW

1. Why to write a package

× Github integration

2. How to write a package

× CRAN submission

- Resources

× Adding package data

- Workflow overview

× Testing in depth

- Demonstration

× ReadMe file

WHY?

- ✓ Reduce duplication for yourself
- ✓ Future you will thank past you for tidy, well documented code
- ✓ Share code with others
- ✓ Congratulations, you're now a software developer

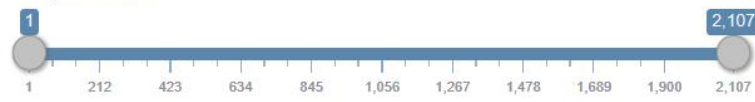
... and it's easier than you'd think!

WHY?

<https://github.com/Sophie-c-h/gpsmartr>

Person Identifier

Suspect Rank



Node Type

- home
- family_immediate
- family_ip
- family_other
- school
- work
- offence
- victim_witness
- incident
- police_other

Frequency

- freq_weekly
- freq_monthly
- freq_yearly

Recency

- rec_1to2d
- rec_3to30d
- rec_1to12m
- rec_1to5y
- rec_over5y

Duration

- dur_1to2d
- dur_3to30d
- dur_1to12m
- dur_1to5y
- dur_over5y

Same crime

- 0
- 1

Same location type

- 0
- 1

Same daypart

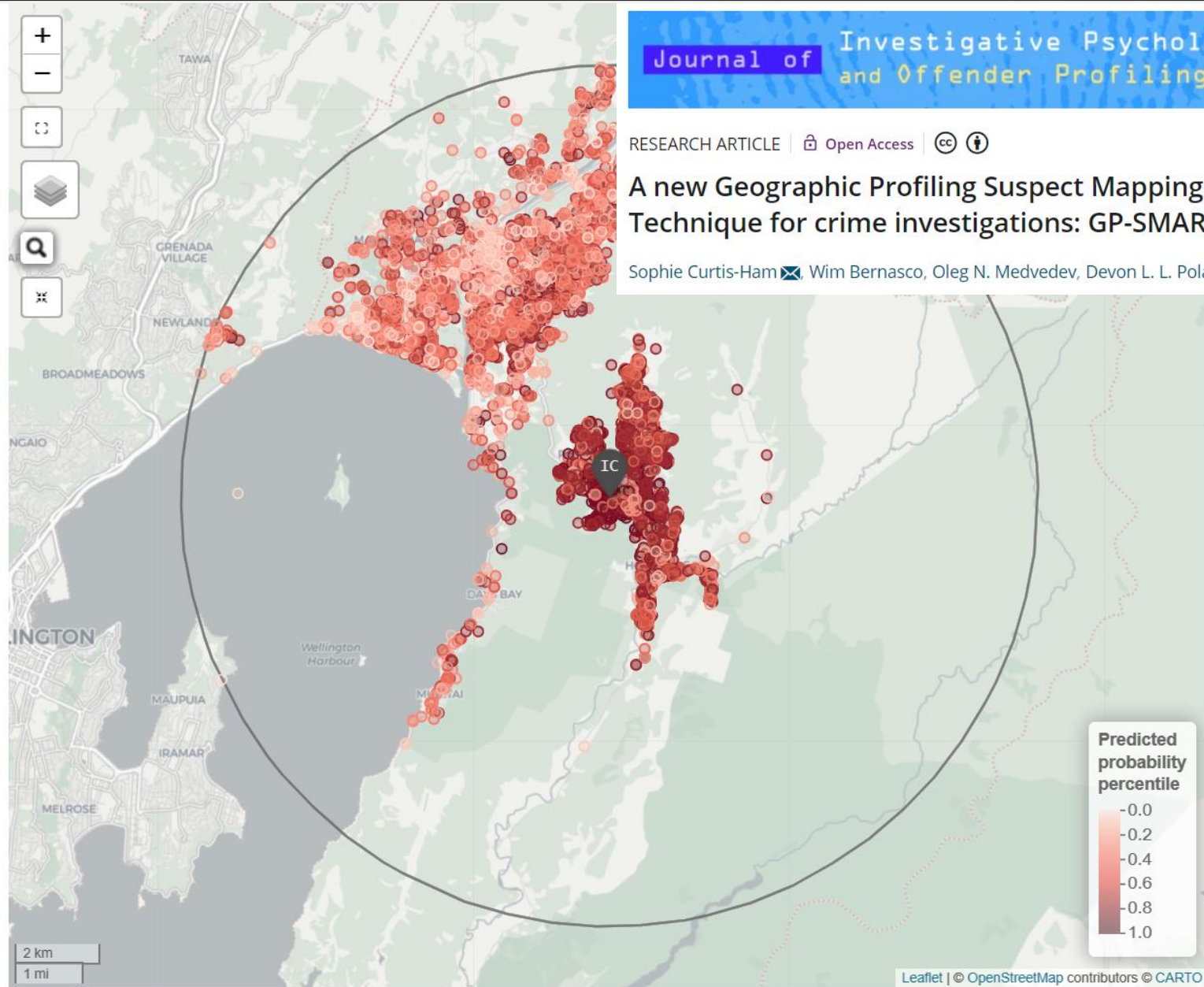
- 0
- 1

Same_weekpart

- 0
- 1

Same season

- 0
- 1

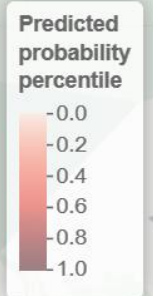


Journal of Investigative Psychology and Offender Profiling

RESEARCH ARTICLE | Open Access | CC BY

A new Geographic Profiling Suspect Mapping And Ranking Technique for crime investigations: GP-SMART

Sophie Curtis-Ham, Wim Bernasco, Oleg N. Medvedev, Devon L. L. Polaschek



HOW - RESOURCES

Videos

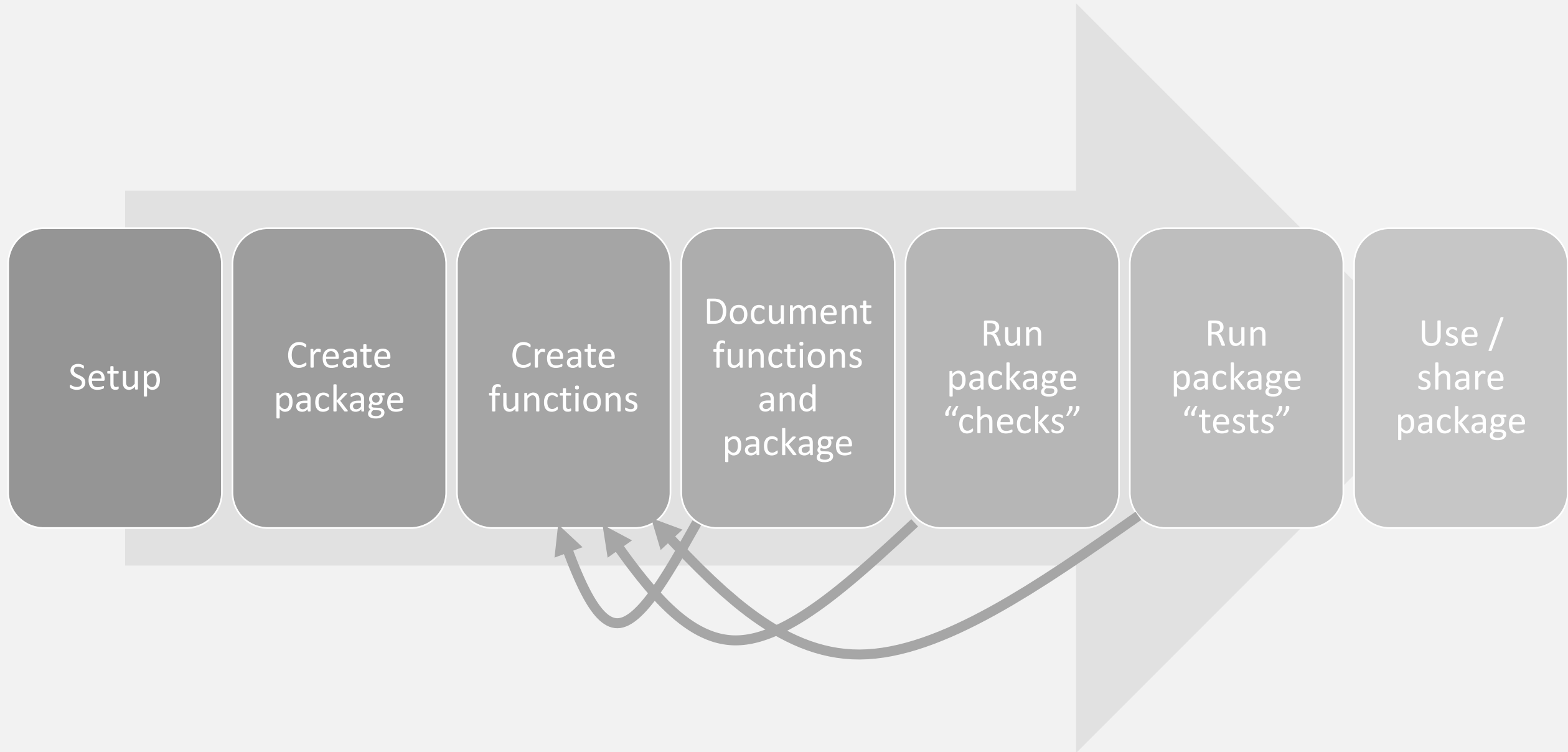
- <https://datawarrior.wordpress.com/2016/09/08/developing-r-packages/> quick video tutorial
- https://www.youtube.com/watch?v=47PN2VG3Rml&ab_channel=StatistikinDD create a package in 2 minutes!
- <https://www.rstudio.com/resources/rstudioconf-2018/you-can-make-a-package-in-20-minutes-jim-hester/>
- <https://www.youtube.com/watch?v=xcXzaEmZ-m4> and accompanying blog <https://www.pipinghotdata.com/posts/2020-10-25-your-first-r-package-in-1-hour/>
- <https://www.youtube.com/watch?v=OlrKRglSdc> another hour tutorial
- https://www.youtube.com/watch?v=1ZrjWKcG1C4&ab_channel=RStudio testthat tutorial

HOW - RESOURCES

Web books/blogs

- <https://hilaryparker.com/2014/04/29/writing-an-r-package-from-scratch/> blog with a nice simple example
- <https://r-pkgs.org/index.html> Hadley Wickham and Jenny Bryan's detailed guide to developing packages
- <https://support.rstudio.com/hc/en-us/articles/200486488-Developing-Packages-with-the-RStudio-IDE> RStudio guide
- <https://cran.r-project.org/doc/manuals/R-exts.html#Data-in-packages> R Core Team very detailed guide
- <https://bookdown.org/rdpeng/RProgDA/building-r-packages.html> Roger Peng chapters on packages and mapping functions
- <https://style.tidyverse.org/documentation.html> tidyverse style guide - covers both code and documentation
- <https://cran.r-project.org/web/packages/roxygen2/vignettes/rd-formatting.html> roxygen documentation guide
- <https://www.r-bloggers.com/2019/03/unit-tests-in-r/> more on constructing tests

HOW - WORKFLOW



HOW - DEMONSTRATION

Setup

Make sure you have these packages installed:

- devtools
- usethis
- tidyverse
- rlang
- broom
- here

For the demo, we will loosely follow these instructions (but create different functions):

