Welcome to the **Tidy Tools** Workshop!

I have free parking passes if you drove

1. Get connected to the wifi

   Westin meetings / tidytools

2. Get the course materials

   usethis::use_course("http://rstd.io/tidy-tools")

3. Follow the instructions in **setup.R**

4. Stuck? Please ask for help!

Preliminaries


HELLO
my name is
Hadley

Charlotte Wickham

Oregon State University +

Independent Consultant

Dave Hurst

Director of Business Development,
RStudio

Karthik Ram

Senior scientist,

Berkeley Institute for Data Science

Your turn

This course is very hands on, and while we're here to help you, the best resource is often the person sitting next to you.

Introduce yourself to your neighbours. Who are you and what are you using R for?

This means that you have to work!

Goal: help build tidy tools

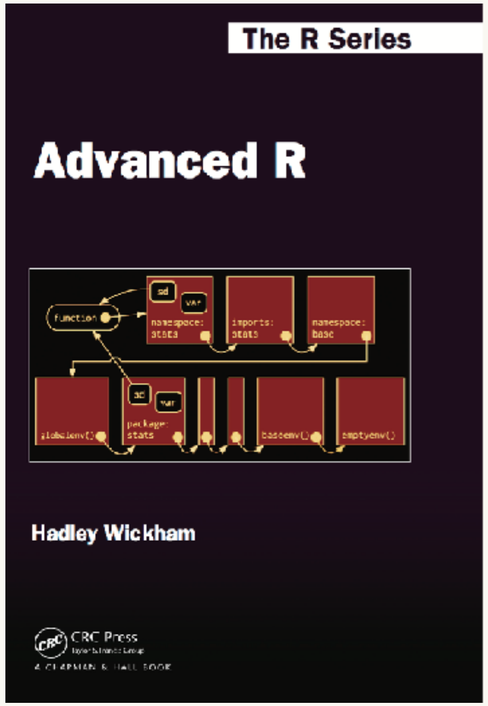
Writing functions

Individual functions to solve individual problems

Designing APIs

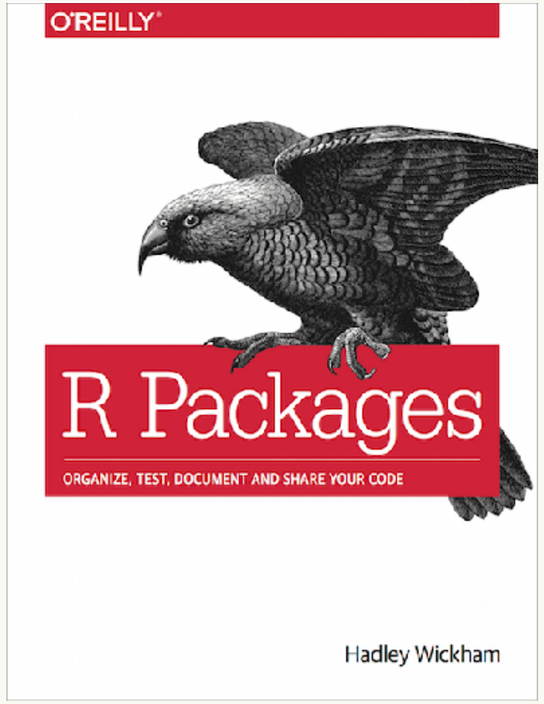
Family of functions that work together to solve family of problems

Much of the course is drawn from existing books



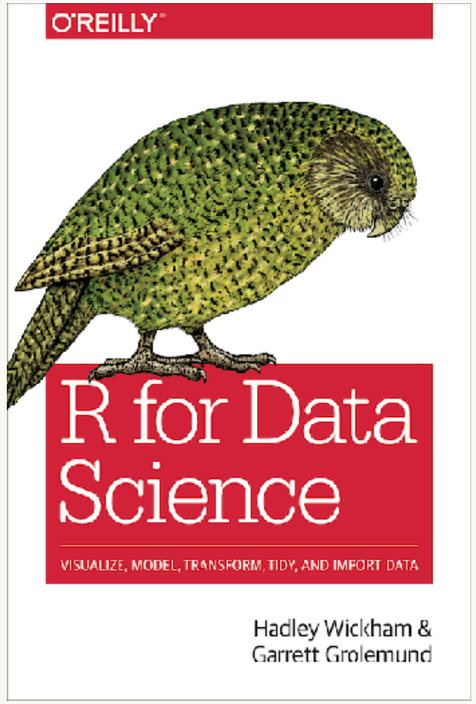
http://adv-r.hadley.nz/

Working on 2nd ed



http://r-pkgs.had.co.nz

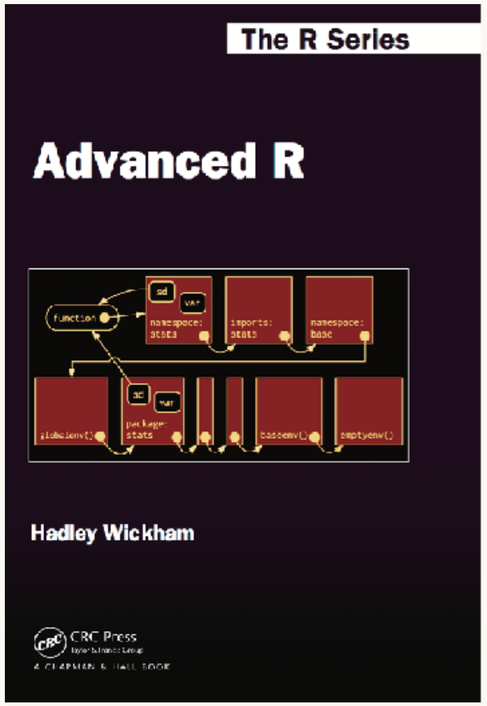
https://amzn.com/1491910399



http://r4ds.had.co.nz

https://amzn.com/1491910399

But the primary book does not yet exist



How R works

Tidy tools

Hadley Wickham

How to solve real problems with R

Warmups

Don’t expect to know all the answers!

Your turn

What are the four common types of atomic vectors? (Bonus points for the two uncommon types)

What are the three primary properties of a vector?

Four common types: logical, integer, double, character

```
typeof(TRUE)
typeof(1L)
typeof(1.5)
typeof("a")

# We'll talk about this (S3) later:
typeof(factor(1:10))
typeof(Sys.Date())
```

Every vector has three properties:

```
x <- 1:5

# 1. Type:
typeof(x)

# 2. Length
length(x)

# 3. Attributes
attributes(x)
# (we'll come back to those later)
```

Missing values

What does NA == NA return? Why?

What should you use instead?

There isn’t a single unknown value

```
age_john <- NA
age_mary <- NA
age_john == age_mary


is.na(x)
```

```
sum(is.na(x))
mean(is.na(x))
```

Your turn

What are the six types of thing that you can put inside []?

blank	include all
integer	**+ve**: include **0**: drop all **-ve**: exclude
logical	keep TRUEs
character	lookup by name

Use character subsetting for simple look ups

```
x <- c("m", "f", "u", "f", "f", "m", "m")

lookup <- c(m = "Male", f = "Female", u = NA)
lookup[x]

unname(lookup[x])
```

