

Workshop 3:

The tidyverse and beyond

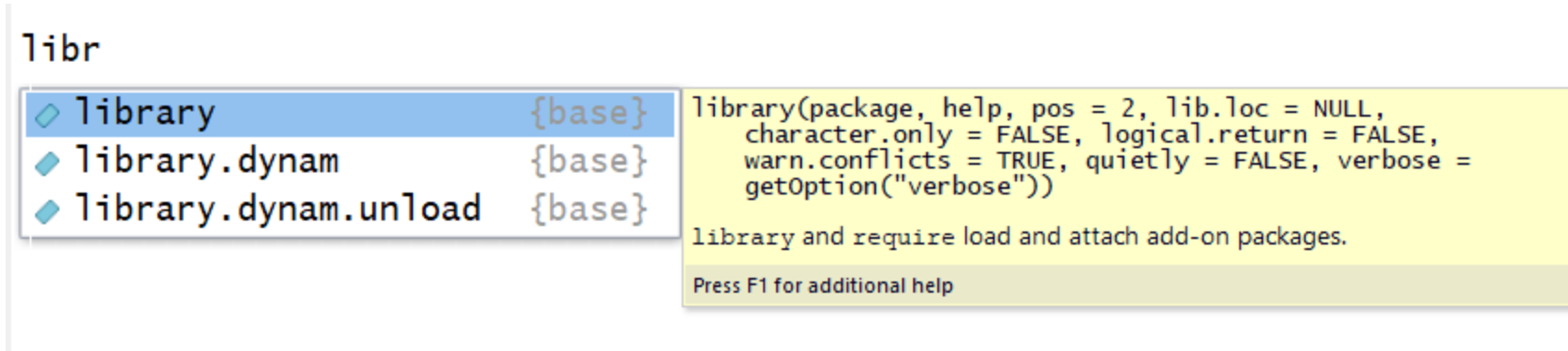
- Send an SOS to the world



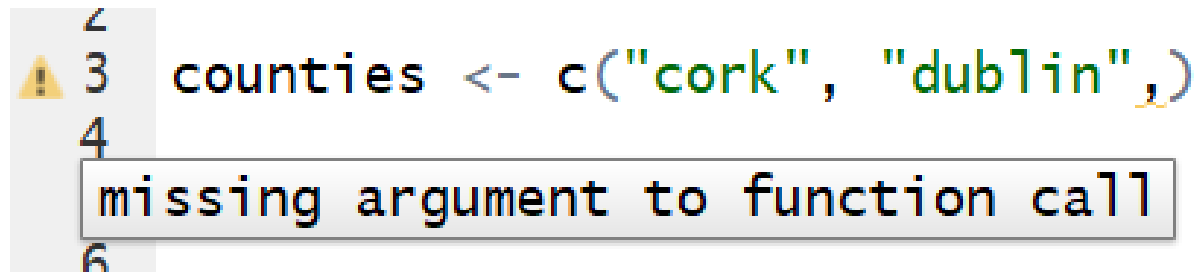
Brendan Palmer,
Statistics & Data Analysis Unit,
Clinical Research Facility - Cork

Levels of help within RStudio

1. Let R help you write your code using Tab



2. Hover over the error symbols to identify what the error is



```
> counties <- c("cork", "dublin",)
Error in c("cork", "dublin", ) : argument 3 is empty
> |
```

Levels of help within RStudio

3. Watch out for capitalisation and naming errors as the code might run, but all the arguments supplied may return "FALSE"

```
> counties <- c("cork", "dublin")
> cork_dublin_df <- house_reg_df %>%
+   filter(county %in% counties)
> |

> ncol(cork_dublin_df)
[1] 3
> nrow(cork_dublin_df)
[1] 0
> unique(hse_df$county)
 [1] "Carlow"      "Cavan"      "Clare"      "Cork"      "Donegal"    "Dublin"     "Galway"
 [8] "Kerry"      "Kildare"    "Kilkenny"   "Laois"     "Leitrim"    "Limerick"   "Longford"
[15] "Louth"      "Mayo"      "Meath"      "Monaghan"  "Offaly"     "Roscommon"  "Sligo"
[22] "Tipperary"  "Waterford" "Westmeath"  "Wexford"   "Wicklow"
> |

> unique(hse_df$county) %in% counties
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[15] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
> |
```

Levels of help within RStudio

4. `> ?gather`

- If you're unsure about a function, place the question mark before it to retrieve some help documentation

`gather {tidyr}`

R Documentation

Gather columns into key-value pairs.

Description

Gather takes multiple columns and collapses into key-value pairs, duplicating all other columns as needed. You use `gather()` when you notice that you have columns that are not variables.

- describes the various arguments to the function
- provides useful examples to guide you through common operations

```
# get first observation for each Species in iris data -- base R
mini_iris <- iris[c(1, 51, 101), ]
# gather Sepal.Length, Sepal.Width, Petal.Length, Petal.Width
gather(mini_iris, key = flower_att, value = measurement,
       Sepal.Length, Sepal.Width, Petal.Length, Petal.Width)
```

Worksheet A

Open script1_ws3_correct_the_errors.R

Package Cheatsheets

- Cheatsheets are available for commonly used packages
 - c.f. last week we explored the baseR cheatsheet
- Useful for quick reference to the most commonly used functions for that package
 - <https://www.rstudio.com/resources/cheatsheets/>
- Open the ggplot2 cheatsheet

Package Vignettes

- A vignette is a long-form guide to your package
- Before R 3.0.0, vignettes were standard pdfs
 - The development of RMarkdown has made vignette building and navigation more accessible
- A vignette should divide functions into useful categories, and demonstrate how to coordinate multiple functions to solve problems (but this may not always be the case)
- You can see all the installed vignettes with;
`browseVignettes()`
- Try it!
 - Click on some hyperlinks to explore the content

Package Webpages

- Many packages are one offs;
 - developed by individuals/labs to solve specific problems
 - once funding expires, package development ends
 - may get released but never updated
- Many of the commonly used packages are updated regularly;
 - new versions released periodically
 - dedicated webpages
 - e.g. lets now explore <http://ggplot2.tidyverse.org>

Worksheet B

Open `script2_ws3_exploring_ggplot2_resources.R`

Stackoverflow

- <https://stackoverflow.com/>

Worksheet C

Open script3_ws3_add_plot_layers.R