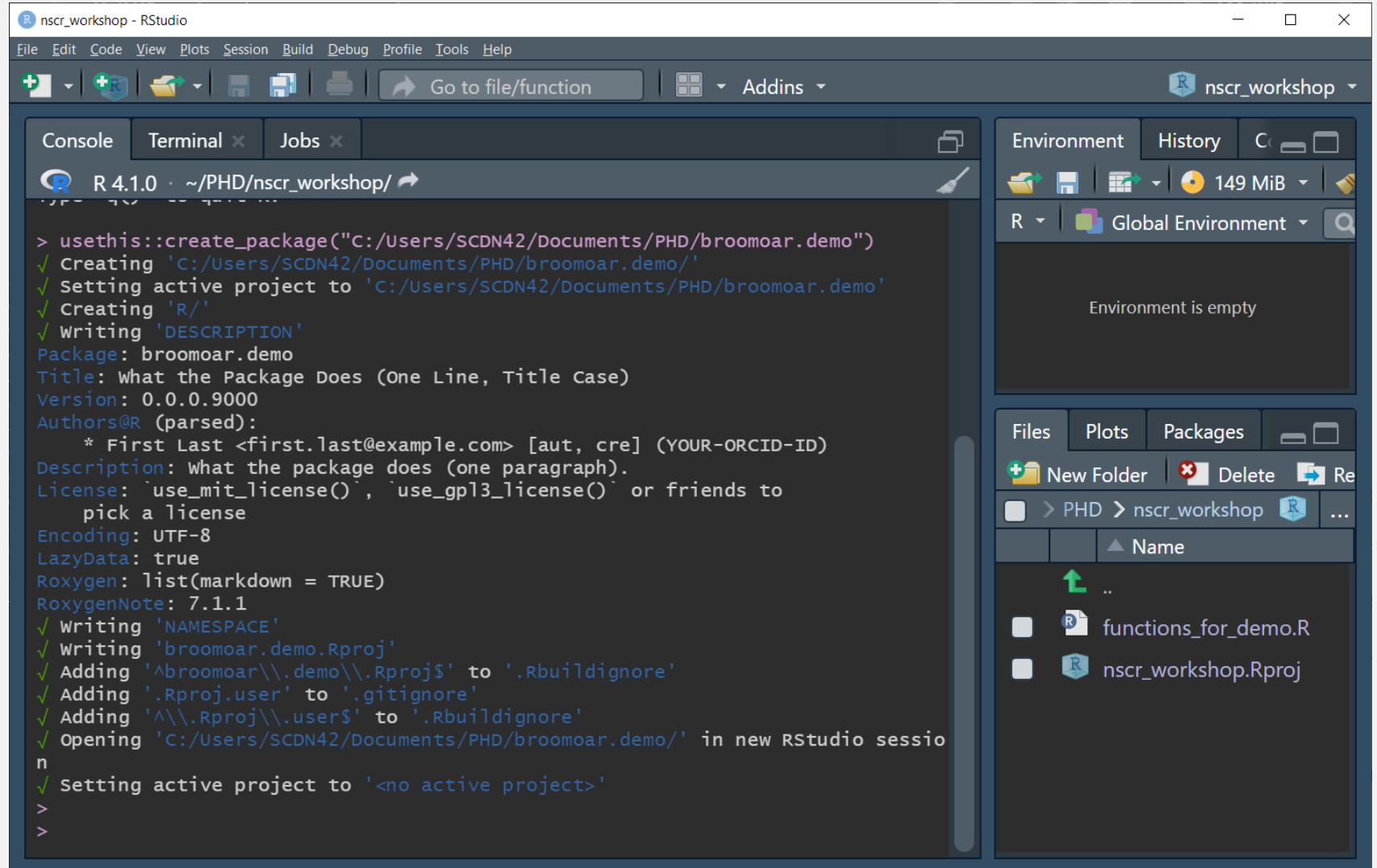
<https://www.pipinghotdata.com/posts/2020-10-25-your-first-r-package-in-1-hour/>

HOW - DEMONSTRATION

Grey background =
enter in console

Create
package

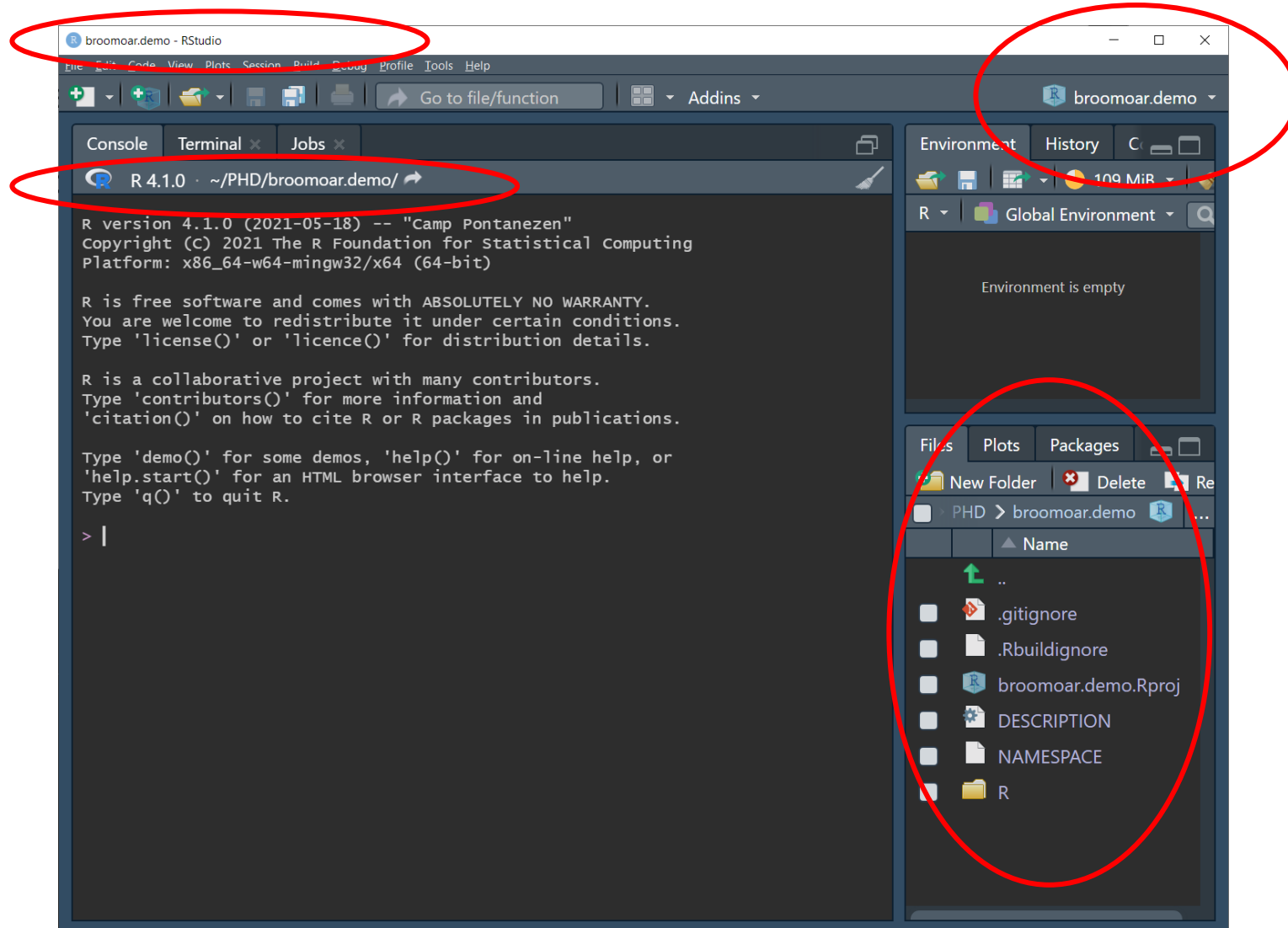
```
here::here()  
usethis::create_package("path/to/package/broomoar.demo")
```



The screenshot shows the RStudio interface with the console pane active. The console displays the output of the `usethis::create_package` function, which creates a new R package named `broomoar.demo`. The output includes details about the package name, title, version, authors, description, license, encoding, and the creation of project files like `NAMESPACE`, `broomoar.demo.Rproj`, and `.Rproj.user`. The console also shows the creation of a new RStudio session for the package and the setting of the active project to `<no active project>`.

```
R 4.1.0 · ~/PHD/nscr_workshop/  
> usethis::create_package("C:/Users/SCDN42/Documents/PHD/broomoar.demo")  
✓ Creating 'C:/Users/SCDN42/Documents/PHD/broomoar.demo/'  
✓ Setting active project to 'C:/Users/SCDN42/Documents/PHD/broomoar.demo/'  
✓ Creating 'R/'  
✓ Writing 'DESCRIPTION'  
Package: broomoar.demo  
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] (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 'broomoar.demo.Rproj'  
✓ Adding '^broomoar\\.demo\\.Rproj$' to '.Rbuildignore'  
✓ Adding '.Rproj.user' to '.gitignore'  
✓ Adding '\\\\.Rproj\\.user$' to '.Rbuildignore'  
✓ Opening 'C:/Users/SCDN42/Documents/PHD/broomoar.demo/' in new RStudio session  
n  
✓ Setting active project to '<no active project>'  
>  
>
```

HOW - DEMONSTRATION



HOW - DEMONSTRATION

Run check



```
devtools::check()

broomoar.demo - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ Addins Go to file/function broomoar.demo

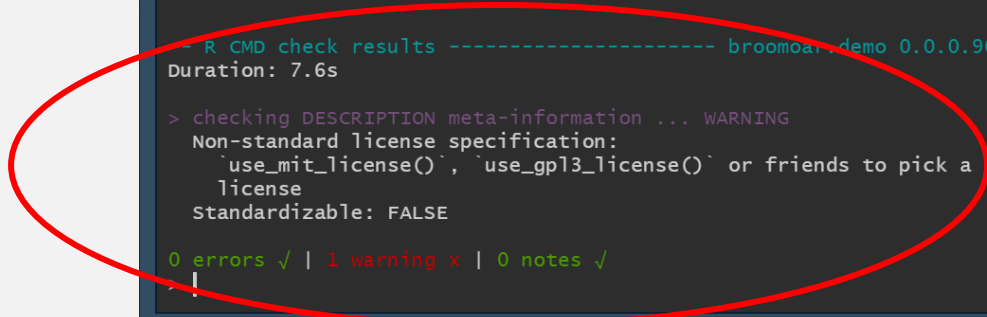
Console Terminal Jobs
R 4.1.0 ~/PHD/broomoar.demo/
✓ checking for future file timestamps ...
W checking DESCRIPTION meta-information (422ms)
Non-standard license specification:
  'use_mit_license()', 'use_gp13_license()' or friends to pick a
  license
Standardizable: FALSE
✓ checking top-level files
✓ checking for left-over files ...
✓ checking index information
✓ checking package subdirectories ...
✓ checking whether the package can be loaded ...
✓ checking whether the package can be loaded with stated dependencies ...
✓ checking whether the package can be unloaded cleanly ...
✓ checking whether the namespace can be loaded with stated dependencies ...
✓ checking whether the namespace can be unloaded cleanly ...
✓ checking loading without being on the library search path ...
- checking examples ... NONE
✓ checking for non-standard things in the check directory
✓ checking for detritus in the temp directory

See
'c:/Users/SCDN42/AppData/Local/Temp/RtmpU1spCa/broomoar.demo.Rcheck/00che
ck.log'
for details.

R CMD check results ----- broomoar.demo 0.0.0.9000 ----
Duration: 7.6s

> checking DESCRIPTION meta-information ... WARNING
Non-standard license specification:
  'use_mit_license()', 'use_gp13_license()' or friends to pick a
  license
Standardizable: FALSE

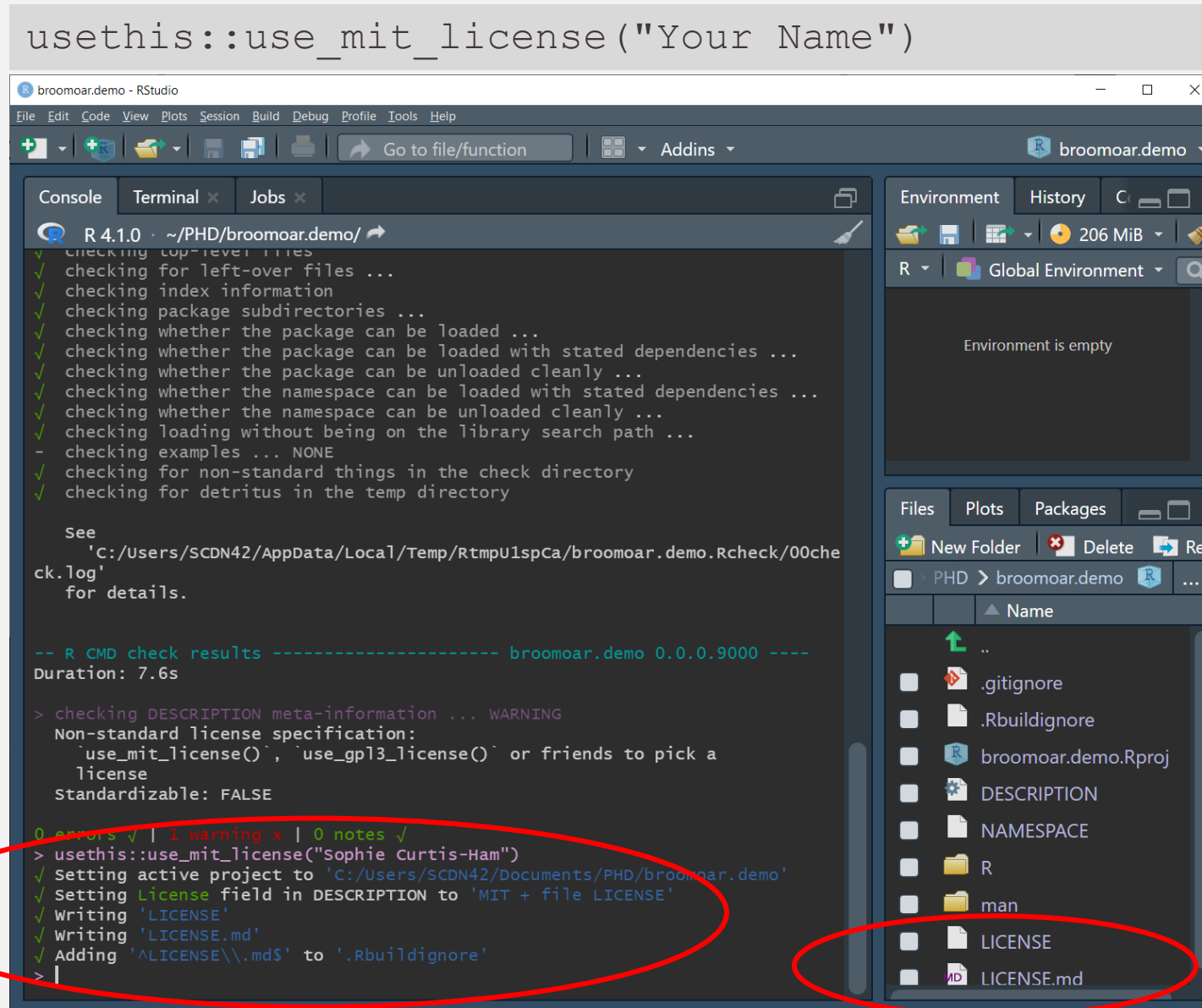
0 errors ✓ | 1 warning x | 0 notes ✓
```