Your turn

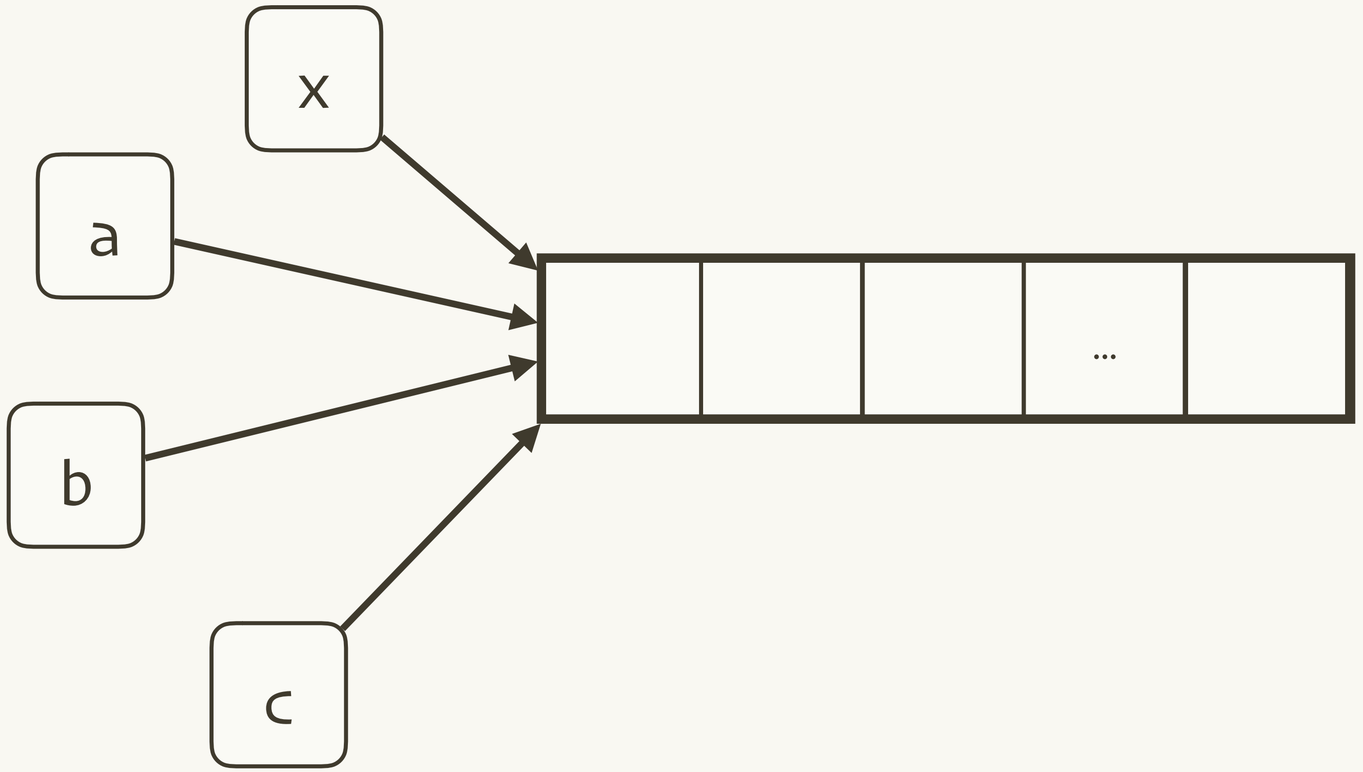
```
x <- runif(1e6)
lobstr::obj_size(x)
#> 8,000,040 B

y <- list(x, x, x)
lobstr::obj_size(y)
#> ???

y[[1]][[1]] <- NA
lobstr::obj_size(y)
#> ???
```

A **name** is a reference to a **value**

```
x <- runif(1e6)
```

Many references can point to one object

```
a <- b <- c <- x
```

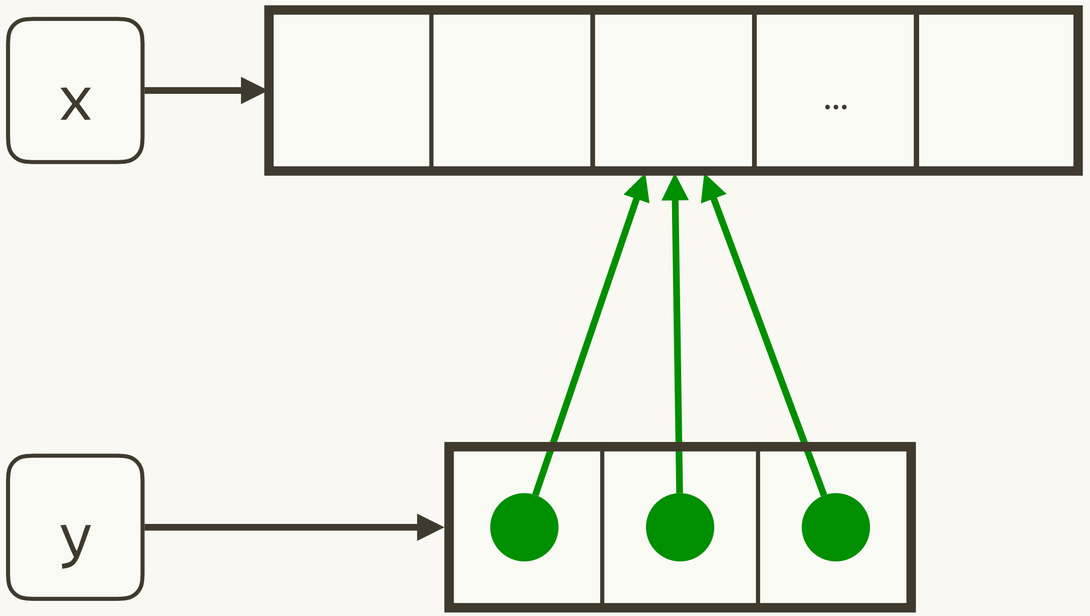


How big is **a**? How big is **b**? How big are **a** and **b** together?

Elements of lists are also references

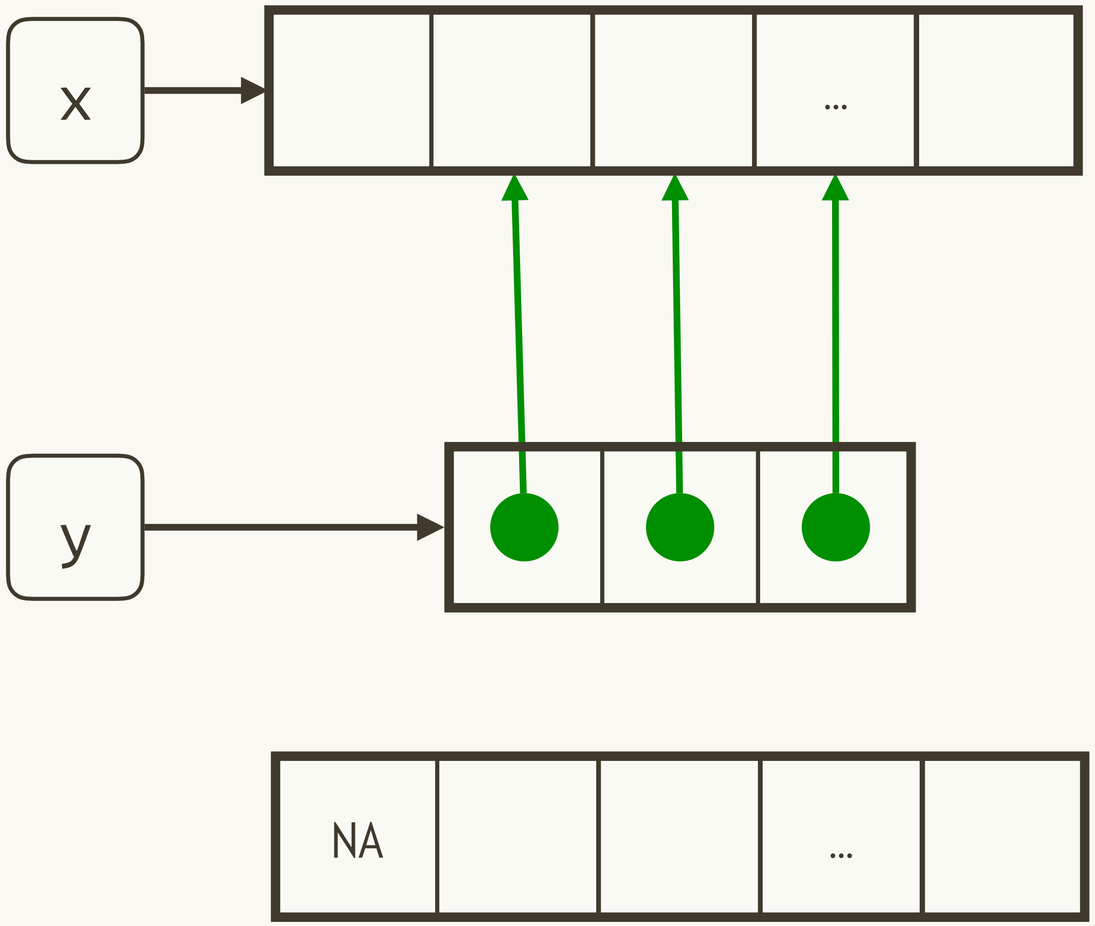
```
y <- list(x, x, x)
```

How big is **x**? How big is **y**? How big are **x** and **y** together?