HOW - DEMONSTRATION

Add licence

```
usethis::use_mit_license("Your Name")
```



The screenshot displays the RStudio interface for a project named 'broomoar.demo'. The console window shows the execution of the `usethis::use_mit_license("Your Name")` command. The output indicates that the package is being checked for various requirements, and a warning is issued regarding the non-standard license specification. The command successfully sets the active project, updates the DESCRIPTION file with the license information, and creates the LICENSE and LICENSE.md files. The Files pane on the right shows the project structure, with the newly created LICENSE and LICENSE.md files highlighted by red circles.

```
✓ checking top-level files
✓ checking for left-over files ...
✓ checking index information
✓ checking package subdirectories ...
✓ checking whether the package can be loaded ...
✓ checking whether the package can be loaded with stated dependencies ...
✓ checking whether the package can be unloaded cleanly ...
✓ checking whether the namespace can be loaded with stated dependencies ...
✓ checking whether the namespace can be unloaded cleanly ...
✓ checking loading without being on the library search path ...
- checking examples ... NONE
✓ checking for non-standard things in the check directory
✓ checking for detritus in the temp directory

See
'C:/Users/SCDN42/AppData/Local/Temp/RtmpU1spCa/broomoar.demo.Rcheck/00che
ck.log'
for details.

-- R CMD check results ----- broomoar.demo 0.0.0.9000 ----
Duration: 7.6s

> checking DESCRIPTION meta-information ... WARNING
Non-standard license specification:
`use_mit_license()`, `use_gpl3_license()` or friends to pick a
license
Standardizable: FALSE

0 errors ✓ | 1 warning x | 0 notes ✓
> usethis::use_mit_license("Sophie Curtis-Ham")
✓ Setting active project to 'C:/Users/SCDN42/Documents/PHD/broomoar.demo'
✓ Setting License field in DESCRIPTION to 'MIT + file LICENSE'
✓ Writing 'LICENSE'
✓ Writing 'LICENSE.md'
✓ Adding '^LICENSE\\.md$' to '.Rbuildignore'
>
```

HOW - DEMONSTRATION

Create
function

```
fn_sumstats <- function(model) {  
  # tidy the summary statistics  
  broom::glance(model) %>%  
  # add another summary statistic  
  mutate(pseudo_r2 = 1-(deviance/null.deviance))  
}  
usethis::use_r("fn_sumstats")
```

THEN COPY AND PASTE THE FUNCTION TO THE NEW BLANK R SCRIPT
THEN SAVE THE SCRIPT

HOW - DEMONSTRATION

The screenshot displays the RStudio interface for a project named 'broomoar.demo'. The main editor window shows the source code for a function named 'fn_sumstats.R'. The code is as follows:

```
1 fn_sumstats <- function(model) {  
2   # tidy the summary statistics  
3   broom::glance(model) %>%  
4   # add another summary statistic  
5   mutate(pseudo_r2 = 1-(deviance/null.deviance))  
6 }  
7
```

The Environment pane on the right shows the 'Global Environment' with a function 'fn_su...' listed under the 'Functions' section. The Files pane on the right shows the project structure, with the file 'fn_sumstats.R' circled in red. The Console pane at the bottom shows the following commands and output:

```
> usethis::use_r("fn_sumstats")  
* Modify 'R/fn_sumstats.R'  
* Call 'use_test()' to create a matching test file
```

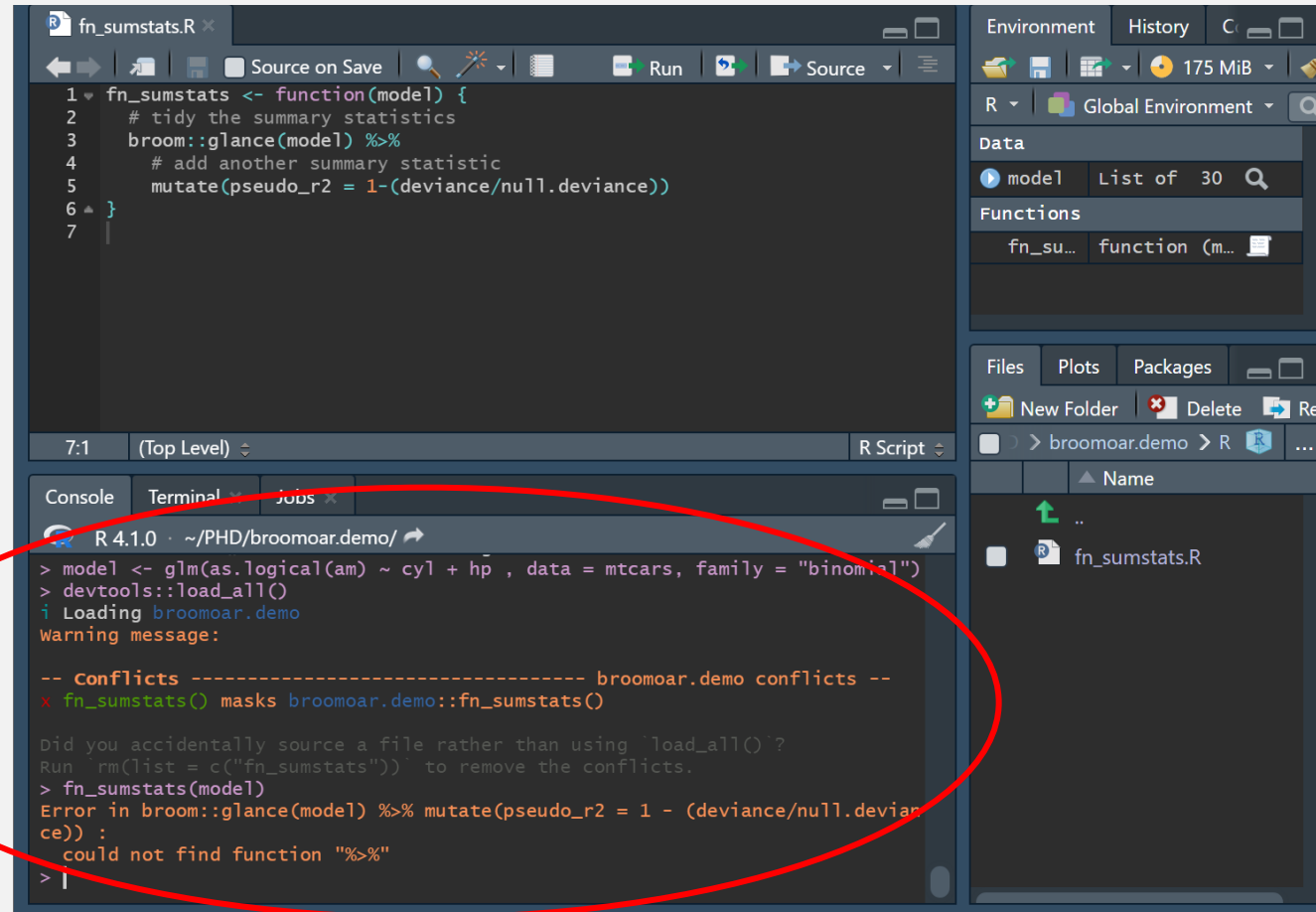
Red circles highlight the function code in the editor, the 'fn_sumstats.R' file in the Files pane, and the console output.

HOW - DEMONSTRATION

Execute

function

```
model <- glm(as.logical(am) ~ cyl + hp,  
            data = mtcars, family = "binomial")  
devtools::load_all()  
fn_sumstats(model)  
rm(fn_sumstats) # if you get the 'conflicts' error
```



```
1 fn_sumstats <- function(model) {  
2   # tidy the summary statistics  
3   broom::glance(model) %>%  
4   # add another summary statistic  
5     mutate(pseudo_r2 = 1 - (deviance/null.deviance))  
6 }  
7
```

Console

```
R 4.1.0 · ~/PHD/broomoar.demo/  
> model <- glm(as.logical(am) ~ cyl + hp , data = mtcars, family = "binomial")  
> devtools::load_all()  
i Loading broomoar.demo  
warning message:  
  
-- Conflicts ----- broomoar.demo conflicts --  
x fn_sumstats() masks broomoar.demo::fn_sumstats()  
  
Did you accidentally source a file rather than using 'load_all()'?  
Run 'rm(list = c("fn_sumstats"))' to remove the conflicts.  
> fn_sumstats(model)  
Error in broom::glance(model) %>% mutate(pseudo_r2 = 1 - (deviance/null.devian  
ce)) :  
  could not find function "%>%"  
>
```



HOW - DEMONSTRATION

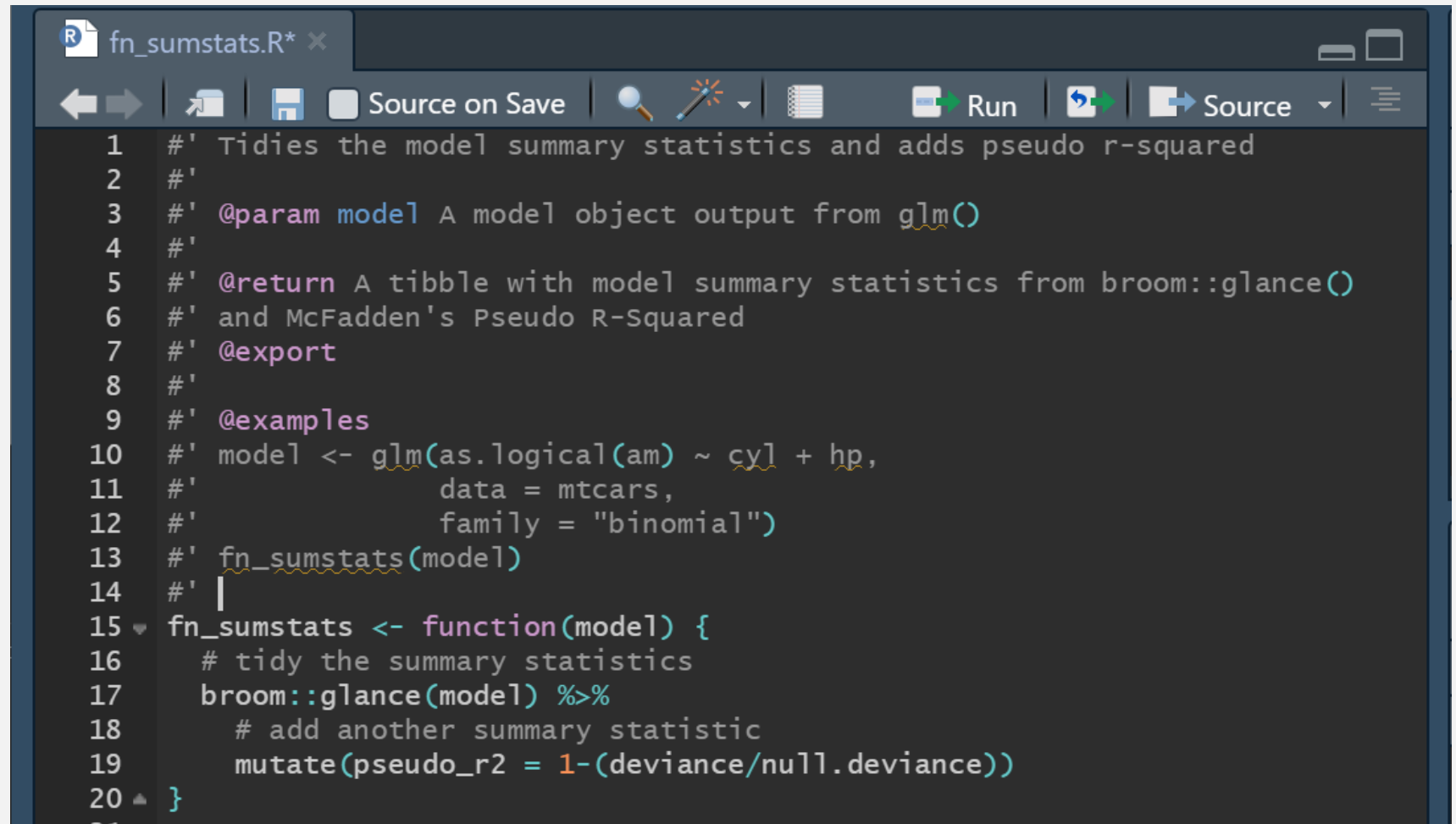
PUT CURSOR INSIDE THE FUNCTION IN THE SCRIPT

Code > Insert Roxygen Skeleton

FILL IN DOCUMENTATION DETAILS

Document

function



```
fn_sumstats.R* x
Source on Save Run Source
1 #' Tidies the model summary statistics and adds pseudo r-squared
2 #'
3 #' @param model A model object output from glm()
4 #'
5 #' @return A tibble with model summary statistics from broom::glance()
6 #' and McFadden's Pseudo R-Squared
7 #' @export
8 #'
9 #' @examples
10 #' model <- glm(as.logical(am) ~ cyl + hp,
11 #'             data = mtcars,
12 #'             family = "binomial")
13 #' fn_sumstats(model)
14 #'
15 fn_sumstats <- function(model) {
16   # tidy the summary statistics
17   broom::glance(model) %>%
18     # add another summary statistic
19     mutate(pseudo_r2 = 1-(deviance/null.deviance))
20 }
```


HOW - DEMONSTRATION

```
usethis::use_package("broom")  
usethis::use_package("dplyr")
```

Document
function

The screenshot displays the RStudio interface with the following components:

- Editor Window:** Shows the `DESCRIPTION` file with the following content:

```
3 Version: 0.0.0.9000  
4 Authors@R:  
5   person(given = "First",  
6         family = "Last",  
7         role = c("aut", "cre"),  
8         email = "first.last@example.com",  
9         comment = c(ORCID = "YOUR-ORCID-ID"))  
10 Description: What the package does (one paragraph).  
11 License: MIT + file LICENSE  
12 Encoding: UTF-8  
13 LazyData: true  
14 Roxygen: list(markdown = TRUE)  
15 RoxygenNote: 7.1.1  
16 Imports:  
17   broom,  
18   dplyr  
19
```
- Console:** Shows the execution of the following commands and their output:

```
> usethis::use_package("broom")  
✓ Adding 'broom' to Imports field in DESCRIPTION  
* Refer to functions with broom::fun()  
> usethis::use_package("dplyr")  
✓ Adding 'dplyr' to Imports field in DESCRIPTION  
* Refer to functions with dplyr::fun()  
>
```
- File Explorer:** Shows the project structure with the `DESCRIPTION` file highlighted.

HOW - DEMONSTRATION

```
usethis::use_pipe()  
devtools::document()
```

Document
function

The screenshot displays the R IDE interface. At the top, the DESCRIPTION file is open, showing the following content:

```
10 Description: what the package does (one paragraph).  
11 License: MIT + file LICENSE  
12 Encoding: UTF-8  
13 LazyData: true  
14 Roxygen: list(markdown = TRUE)  
15 RoxygenNote: 7.1.1  
16 Imports:  
17   broom,  
18   dplyr,  
19   magrittr  
20
```

The console output shows the following commands and their results:

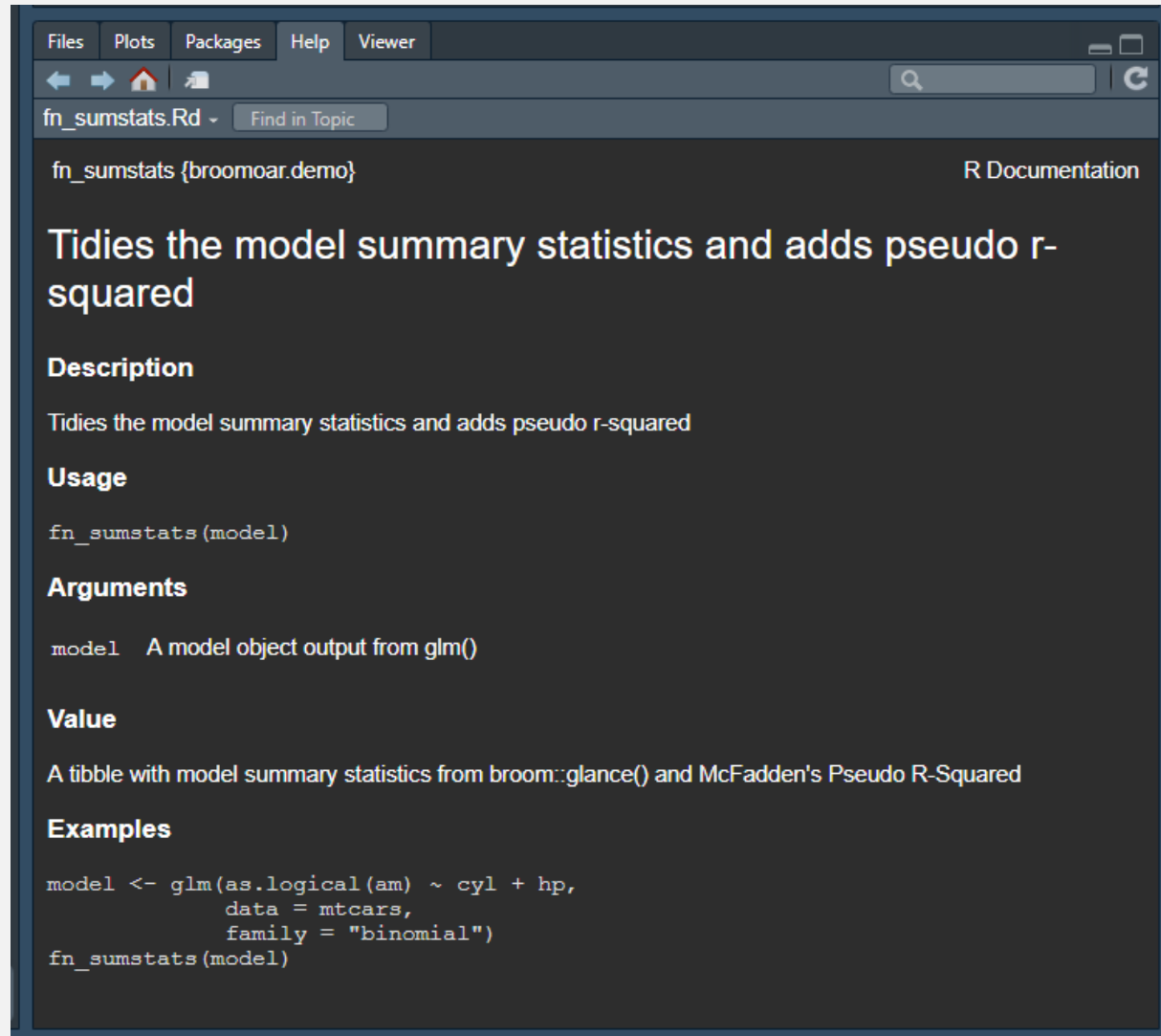
```
R 4.1.0 ~ /PHD/broomoar.demo/  
could not find function "%>%"  
> usethis::use_package("broom")  
✓ Adding 'broom' to Imports field in DESCRIPTION  
* Refer to functions with 'broom::fun()'   
> usethis::use_package("dplyr")  
✓ Adding 'dplyr' to Imports field in DESCRIPTION  
* Refer to functions with 'dplyr::fun()'   
> usethis::use_pipe()  
✓ Adding 'magrittr' to Imports field in DESCRIPTION  
✓ Writing 'R/utils-pipe.R'  
* Run 'devtools::document()' to update 'NAMESPACE'  
> devtools::document()  
i Updating broomoar.demo documentation  
i Loading broomoar.demo  
Writing NAMESPACE  
Writing NAMESPACE  
Writing fn_sumstats.Rd  
Writing pipe.Rd
```

The IDE interface also shows the Environment pane with 157 MiB of memory used, the Data pane with a model object, and the Files pane showing the project structure with files like 'fn_sumstats.Rd' and 'pipe.Rd' circled in red.

HOW - DEMONSTRATION

Document
function

```
?fn_sumstats
```



The screenshot shows the R Documentation window for the `fn_sumstats` function. The window title is `fn_sumstats.Rd` and it includes a search bar and navigation icons. The content is as follows:

`fn_sumstats {broomar.demo}` R Documentation

Tidies the model summary statistics and adds pseudo r-squared

Description

Tidies the model summary statistics and adds pseudo r-squared

Usage

```
fn_sumstats(model)
```

Arguments

`model` A model object output from `glm()`

Value

A tibble with model summary statistics from `broom::glance()` and McFadden's Pseudo R-Squared

Examples

```
model <- glm(as.logical(am) ~ cyl + hp,
             data = mtcars,
             family = "binomial")
fn_sumstats(model)
```