Modifying an object creates a copy

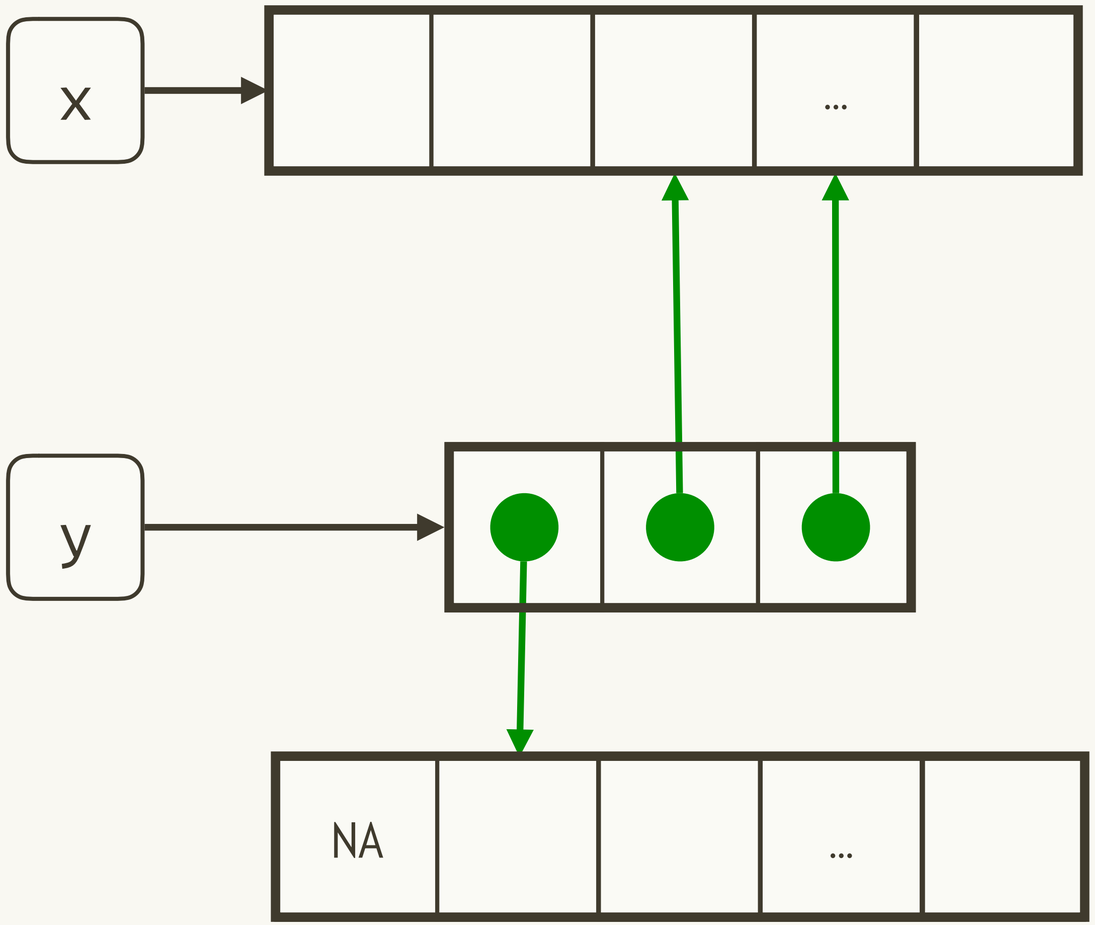
```
y[[1]][[1]] <- NA
```



Modifying an object creates a copy

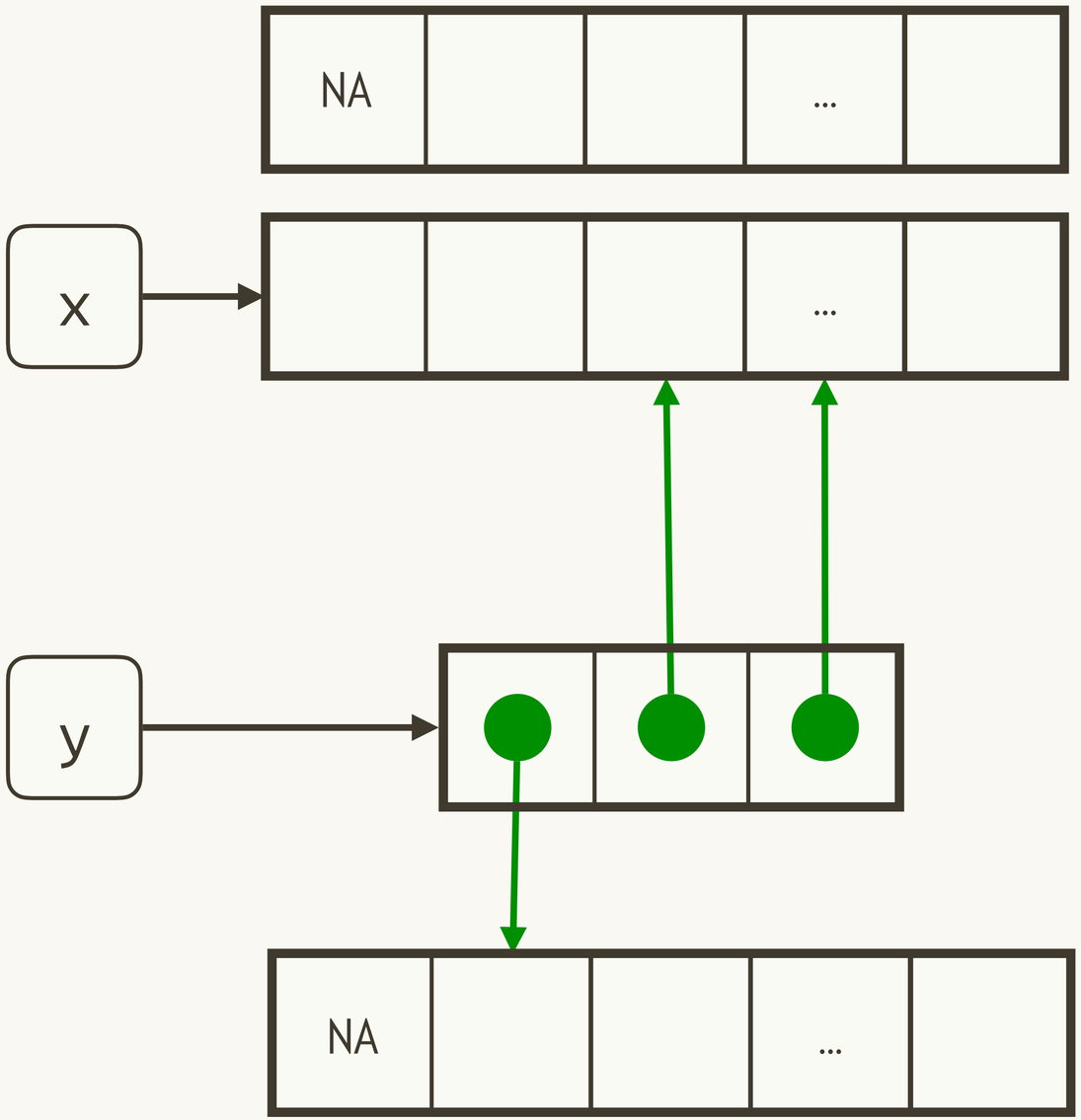
```
y[[1]][[1]] <- NA
```

How big is **x**? How big is **y**? How big are **x** and **y** together?