HOW - DEMONSTRATION

```
1 #' Run GP-SMART for an input crime and input suspects
2 #'
3 #' @description
4 #' Calculates suspect variables relative to the input crime,
5 #' calculates probability of crime per suspect activity location (node) within the user's input search radius,
6 #' ranks suspects according to their highest probability activity node,
7 #' and returns ranked suspects and (optionally) their activity nodes for mapping.
8 #'
9 #' @param input_crime A data frame with 1 row and 16 columns, created by [fn_prepare_input_crime()].
10 #'
11 #' @param input_suspects A data frame with 1 row per suspect activity node and 21 columns, created by [fn_prepare_suspect_data()].
12 #'
13 #' @param search_radius The distance (in kilometres) around the input crime to search for suspect nodes. Suspects with nodes within this distance are returned.
14 #'
15 #' @param weights A list of 8 data frames, one per weighting variable (node attribute),
16 #' specifying the weight to use for each value of each attribute for each input crime type.
17 #' If 'NULL', the default, built-in weights will be used (see Details).
18 #'
19 #' @param return_node_predictions Logical. Whether to return per-node predictions for the shortlisted suspects, required as input for [fn_map_gpsmart_output()].
20 #'
21 #' @author Sophie Curtis-Ham
22 #'
23 #' @importFrom rlang .data
24 #' @importFrom utils data
25 #'
26 #' @rawNamespace import(data.table, except = c(month, hour, quarter, week, year, wday, second, minute, mday, yday, isoweek))
27 #'
28 #' @return A list with four elements:
29 #'
30 #' \describe{
31 #'   \item{`suspects_ranked`} {A data frame with n rows = n suspects with nodes within the `search_radius` of the input crime, ranked by their maximum predicted crime probability.}
32 #'   \item{`node_predictions`} {If `return_node_predictions` = `TRUE`, a data frame with n rows = n suspect nodes within search radius, including all node attributes and predicted crime probability per node; if `return_node_predictions` = `FALSE`, a character string "no node predictions".}
33 #'   \item{`input_crime`} {The data frame entered as input for [fn_gpsmart()].}
34 #'   \item{`search_radius`} {The search radius (in kilometres), entered as input for [fn_gpsmart()].}
35 #' }
36 #' It is included in the output for use in the [fn_map_gpsmart_output()] function.
37 #'
38 #' \item{`season_variables`} {The search radius (in kilometres), entered as input for [fn_gpsmart()].}
39 #' It is included in the output for use in the [fn_map_gpsmart_output()] function.
40 #' }
41 #'
42 #' @export
43 #'
44 #' @details
45 #' \itemize{
46 #'   \item{In calculating the probability of crime per suspect activity node, the function weights the node (adjusts the probability) based on its values across 7 attributes (weighting variables). The weighting variables are: frequency, recency, duration, behaviour similarity, location type similarity, daypart similarity, weekpart similarity. The default (built in) weights were estimated using a New Zealand sample of offenders and activity locations as described in Curtis-Ham et al (2022). To create your own weights, use the built in list of data frames as a template, replacing only the weight values. (use \code{gpsmartr::private::node_weights_lookups} to access the built in weights).}
47 #'   \item{The function will return an error if there are no suspect nodes within the search radius.}
48 #'   \item{Season variables are based on Southern Hemisphere.}
49 #' }
50 #'
51 #' @seealso
52 #' * [fn_prepare_input_crime()] checks that the minimum necessary variables are present and creates the `input_crime` data frame for use in this function.
53 #' * [fn_prepare_suspect_data()] checks that the minimum necessary variables are present and creates the `input_suspects` data frame for use in this function.
54 #' * [fn_map_gpsmart_output()] creates an interactive map visualising the output of this function.
55 #'
56 #' @references
57 #' * Curtis-Ham S., Bernasco, W., Medvedev, O. N., & Polaschek, D. L. L. (2022). 'A new geographic profiling method for mapping and ranking suspects in crime investigations: GP-SMART'. Journal of Investigative Psychology and Offender Profiling. https://doi.org/10.1002/jip.1585
58 #'
59 #' @examples
60 #' \dontrun{
61 #' data(example_input_crime_raw)
62 #' data(example_input_suspects_raw)
63 #' fn_gpsmart(
64 #'   input_crime = fn_prepare_input_crime(example_input_crime_raw),
65 #'   input_suspects = fn_prepare_suspect_data(example_input_suspects_raw),
66 #'   search_radius = 10,
67 #'   return_node_predictions = FALSE
68 #' )
69 #' }
70 #'
71 #' fn_gpsmart <- function(input_crime, input_suspects, search_radius = 10, weights = NULL, return_node_predictions = FALSE) {
```



```
fn_gpsmartRd - Find in Topic
R Documentation

fn_gpsmart(gpsmart)

Run GP-SMART for an input crime and input suspects

Description
Calculates suspect variables relative to the input crime, calculates probability of crime per suspect activity location (node) within the user's input search radius, ranks suspects according to their highest probability activity node, and returns ranked suspects and (optionally) their activity nodes for mapping.

Usage
fn_gpsmart(
  input_crime,
  input_suspects,
  search_radius = 10,
  weights = NULL,
  return_node_predictions = FALSE
)

Arguments
input_crime      A data frame with 1 row and 16 columns, created by fn_prepare_input_crime().
input_suspects  A data frame with 1 row per suspect activity node and 21 columns, created by fn_prepare_suspect_data().
search_radius   The distance (in kilometres) around the input crime to search for suspect nodes. Suspects with nodes within this distance are shortlisted. The default is 10km.
weights         A list of 8 data frames, one per weighting variable (node attribute), specifying the weight to use for each value of each attribute for each input crime type. If NULL, the default, built-in weights will be used (see Details).
return_node_predictions Logical. Whether to return per-node predictions for the shortlisted suspects, required as input for fn_map_gpsmart_output().

Details
• In calculating the probability of crime per suspect activity node, the function weights the node (adjusts the probability) based on its values across 7 attributes (weighting variables). The weighting variables are: frequency, recency, duration, behaviour similarity, location type similarity, daypart similarity, weekpart similarity and season similarity. The default (built in) weights were estimated using a New Zealand sample of offenders and activity locations as described in Curtis-Ham et al (2022). To create your own weights, use the list in data frames as a template, replacing only the weight values. (use gpsmartr::private::node_weights_lookups to access the built in weights).
• The function will return an error if there are no suspect nodes within the search radius.
• Season variables are based on Southern Hemisphere.

Value
A list with four elements:
suspects_ranked
  A data frame with n rows = n suspects with nodes within the search_radius of the input crime, ranked by their maximum predicted crime probability.
node_predictions
  If return_node_predictions = TRUE, a data frame with n rows = n suspect nodes within search radius, including all node attributes and predicted crime probability per node; if return_node_predictions = FALSE, a character string "no node predictions".
input_crime
  The data frame entered as input for fn_gpsmart(). It is included in the output for use in the fn_map_gpsmart_output() function.
search_radius
  The search radius (in kilometres), entered as input for fn_gpsmart(). It is included in the output for use in the fn_map_gpsmart_output() function.

Author(s)
Sophie Curtis-Ham

References
Curtis-Ham S, Bernasco W, Medvedev O N, & Polaschek D L L (2022) 'A new geographic profiling method for mapping and ranking suspects in crime investigations. GP-SMART'. Journal of Investigative Psychology and Offender Profiling. https://doi.org/10.1002/jip.1585

See Also
• fn_prepare_input_crime() checks that the minimum necessary variables are present and creates the input_crime data frame for use in this function.
• fn_prepare_suspect_data() checks that the minimum necessary variables are present and creates the input_suspects data frame for use in this function.
• fn_map_gpsmart_output() creates an interactive map visualising the output of this function.

Examples
## Not run:
data(example_input_crime_raw)
data(example_input_suspects_raw)
fn_gpsmart(
  input_crime = fn_prepare_input_crime(example_input_crime_raw),
  input_suspects = fn_prepare_suspect_data(example_input_suspects_raw),
  search_radius = 10,
  return_node_predictions = FALSE
)
```

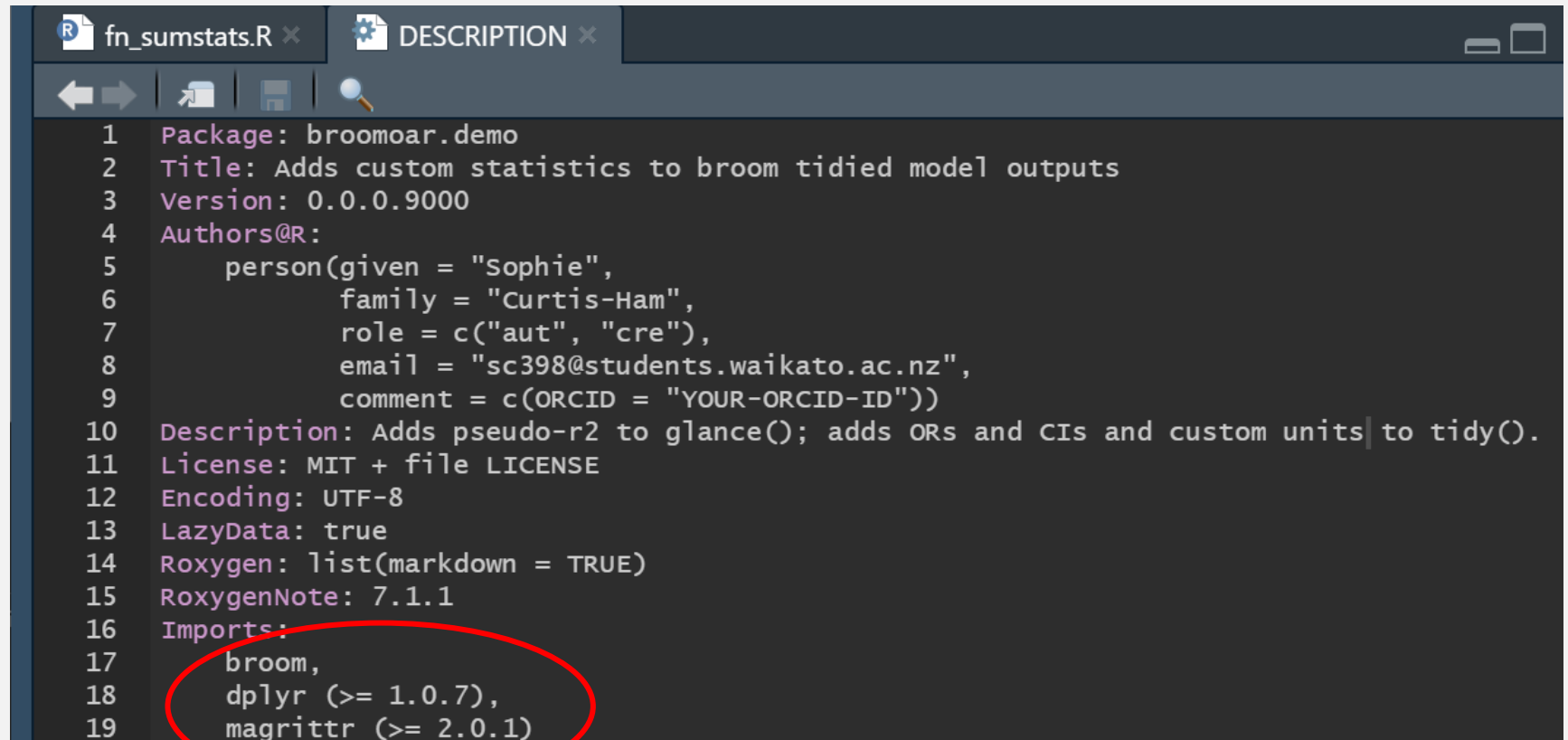
HOW - DEMONSTRATION

FILL IN DESCRIPTION DETAILS

Document

package

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



```
1 Package: broomoar.demo
2 Title: Adds custom statistics to broom tidied model outputs
3 Version: 0.0.0.9000
4 Authors@R:
5   person(given = "Sophie",
6           family = "Curtis-Ham",
7           role = c("aut", "cre"),
8           email = "sc398@students.waikato.ac.nz",
9           comment = c(ORCID = "YOUR-ORCID-ID"))
10 Description: Adds pseudo-r2 to glance(); adds ORs and CIs and custom units to tidy().
11 License: MIT + file LICENSE
12 Encoding: UTF-8
13 LazyData: true
14 Roxygen: list(markdown = TRUE)
15 RoxygenNote: 7.1.1
16 Imports:
17   broom,
18   dplyr (>= 1.0.7),
19   magrittr (>= 2.0.1)
```

HOW - DEMONSTRATION

Run
check



```
devtools::check(cran = FALSE)
```

```
> checking examples ... ERROR
Running examples in 'broomoar.demo-Ex.R' failed
The error most likely occurred in:

> base::assign(".ptime", proc.time(), pos = "CheckExEnv")
> ### Name: fn_sumstats
> ### Title: Tidies the model summary statistics and adds pseudo r-squared
> ### Aliases: fn_sumstats
>
> ### ** Examples
>
> model <- glm(as.logical(am) ~ cyl + hp,
+             data = mtcars,
+             family = "binomial")
> fn_sumstats(model)
Error in mutate(., pseudo_r2 = 1 - (deviance/null.deviance)) :
could not find function "mutate"
Calls: fn_sumstats -> %>%
Execution halted

> checking dependencies in R code ... NOTE
Namespace in Imports field not imported from: 'dplyr'
All declared Imports should be used.

> checking R code for possible problems ... NOTE
fn_sumstats: no visible global function definition for 'mutate'
fn_sumstats: no visible binding for global variable 'deviance'
fn_sumstats: no visible binding for global variable 'null.deviance'
Undefined global functions or variables:
  deviance mutate null.deviance
Consider adding
  importFrom("stats", "deviance")
to your NAMESPACE file.

1 error x | 0 warnings v | 2 notes x
```

HOW - DEMONSTRATION

ADJUST ROXYGEN HEADER

ADD rlang TO DESCRIPTION IMPORTS [NOT SHOWN]

ADD .data\$ TO FUNCTION

Fix

issues

```
fn_sumstats.R x DESCRIPTION x
Source on Save Run Source
1 #' Tidies the model summary statistics and adds pseudo r-squared
2 #'
3 #' @param model A model object output from glm()
4 #'
5 #' @return A tibble with model summary statistics from broom::glance()
6 #' and McFadden's Pseudo R-Squared
7 #' @export
8 #' @importFrom dplyr mutate
9 #' @importFrom rlang .data
10 #'
11 #' @examples
12 #' model <- glm(as.logical(am) ~ cyl + hp,
13 #'             data = mtcars,
14 #'             family = "binomial")
15 #' fn_sumstats(model)
16 #'
17 fn_sumstats <- function(model) {
18   # tidy the summary statistics
19   broom::glance(model) %>%
20   # add another summary statistic
21   mutate(pseudo_r2 = 1-(.data$deviance/.data$null.deviance))
22 }
23
```

HOW - DEMONSTRATION



```
devtools::document()  
devtools::check(cran = FALSE)
```

```
Console Terminal x Jobs x  
R 4.1.0 ~/PHD/broomoar.demo/ →  
v checking DESCRIPTION meta-information ...  
v checking top-level files  
v checking for left-over files  
v checking index information  
v checking package subdirectories ...  
v checking R files for non-ASCII characters ...  
v checking R files for syntax errors ...  
v checking whether the package can be loaded (711ms)  
v checking whether the package can be loaded with stated dependencies (608ms)  
v checking whether the package can be unloaded cleanly (712ms)  
v checking whether the namespace can be loaded with stated dependencies (611ms)  
v checking whether the namespace can be unloaded cleanly (818ms)  
v checking loading without being on the library search path (699ms)  
v checking dependencies in R code (711ms)  
v checking S3 generic/method consistency (1s)  
v checking replacement functions (610ms)  
v checking foreign function calls (609ms)  
v checking R code for possible problems (2s)  
v checking Rd files ...  
v checking Rd metadata ...  
v checking Rd cross-references ...  
v checking for missing documentation entries (612ms)  
v checking for code/documentation mismatches (1.9s)  
v checking Rd \usage sections (1.2s)  
v checking Rd contents ...  
v checking for unstated dependencies in examples ...  
v checking examples (1.1s)  
  
-- R CMD check results ----- broomoar.demo 0.0.0.9000 ----  
Duration: 22.6s  
0 errors v | 0 warnings v | 0 notes v  
> |
```


HOW - DEMONSTRATION

Create
another
function

THEN COPY AND
PASTE THE
FUNCTION TO
THE NEW BLANK
R SCRIPT, THEN
SAVE THE SCRIPT

```
fn_coefs <- function(model, vars_to_adjust = NULL, multiply_by = 1) {
  broom::tidy(model) %>%
  dplyr::rename(variable = .data$term) %>% # rename the variable column
  # adjust units, create OR and CIs, adjust decimals, relabel adjusted variable
  mutate(estimate = ifelse(.data$variable %in% vars_to_adjust,
    .data$estimate*multiply_by,
    .data$estimate),
    std.error = ifelse(.data$variable %in% vars_to_adjust,
    .data$std.error*multiply_by,
    .data$std.error),
    conf.low = ifelse(.data$variable %in% vars_to_adjust,
    confint(model)[,1]*multiply_by,
    confint(model)[,1]),
    conf.high = ifelse(.data$variable %in% vars_to_adjust,
    confint(model)[,2]*multiply_by,
    confint(model)[,2]),
    OR = exp(data$estimate),
    OR_CI_low = exp(.data$conf.low),
    OR_CI_high = exp(.data$conf.high),
    p.value = round(.data$p.value, 5), # round p-value
    variable = ifelse(.data$variable %in% vars_to_adjust,
    paste0(.data$variable, " (per ", multiply_by, ")"),
    .data$variable))
  }
  usethis::use_r("fn_coefs")
}
```

HOW - DEMONSTRATION

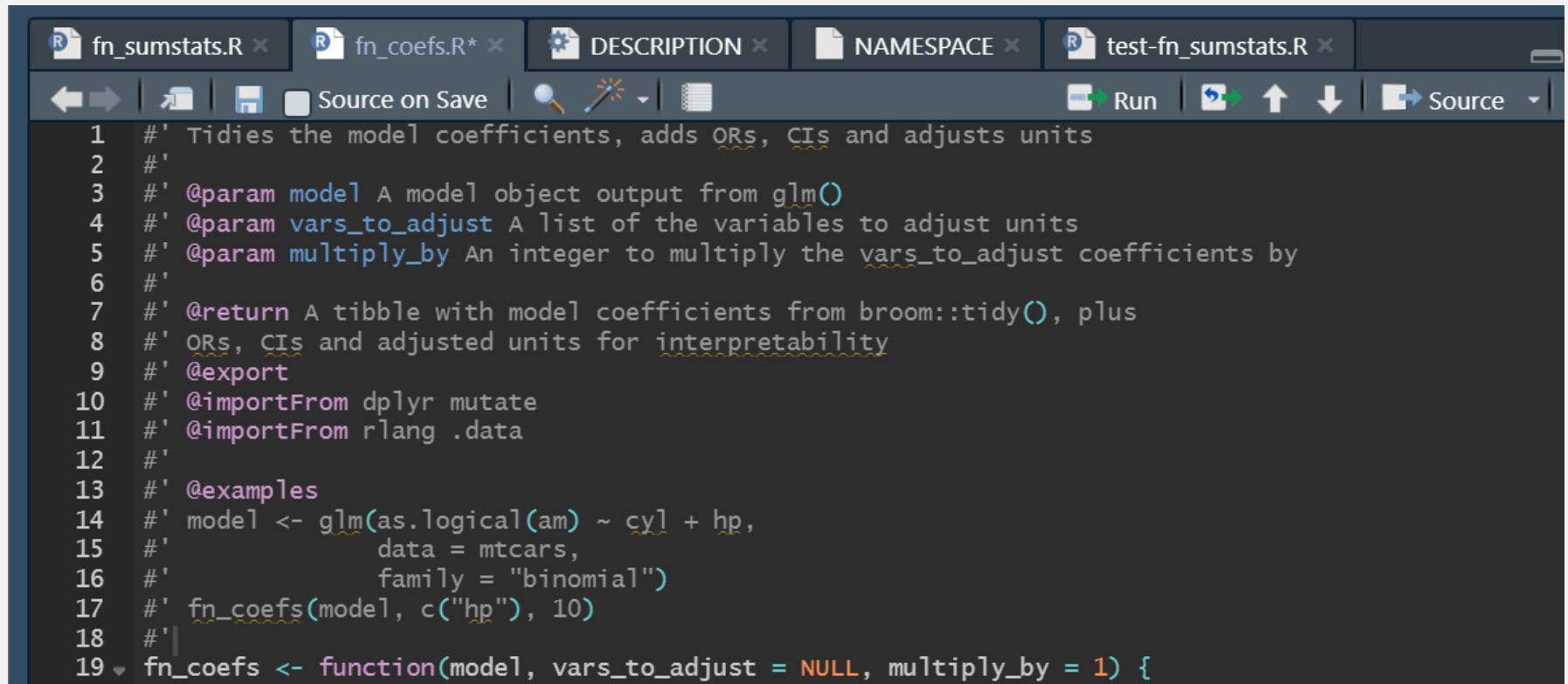
PUT CURSOR INSIDE THE FUNCTION IN THE SCRIPT

Code > Insert Roxygen Skeleton

FILL IN DOCUMENTATION DETAILS

Document

function



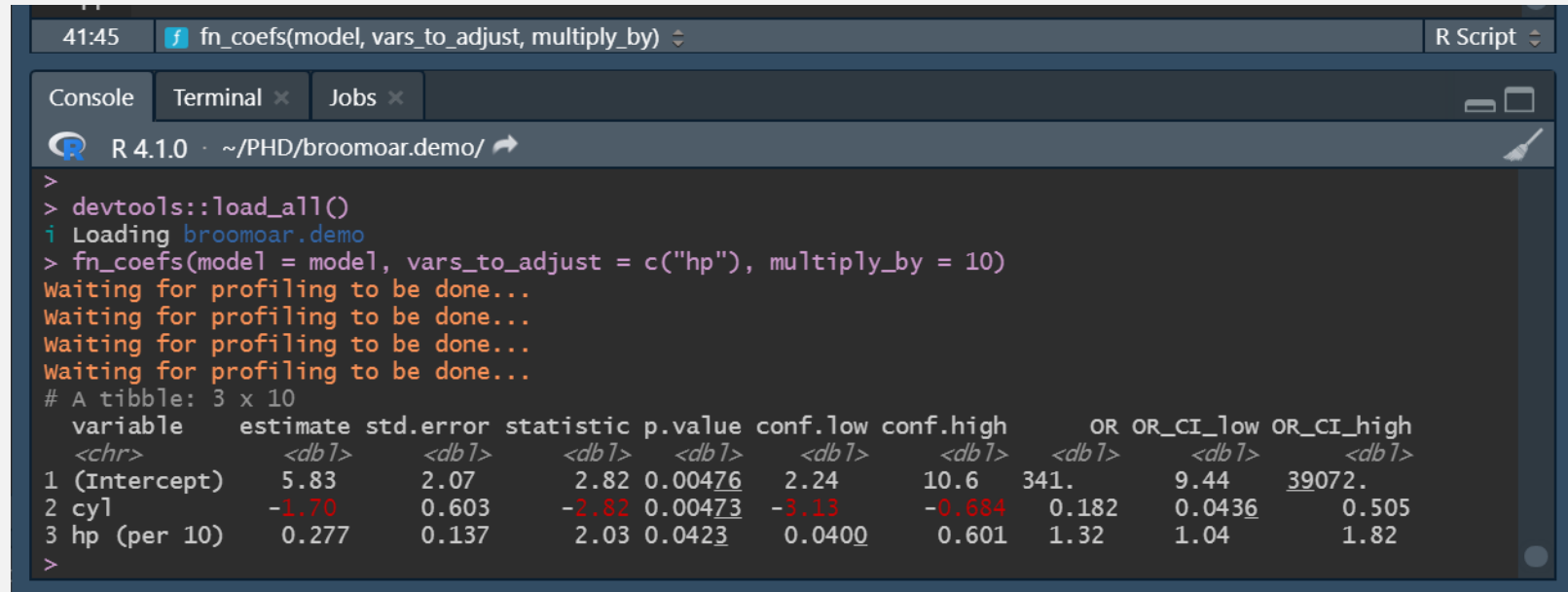
```
fn_sumstats.R x  fn_coefs.R* x  DESCRIPTION x  NAMESPACE x  test-fn_sumstats.R x
Source on Save  Run  Source
1 #' Tidies the model coefficients, adds ORs, CIs and adjusts units
2 #'
3 #' @param model A model object output from glm()
4 #' @param vars_to_adjust A list of the variables to adjust units
5 #' @param multiply_by An integer to multiply the vars_to_adjust coefficients by
6 #'
7 #' @return A tibble with model coefficients from broom::tidy(), plus
8 #' ORs, CIs and adjusted units for interpretability
9 #' @export
10 #' @importFrom dplyr mutate
11 #' @importFrom rlang .data
12 #'
13 #' @examples
14 #' model <- glm(as.logical(am) ~ cyl + hp,
15 #'             data = mtcars,
16 #'             family = "binomial")
17 #' fn_coefs(model, c("hp"), 10)
18 #'
19 fn_coefs <- function(model, vars_to_adjust = NULL, multiply_by = 1) {
```