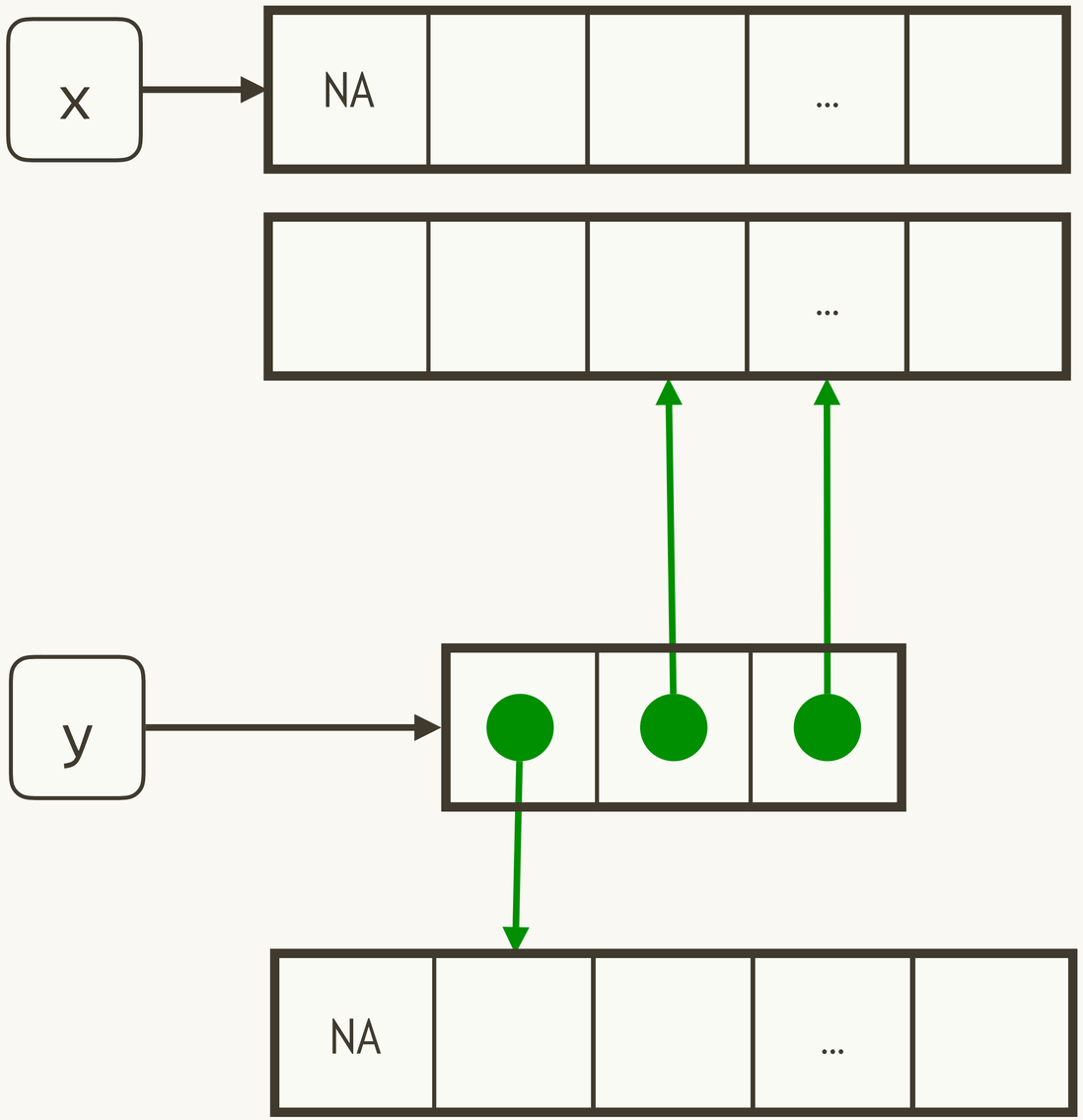
Modifying an object creates a copy

```
x[[1]] <- NA
```



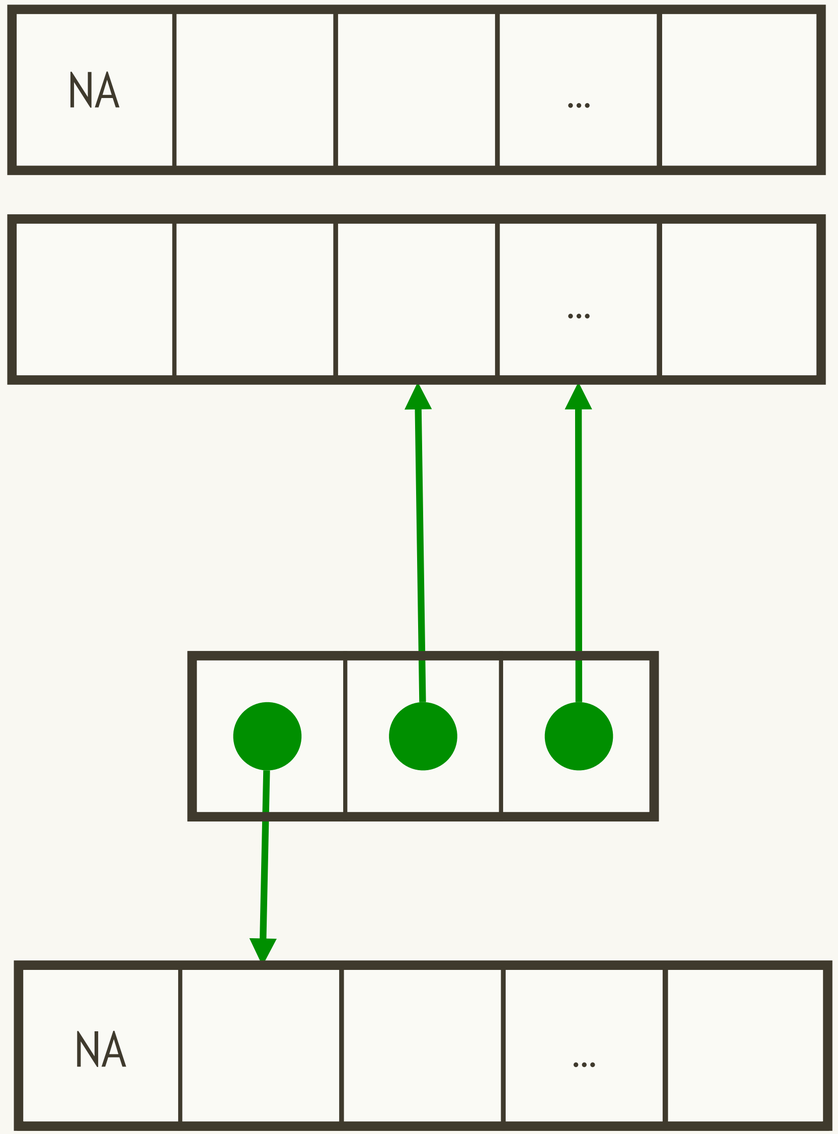
Modifying an object creates a copy

```
x[[1]] <- NA
```