HOW - DEMONSTRATION

Execute

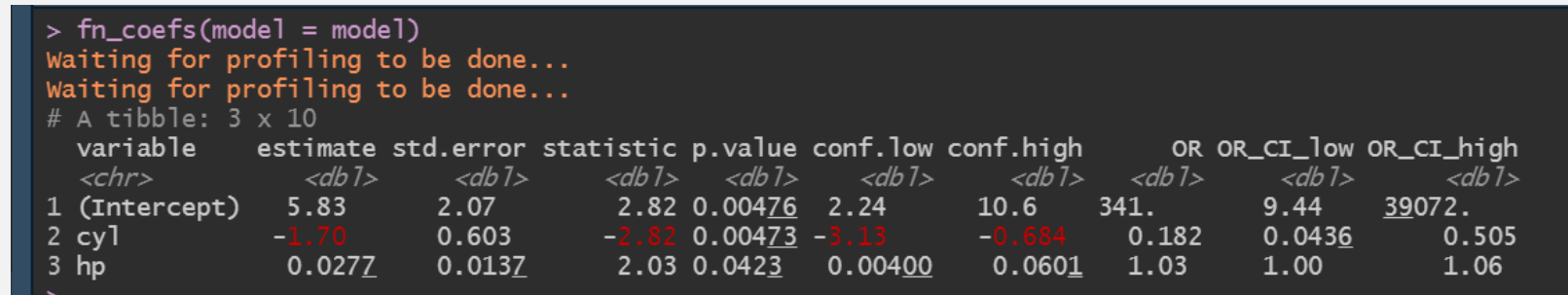
function

```
model <- glm(as.logical(am) ~ cyl + hp,  
            data = mtcars, family = "binomial")  
rm(fn_coefs) # prevents the 'conflicts' error  
devtools::load_all()  
fn_coefs(model = model, vars_to_adjust = c("hp"), multiply_by = 10)
```



```
41:45 f fn_coefs(model, vars_to_adjust, multiply_by) R Script  
Console Terminal Jobs  
R 4.1.0 ~/PHD/broomoar.demo/  
>  
> devtools::load_all()  
i Loading broomoar.demo  
> fn_coefs(model = model, vars_to_adjust = c("hp"), multiply_by = 10)  
Waiting for profiling to be done...  
Waiting for profiling to be done...  
Waiting for profiling to be done...  
Waiting for profiling to be done...  
# A tibble: 3 x 10  
  variable estimate std.error statistic p.value conf.low conf.high OR OR_CI_low OR_CI_high  
  <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
1 (Intercept) 5.83 2.07 2.82 0.00476 2.24 10.6 341. 9.44 39072.  
2 cyl -1.70 0.603 -2.82 0.00473 -3.13 -0.684 0.182 0.0436 0.505  
3 hp (per 10) 0.277 0.137 2.03 0.0423 0.0400 0.601 1.32 1.04 1.82  
>
```

```
fn_coefs(model = model)
```



```
> fn_coefs(model = model)  
Waiting for profiling to be done...  
Waiting for profiling to be done...  
# A tibble: 3 x 10  
  variable estimate std.error statistic p.value conf.low conf.high OR OR_CI_low OR_CI_high  
  <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
1 (Intercept) 5.83 2.07 2.82 0.00476 2.24 10.6 341. 9.44 39072.  
2 cyl -1.70 0.603 -2.82 0.00473 -3.13 -0.684 0.182 0.0436 0.505  
3 hp 0.0277 0.0137 2.03 0.0423 0.00400 0.0601 1.03 1.00 1.06  
>
```

HOW - DEMONSTRATION

Document function

```
devtools::document()  
?fn_coefs
```

```
3 #' @param model A model object output from glm()
4 #' @param vars_to_adjust A list of the variables to
5 #' @param multiply_by An integer to multiply the va
6 #'
7 #' @return A tibble with model coefficients from br
8 #' ORs, CIs and adjusted units for interpretability
9 #' @export
10 #' @importFrom dplyr mutate
11 #' @importFrom rlang .data
12 #'
13 #' @examples
14 #' model <- glm(as.logical(am) ~ cyl + hp,
15 #'             data = mtcars,
16 #'             family = "binomial")
17 #' fn_coefs(model, c("hp"), 10)
18 #'
19 fn_coefs <- function(model, vars_to_adjust = NULL,
20                      broom::tidy(model) %>%
21                      dplyr::rename(variable = .data$term) %>% # ren
22                      # adjust units, create OR and CIs, adjust decim
23                      mutate(estimate = ifelse(.data$variable %in% va
24                                                .data$estimate*multip
25                                                .data$estimate),
26                                                std.error = ifelse(.data$variable %in%
27                                                                  .data$std.error*multip
28                                                                  .data$std.error),
29                                                conf.low = ifelse(.data$variable %in% v
30                                                                  confint(model)[,1]*multi
31                                                                  confint(model)[,1]),
32                                                conf.high = ifelse(.data$variable %in%
33
```

fn_coefs.Rd - Find in Topic

fn_coefs {broomoar.demo} R Documentation

Tidies the model coefficients, adds ORs, CIs and adjusts units

Description

Tidies the model coefficients, adds ORs, CIs and adjusts units

Usage

```
fn_coefs(model, vars_to_adjust = NULL, multiply_by = 1)
```

Arguments

model	A model object output from glm()
vars_to_adjust	A list of the variables to adjust units
multiply_by	An integer to multiply the vars_to_adjust coefficients by

Value

A tibble with model coefficients from broom::tidy(), plus ORs, CIs and adjusted units for interpretability

Examples

```
model <- glm(as.logical(am) ~ cyl + hp,
             data = mtcars,
             family = "binomial")
fn_coefs(model, c("hp"), 10)
```

18:3 (Top Level) R Script

Console Terminal Jobs

```
R 4.1.0 ~/PHD/broomoar.demo/
>
> devtools::document()
i Updating broomoar.demo documentation
i Loading broomoar.demo
Writing fn_coefs.Rd
> ?fn_coefs
i Rendering development documentation for "fn_coefs"
>
> |
```

HOW - DEMONSTRATION

Run

check



```
devtools::check(cran = FALSE)
```

```
checking tests ...
v Running 'testthat.R' (908ms)

See
'C:/Users/SCDN42/AppData/Local/Temp/Rtmpu4DsQT/broomoar.demo.Rcheck/00check.log'
for details.

-- R CMD check results ----- broomoar.demo 0.0.0.9000 ----
duration: 26.7s

> checking R code for possible problems ... NOTE
fn_coefs: no visible global function definition for 'confint'
Undefined global functions or variables:
  confint
Consider adding
  importFrom("stats", "confint")
to your NAMESPACE file.

0 errors v | 0 warnings v | 1 note x
>
```

ADD @importFrom stats confint TO THE ROXYGEN HEADER

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

OPEN THE NAMESPACE FILE AND CHECK IT'S UPDATED

**** NEVER EDIT THE NAMESPACE FILE ITSELF ****

HOW - DEMONSTRATION

```
devtools::check(cran = FALSE)
```



```
Console Terminal x Jobs x
R 4.1.0 · ~/PHD/broomoar.demo/
v checking DESCRIPTION meta-information ...
v checking top-level files
v checking for left-over files ...
v checking index information
v checking package subdirectories ...
v checking R files for non-ASCII characters ...
v checking R files for syntax errors ...
v checking whether the package can be loaded (711ms)
v checking whether the package can be loaded with stated dependencies (615ms)
v checking whether the package can be unloaded cleanly (712ms)
v checking whether the namespace can be loaded with stated dependencies (611ms)
v checking whether the namespace can be unloaded cleanly (811ms)
v checking loading without being on the library search path (715ms)
v checking dependencies in R code (712ms)
v checking S3 generic/method consistency (1s)
v checking replacement functions (613ms)
v checking foreign function calls (611ms)
v checking R code for possible problems (2s)
v checking Rd files ...
v checking Rd metadata ...
v checking Rd cross-references ...
v checking for missing documentation entries (611ms)
v checking for code/documentation mismatches (1.8s)
v checking Rd \usage sections (1.2s)
v checking Rd contents ...
v checking for unstated dependencies in examples ...
v checking examples (1.2s)
v checking for unstated dependencies in 'tests' ...
- checking tests ...
v Running 'testthat.R' (849ms)

-- R CMD check results ----- broomoar.demo 0.0.0.9000 ----
Duration: 23.5s
0 errors v | 0 warnings v | 0 notes v
>
```

HOW - DEMONSTRATION

Design

and run

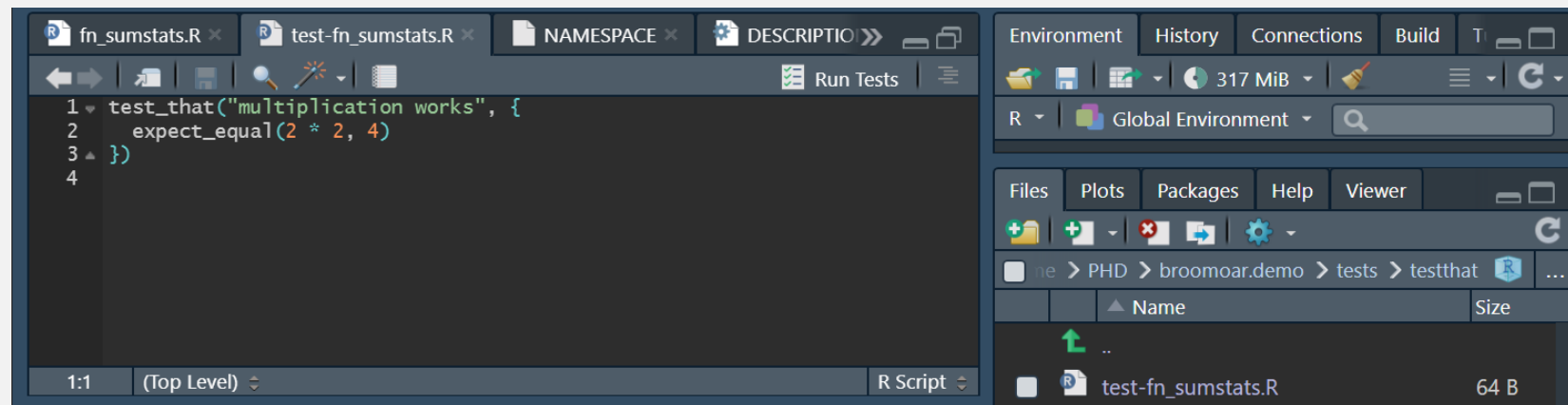
tests

```
usethis::use_testthat()
```

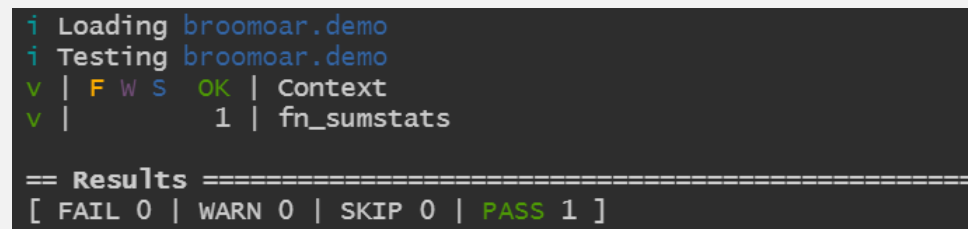
→ updates the DESCRIPTION file to include testthat as a suggested package and creates a test directory that contains:

- a script named testthat.R that contains all the code required to run the tests
- a directory testthat where you will save all of your test scripts.

```
usethis::use_test()
```