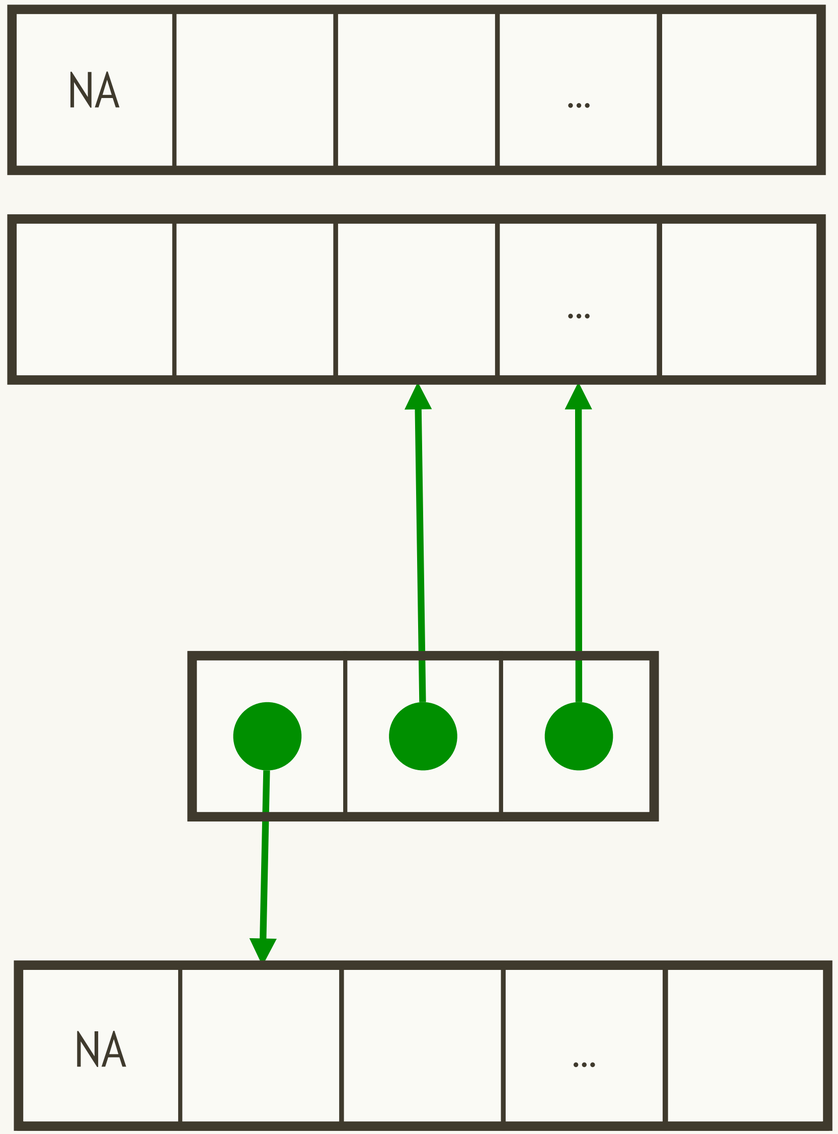


rm() removes references

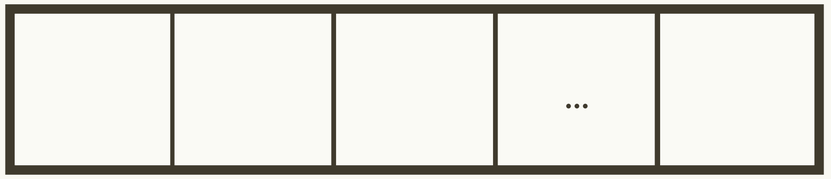
```
rm(x, y)
```

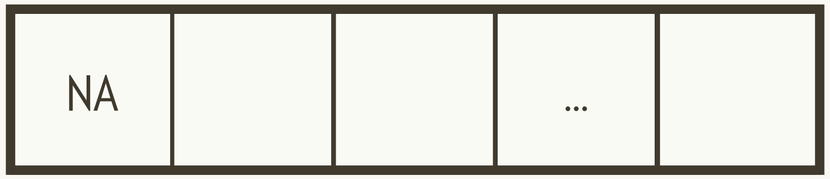


The garbage collector removes values



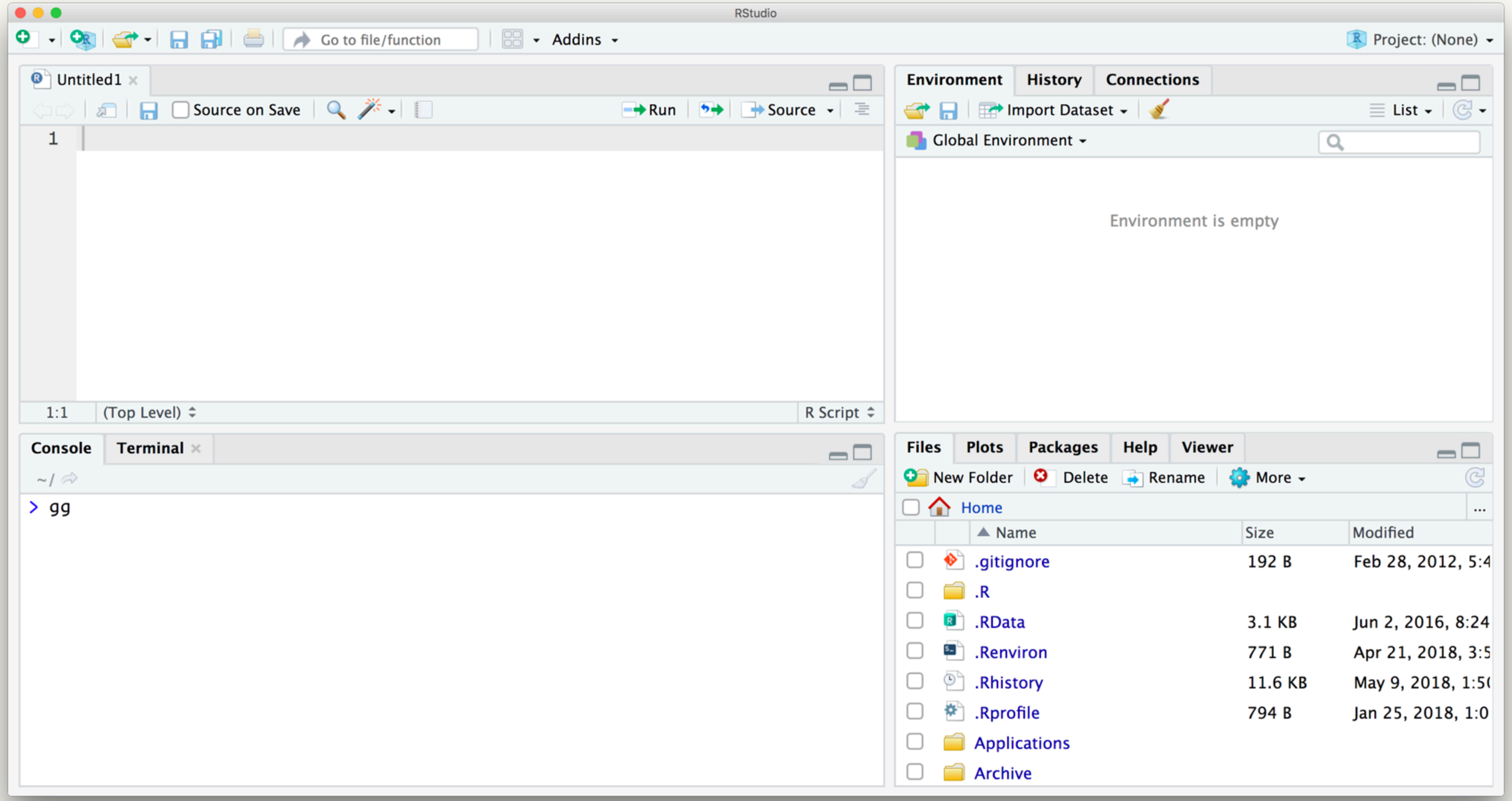
The garbage collector removes values

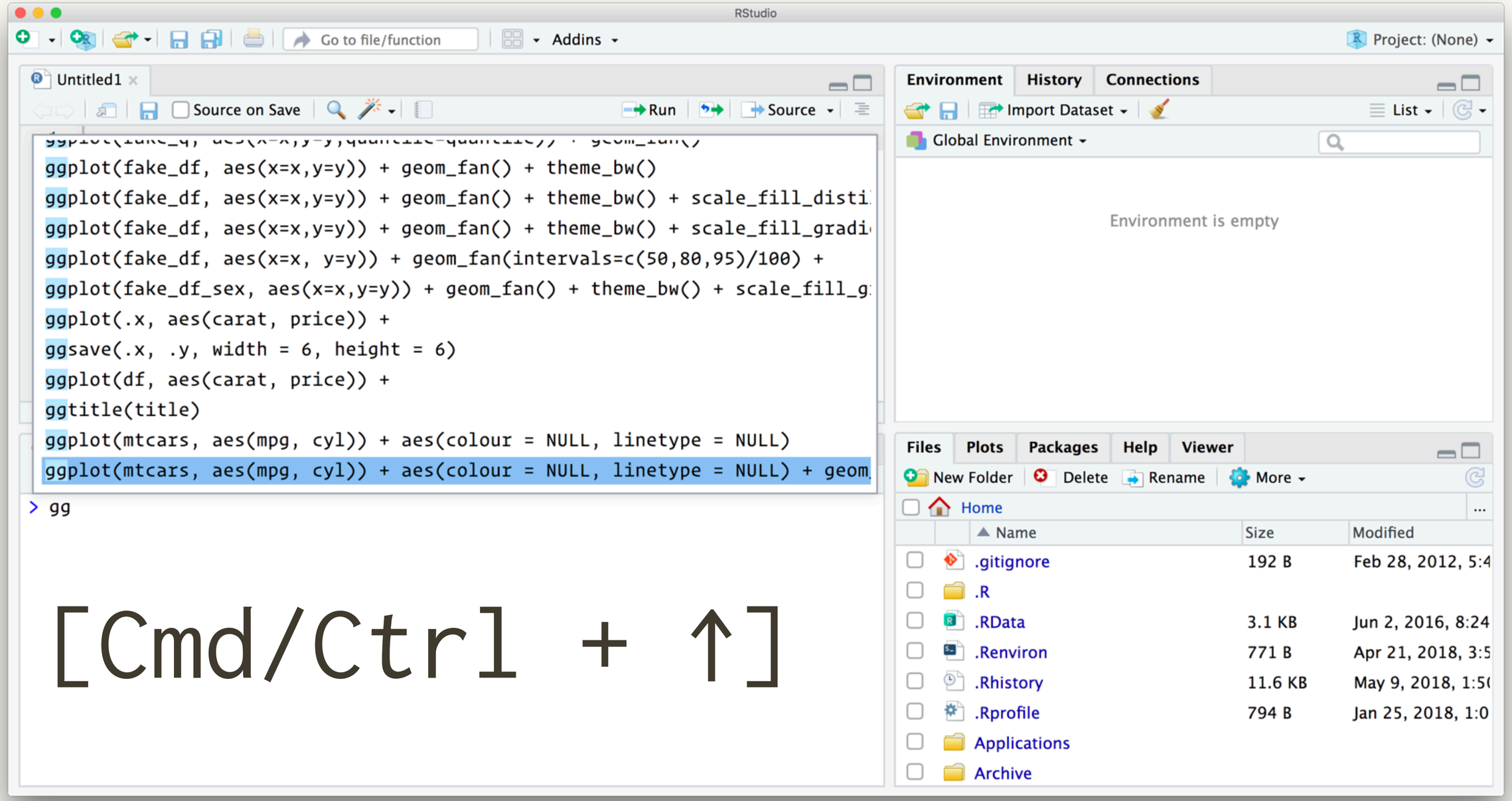