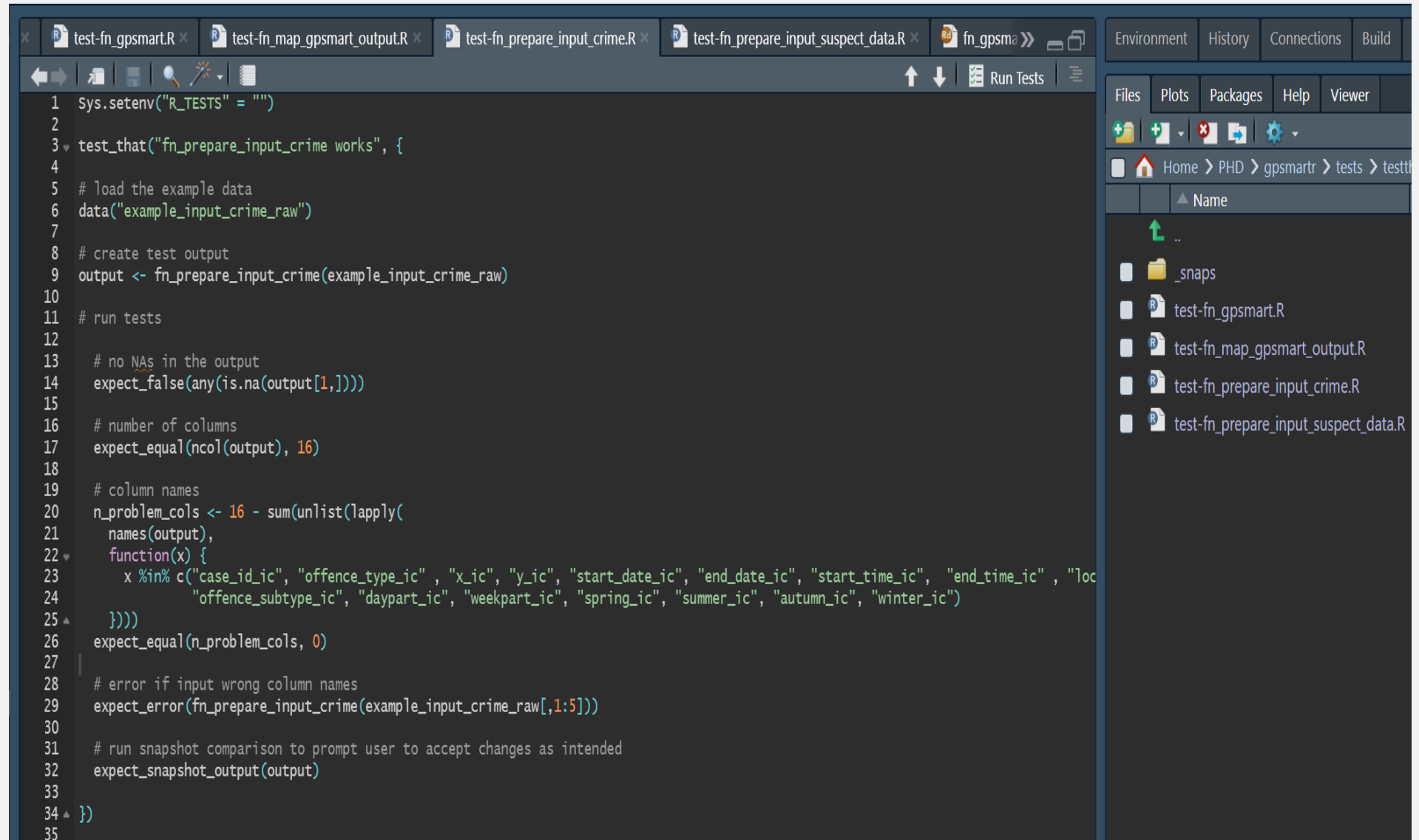


```
devtools::test()
```



HOW - DEMONSTRATION

Design tests



```
1 Sys.setenv("R_TESTS" = "")
2
3 test_that("fn_prepare_input_crime works", {
4
5 # load the example data
6 data("example_input_crime_raw")
7
8 # create test output
9 output <- fn_prepare_input_crime(example_input_crime_raw)
10
11 # run tests
12
13 # no NAs in the output
14 expect_false(any(is.na(output[1,])))
15
16 # number of columns
17 expect_equal(ncol(output), 16)
18
19 # column names
20 n_problem_cols <- 16 - sum(unlist(lapply(
21   names(output),
22   function(x) {
23     x %in% c("case_id_ic", "offence_type_ic", "x_ic", "y_ic", "start_date_ic", "end_date_ic", "start_time_ic", "end_time_ic", "loc
24             "offence_subtype_ic", "daypart_ic", "weekpart_ic", "spring_ic", "summer_ic", "autumn_ic", "winter_ic")
25   })))
26 expect_equal(n_problem_cols, 0)
27
28 # error if input wrong column names
29 expect_error(fn_prepare_input_crime(example_input_crime_raw[,1:5]))
30
31 # run snapshot comparison to prompt user to accept changes as intended
32 expect_snapshot_output(output)
33
34 })
35
```

The screenshot shows the RStudio interface. The top pane displays the R script with line numbers 1 through 35. The script includes a test function that checks for NAs, column count, column names, and snapshot output. The bottom pane shows a file explorer with the following files:

- ..
- .._snaps
- test-fn_gpsmart.R
- test-fn_map_gpsmart_output.R
- test-fn_prepare_input_crime.R
- test-fn_prepare_input_suspect_data.R

HOW - DEMONSTRATION

Design
tests

```
psmart.R x DESCRIPTION x test-fn_gpsmart.R x test-fn_map_gpsmart_output.R x fn_gpsmart.Rd x fn_map_gpsmart_output.R x RE >>
1 Sys.setenv("R_TESTS" = "")
2
3 # load the example data
4 data("example_input_crime_raw")
5 data("example_input_suspects_raw")
6
7 # select a random sample to test ensuring representation of:
8 # - example input crime offender
9 # - records of each node_type for other suspects
10 # - but this time we don't need records of each level of each reliability and relevance factor, to check that the function copes with
11 set.seed(99)
12 input_suspects <- example_input_suspects_raw %>%
13   dplyr::group_by(node_type) %>%
14   dplyr::slice_sample(n=100) %>%
15   dplyr::ungroup() %>%
16   dplyr::bind_rows(
17     example_input_suspects_raw %>%
18     dplyr::filter(stringr::str_detect(person_id, stringr::str_sub(example_input_crime_raw$case_id[[1]], start = 1L, end = 7L)))
19   ) %>% fn_prepare_suspect_data()
20
21 # create test outputs
22 output_default_weights_node_preds <- fn_gpsmart(input_crime = fn_prepare_input_crime(example_input_crime_raw),
23   input_suspects = input_suspects,
24   search_radius = 10,
25   return_node_predictions = TRUE)
26
27 output_defaults <- fn_gpsmart(input_crime = fn_prepare_input_crime(example_input_crime_raw),
28   input_suspects = input_suspects)
29
30 # run tests
31 test_that("output no error", {
32
33   expect_error(fn_map_gpsmart_output(output_default_weights_node_preds), NA)
34 })
35
36 test_that("output no warning", {
37
38   expect_warning(fn_map_gpsmart_output(output_default_weights_node_preds), NA)
39 })
40
41
42 test_that("error if no node predictions in input", {
43
44   expect_error(fn_map_gpsmart_output(output_defaults),
45     "No suspect node predictions to map",
46     fixed = TRUE)
47 })
48
49
50 # run snapshot comparison to prompt user to accept changes as intended
51 test_that("output matches expected output for test input using defaults", {
52
53   expect_snapshot_output(fn_map_gpsmart_output(output_default_weights_node_preds))
54 })
55
56
```

HOW - DEMONSTRATION

Use

package

```
devtools::install("path/to/package/broomoar.demo")
library(broomoar.demo)
model <- glm(as.logical(am) ~ cyl + hp,
             data = mtcars, family = "binomial")
fn_sumstats(model)
```

The screenshot displays the RStudio interface. The left pane shows the R console with the following output:

```
R 4.1.0 ~/PHD/nscr_workshop/
> devtools::install("C:/Users/SCDN42/Documents/PHD/broomoar.demo")
v checking for file 'C:/Users/SCDN42/Documents/PHD/broomoar.demo/DESCRIPTION' ...
- preparing 'broomoar.demo':
v checking DESCRIPTION meta-information ...
- checking for LF line-endings in source and make files and shell scripts
- checking for empty or unneeded directories
Omitted 'LazyData' from DESCRIPTION
- building 'broomoar.demo_0.0.0.9000.tar.gz'

Running "C:/PROGRA~1/R/R-41~1.0/bin/x64/Rcmd.exe" INSTALL \
"C:/Users/SCDN42/AppData/Local/Temp/RtmpApBrRk/broomoar.demo_0.0.0.9000.tar.gz" \
--install-tests
* installing to library 'C:/Program Files/R/R-3.6.3/library'
* installing *source* package 'broomoar.demo' ...
** using staged installation
** R
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
converting help for package 'broomoar.demo'
  fn_sumstats      html
  pipe             html
REDIRECT:topic %%% -> pipe.html [ FAIL ]
** building package indices
** testing if installed package can be loaded from temporary location
*** arch - i386
*** arch - x64
** testing if installed package can be loaded from final location
*** arch - i386
*** arch - x64
** testing if installed package keeps a record of temporary installation path
* DONE (broomoar.demo)
> library(broomoar.demo)
> model <- glm(as.logical(am) ~ cyl + hp, data = mtcars, family = "binomial")
>
> fn_sumstats(model)
# A tibble: 1 x 9
  null.deviance df.null logLik AIC BIC deviance df.residual nobs pseudo_r2
  <dbl> <int> <dbl> <dbl> <dbl> <dbl> <int> <int> <dbl>
1 43.2 31 -14.3 34.6 39.0 28.6 29 32 0.338
```

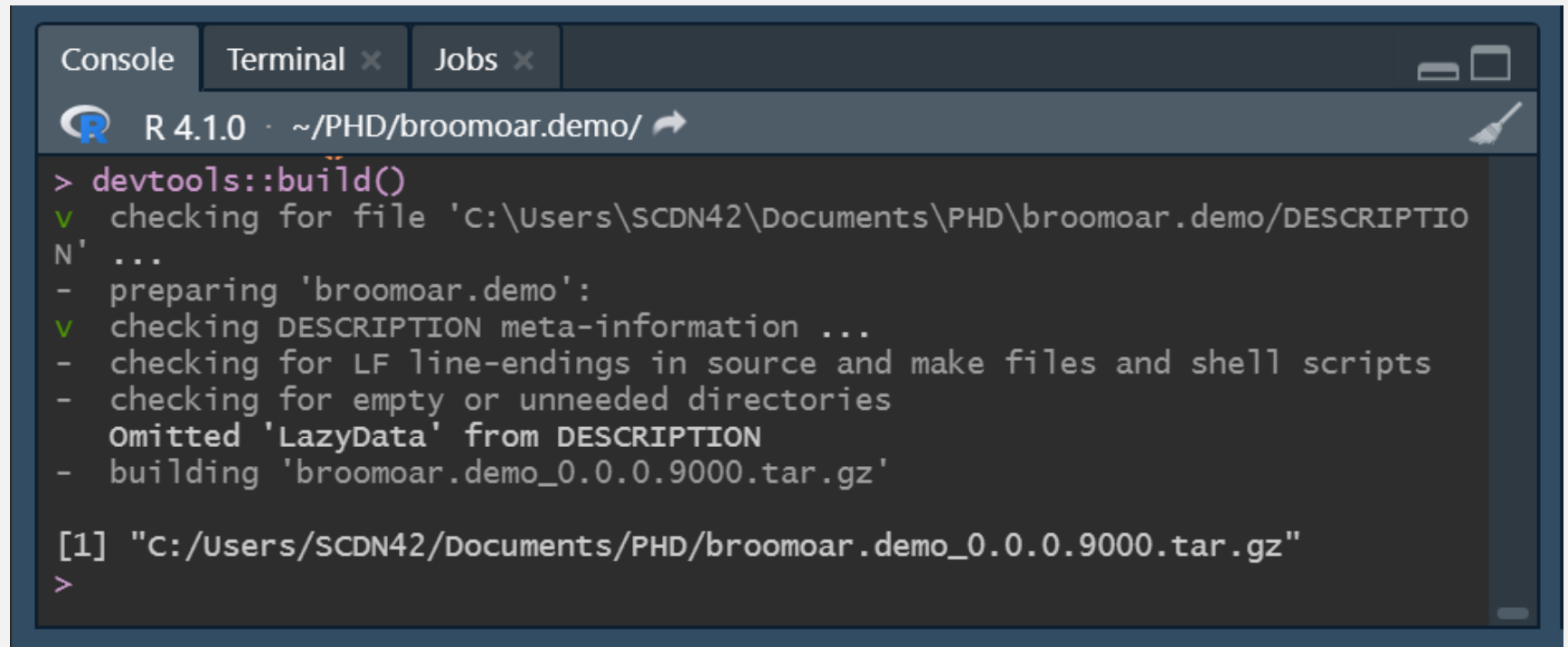
The right pane shows the 'Packages' tab with a list of installed and available packages. The 'broomoar.demo' package is highlighted with a red circle. The table below represents the data shown in the 'Packages' pane:

Name	Description	Version
brew	Templating Framework for Report Generation	1.0-7
brio	Basic R Input Output	1.1.3
broom	Convert Statistical Objects into Tidy Tibbles	0.8.0
broom	Convert Statistical Objects into Tidy Tibbles	0.8.0
<input checked="" type="checkbox"/> broomoar.demo	Adds custom statistics to broom tidied model outputs	0.0.0.9000
bslib	Custom Bootstrap Sass Themes for 'shiny' and 'rmarkdown'	0.3.1
cachem	Cache R Objects with Automatic Pruning	1.0.6
callr	Call R from R	3.7.0
callr	Call R from R	3.7.0
car	Companion to Applied Regression	3.0-13
carData	Companion to Applied Regression Data Sets	3.0-5

HOW - DEMONSTRATION

Share
package

```
devtools::build()
```



The screenshot shows an RStudio terminal window with the following content:

```
Console Terminal x Jobs x  
R 4.1.0 · ~/PHD/broomoar.demo/ →  
> devtools::build()  
✓ checking for file 'C:\Users\SCDN42\Documents\PHD\broomoar.demo\DESCRIPTION'  
...  
- preparing 'broomoar.demo':  
✓ checking DESCRIPTION meta-information ...  
- checking for LF line-endings in source and make files and shell scripts  
- checking for empty or unneeded directories  
Omitted 'LazyData' from DESCRIPTION  
- building 'broomoar.demo_0.0.0.9000.tar.gz'  
  
[1] "C:/Users/SCDN42/Documents/PHD/broomoar.demo_0.0.0.9000.tar.gz"  
>
```

```
install.packages("path/to/file/broomoar.demo_0.0.0.9000.tar.gz")
```

OR PUSH PACKAGE TO GITHUB WITHIN RSTUDIO, THEN INSTALL FROM GITHUB:

```
devtools::install_github("author/packageName")
```