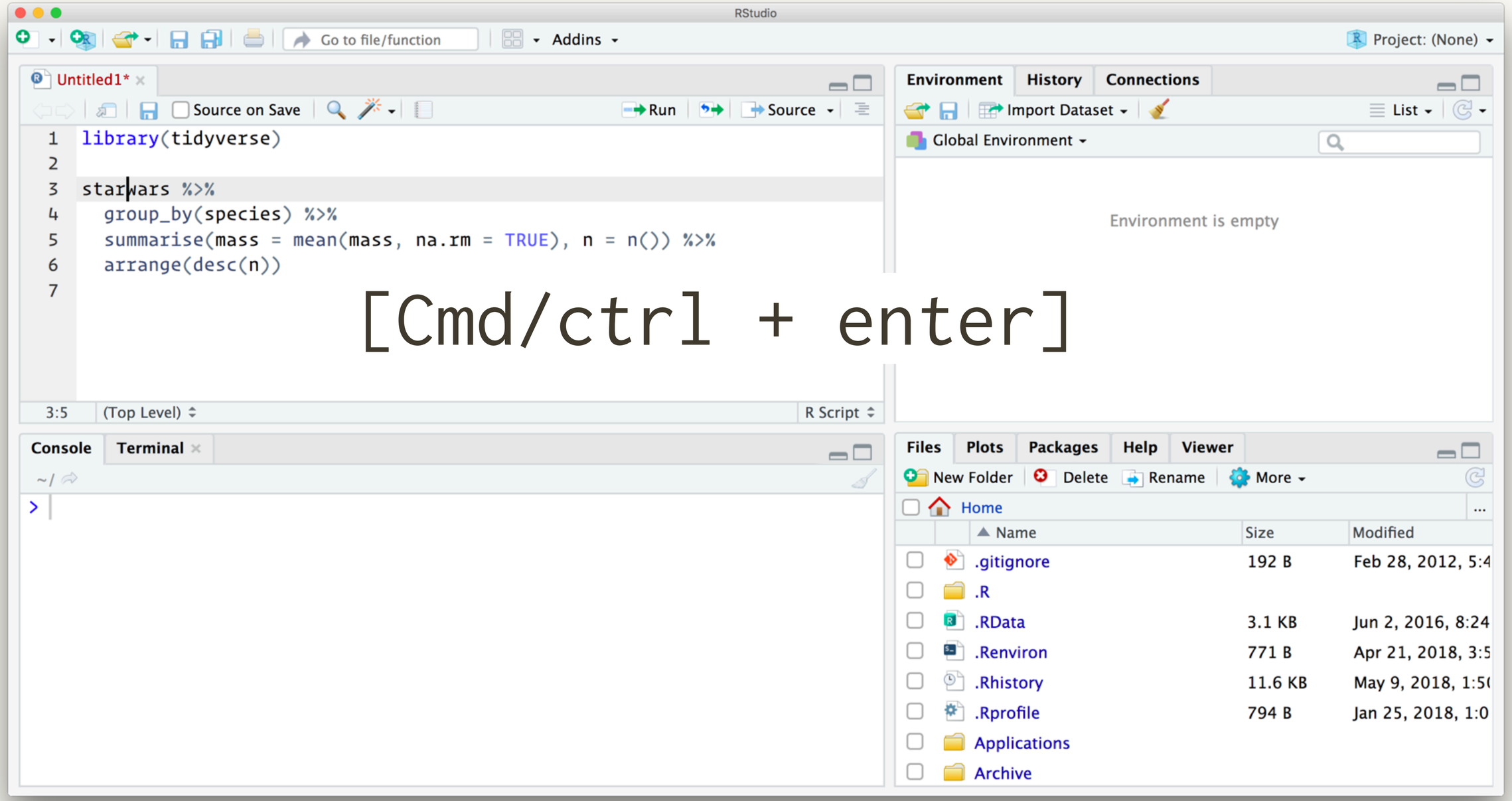


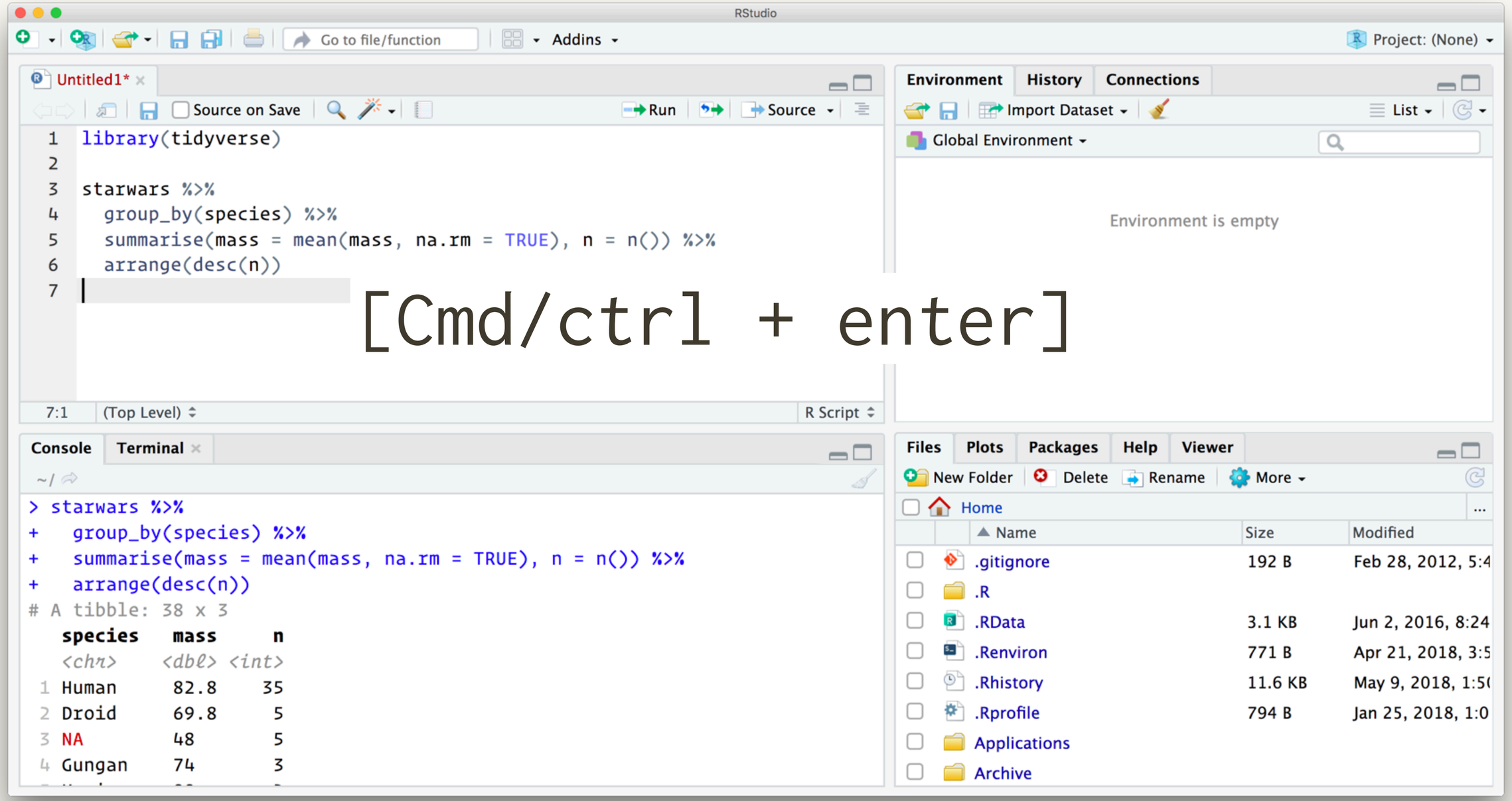
RStudio

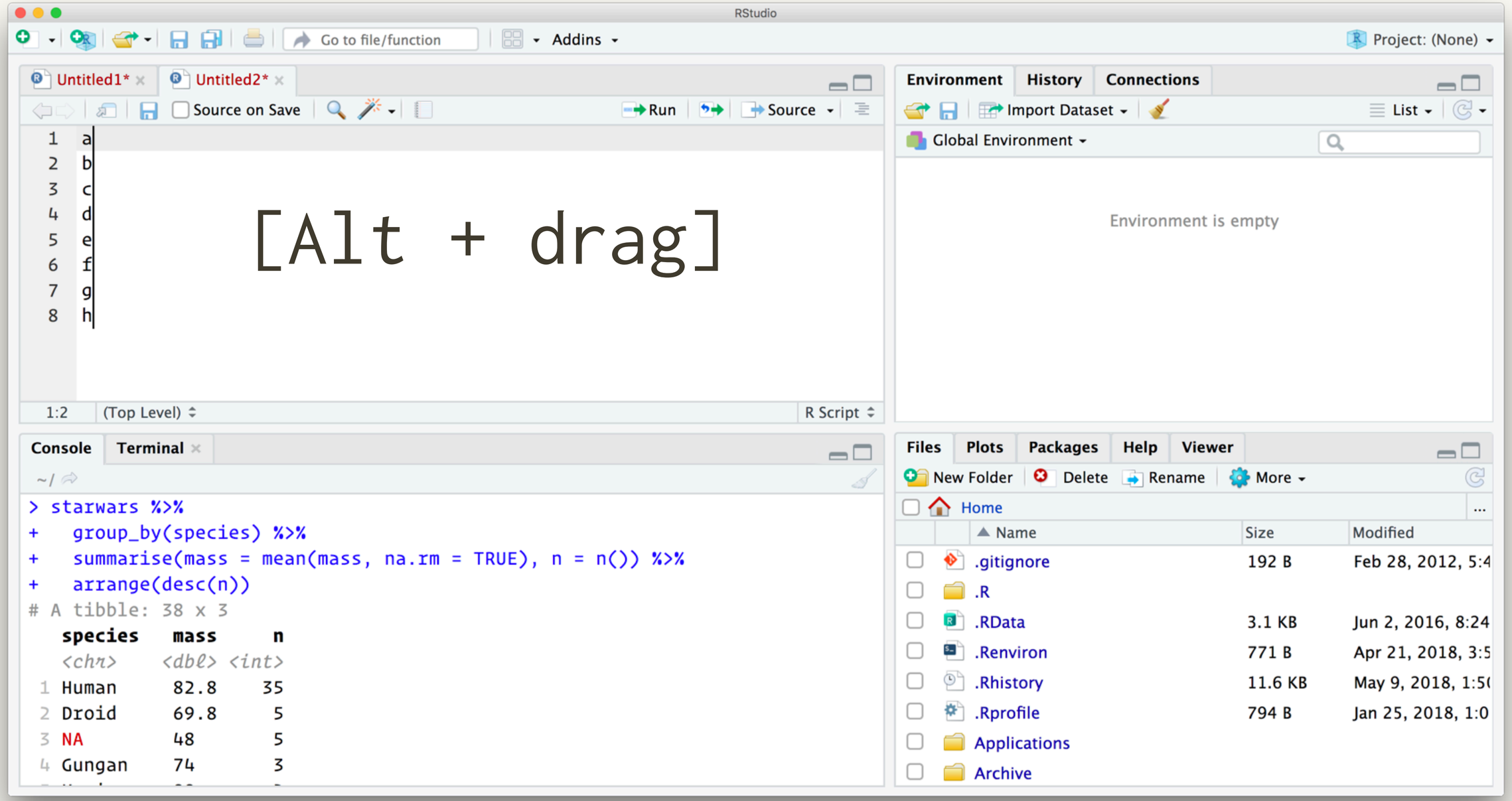
You don’t have to use RStudio,
but if you do, try to master it!

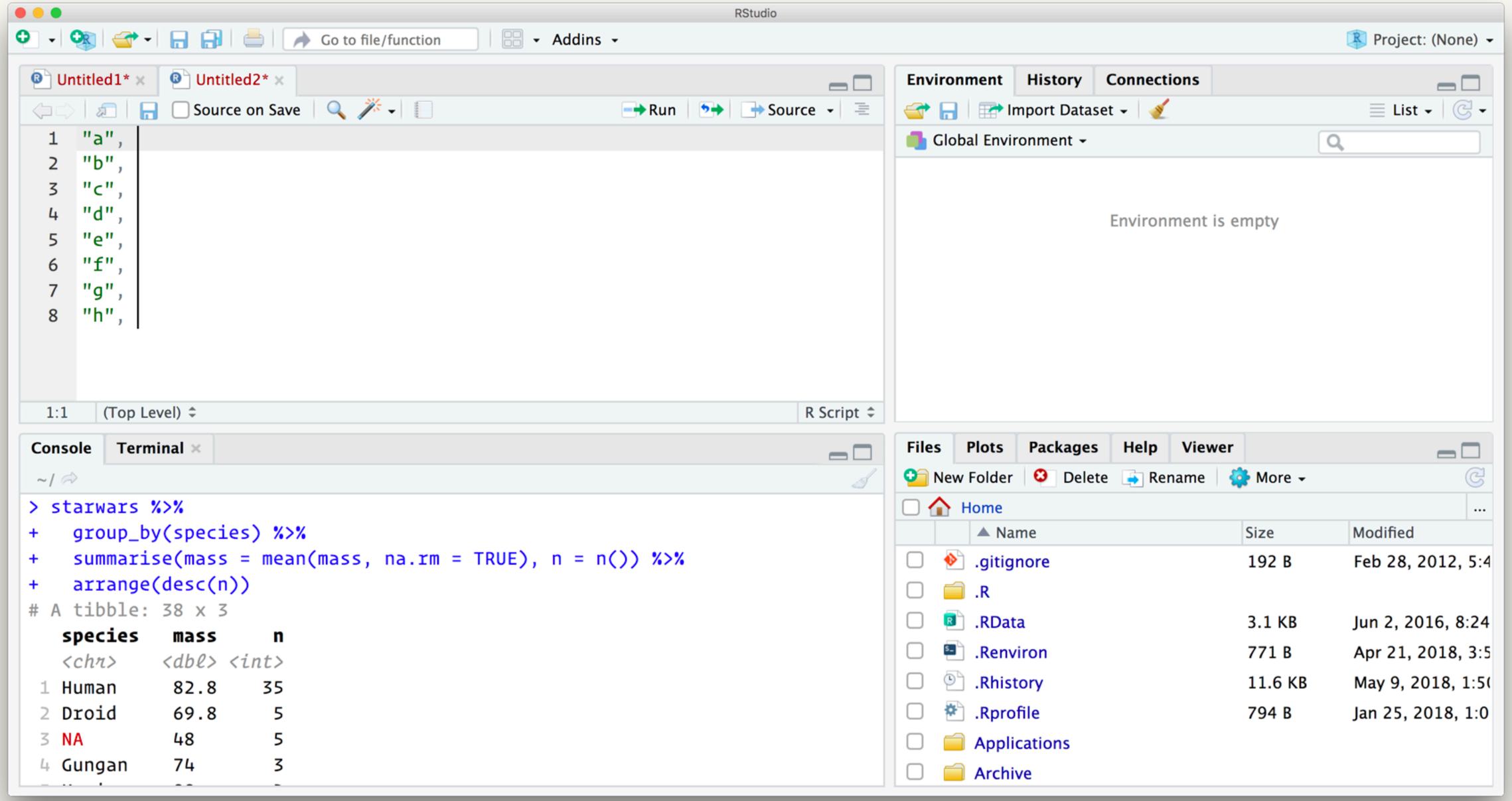

RStudio
Go to file/function
Addins
Project: (None)
Untitled1
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Console
Terminal
~/
> gg
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.gitignore
192 B
Feb 28, 2012, 5:4
.R
.RData
3.1 KB
Jun 2, 2016, 8:24
.Renviron
771 B
Apr 21, 2018, 3:5
.Rhistory
11.6 KB
May 9, 2018, 1:5(
.Rprofile
794 B
Jan 25, 2018, 1:0
Applications
Archive

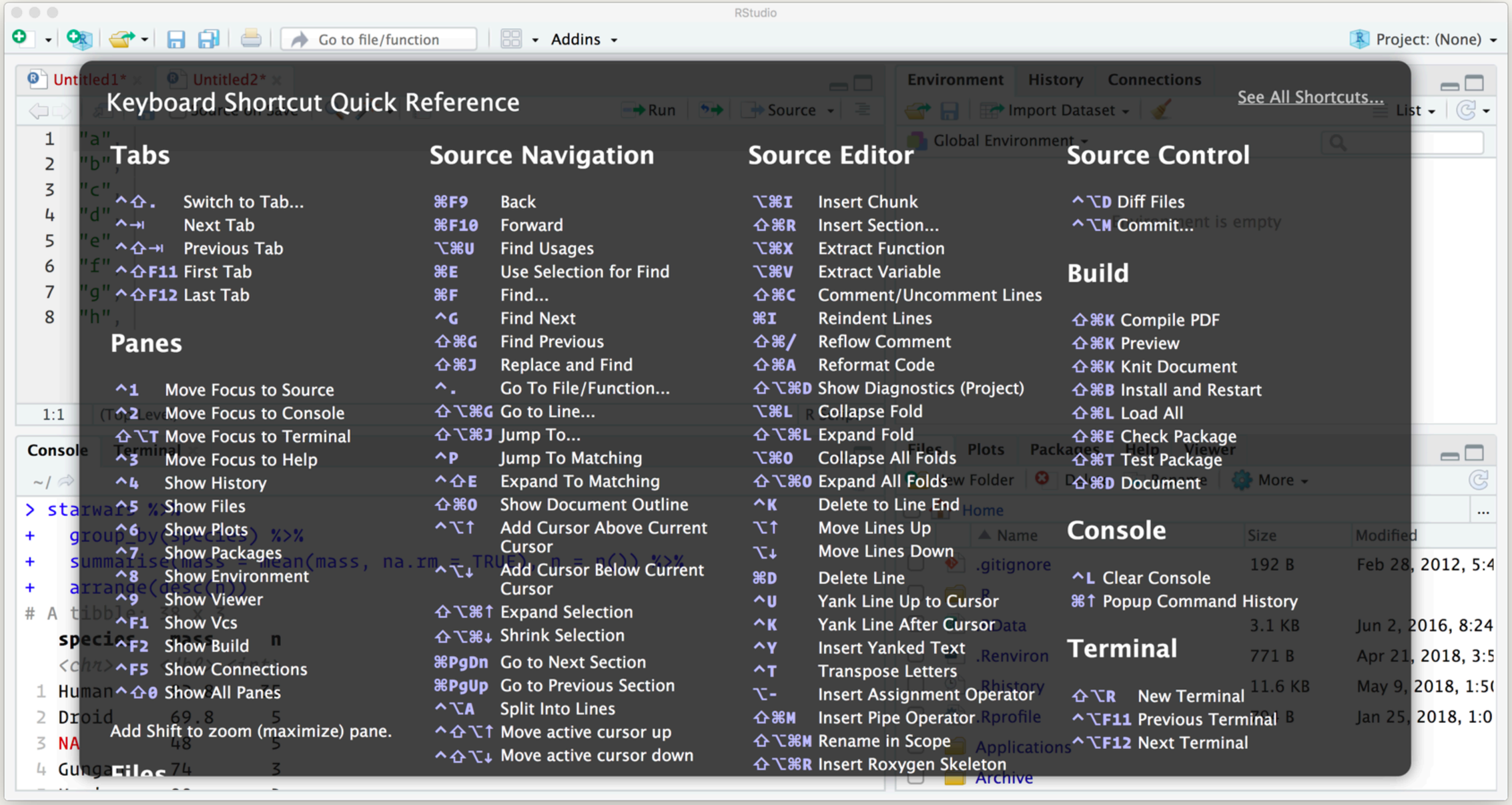

RStudio
Go to file/function
Addins
Project: (None)
Untitled1
Source on Save
Run
Source
ggplot(fake_df, aes(x=x,y=y)) + geom_fan() + theme_bw()
ggplot(fake_df, aes(x=x,y=y)) + geom_fan() + theme_bw() + scale_fill_disti
ggplot(fake_df, aes(x=x,y=y)) + geom_fan() + theme_bw() + scale_fill_gradi
ggplot(fake_df, aes(x=x, y=y)) + geom_fan(intervals=c(50,80,95)/100) +
ggplot(fake_df_sex, aes(x=x,y=y)) + geom_fan() + theme_bw() + scale_fill_g
ggplot(.x, aes(carat, price)) +
ggsave(.x, .y, width = 6, height = 6)
ggplot(df, aes(carat, price)) +
ggtitle(title)
ggplot(mtcars, aes(mpg, cyl)) + aes(colour = NULL, linetype = NULL)
ggplot(mtcars, aes(mpg, cyl)) + aes(colour = NULL, linetype = NULL) + geom
> gg
[Cmd/Ctrl + ↑]
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.gitignore
192 B
Feb 28, 2012, 5:4
.R
.RData
3.1 KB
Jun 2, 2016, 8:24
.Renviron
771 B
Apr 21, 2018, 3:5
.Rhistory
11.6 KB
May 9, 2018, 1:5(
.Rprofile
794 B
Jan 25, 2018, 1:0
Applications
Archive


RStudio
Go to file/function
Addins
Project: (None)
Untitled1*
Source on Save
Run
Source
1 library(tidyverse)
2
3 starwars %>%
4 group_by(species) %>%
5 summarise(mass = mean(mass, na.rm = TRUE), n = n()) %>%
6 arrange(desc(n))
7
[Cmd/ctrl + enter]
3:5
(Top Level)
R Script
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Console
Terminal
~/
>
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.gitignore
192 B
Feb 28, 2012, 5:4
.R
.RData
3.1 KB
Jun 2, 2016, 8:24
.Renviron
771 B
Apr 21, 2018, 3:5
.Rhistory
11.6 KB
May 9, 2018, 1:5
.Rprofile
794 B
Jan 25, 2018, 1:0
Applications
Archive


RStudio
Go to file/function
Addins
Project: (None)
Untitled1*
Source on Save
Run
Source
1 library(tidyverse)
2
3 starwars %>%
4 group_by(species) %>%
5 summarise(mass = mean(mass, na.rm = TRUE), n = n()) %>%
6 arrange(desc(n))
7
[Cmd/ctrl + enter]
7:1
(Top Level)
R Script
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Console
Terminal
~/
> starwars %>%
+ group_by(species) %>%
+ summarise(mass = mean(mass, na.rm = TRUE), n = n()) %>%
+ arrange(desc(n))
# A tibble: 38 x 3
species mass n
<chr> <dbl> <int>
1 Human 82.8 35
2 Droid 69.8 5
3 NA 48 5
4 Gungan 74 3
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.gitignore 192 B Feb 28, 2012, 5:4
.R
.RData 3.1 KB Jun 2, 2016, 8:24
.Renviron 771 B Apr 21, 2018, 3:5
.Rhistory 11.6 KB May 9, 2018, 1:5(
.Rprofile 794 B Jan 25, 2018, 1:0
Applications
Archive


RStudio
Go to file/function
Addins
Project: (None)
Untitled1*
Untitled2*
Source on Save
Run
Source
1 a
2 b
3 c
4 d
5 e
6 f
7 g
8 h
[Alt + drag]
1:2
(Top Level)
R Script
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Console
Terminal
~/
> starwars %>%
+ group_by(species) %>%
+ summarise(mass = mean(mass, na.rm = TRUE), n = n()) %>%
+ arrange(desc(n))
# A tibble: 38 x 3
species mass n
<chr> <dbl> <int>
1 Human 82.8 35
2 Droid 69.8 5
3 NA 48 5
4 Gungan 74 3
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.gitignore 192 B Feb 28, 2012, 5:4
.R
.RData 3.1 KB Jun 2, 2016, 8:24
.Renviron 771 B Apr 21, 2018, 3:5
.Rhistory 11.6 KB May 9, 2018, 1:5(
.Rprofile 794 B Jan 25, 2018, 1:0
Applications
Archive


RStudio
Go to file/function
Addins
Project: (None)
Untitled1*
Untitled2*
Source on Save
Run
Source
"a",
"b",
"c",
"d",
"e",
"f",
"g",
"h",
1:1
(Top Level)
R Script
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Console
Terminal
~/
> starwars %>%
+ group_by(species) %>%
+ summarise(mass = mean(mass, na.rm = TRUE), n = n()) %>%
+ arrange(desc(n))
# A tibble: 38 x 3
species mass n
<chr> <dbl> <int>
1 Human 82.8 35
2 Droid 69.8 5
3 NA 48 5
4 Gungan 74 3
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.gitignore 192 B Feb 28, 2012, 5:4
.R
.RData 3.1 KB Jun 2, 2016, 8:24
.Renviron 771 B Apr 21, 2018, 3:5
.Rhistory 11.6 KB May 9, 2018, 1:5(
.Rprofile 794 B Jan 25, 2018, 1:0
Applications
Archive

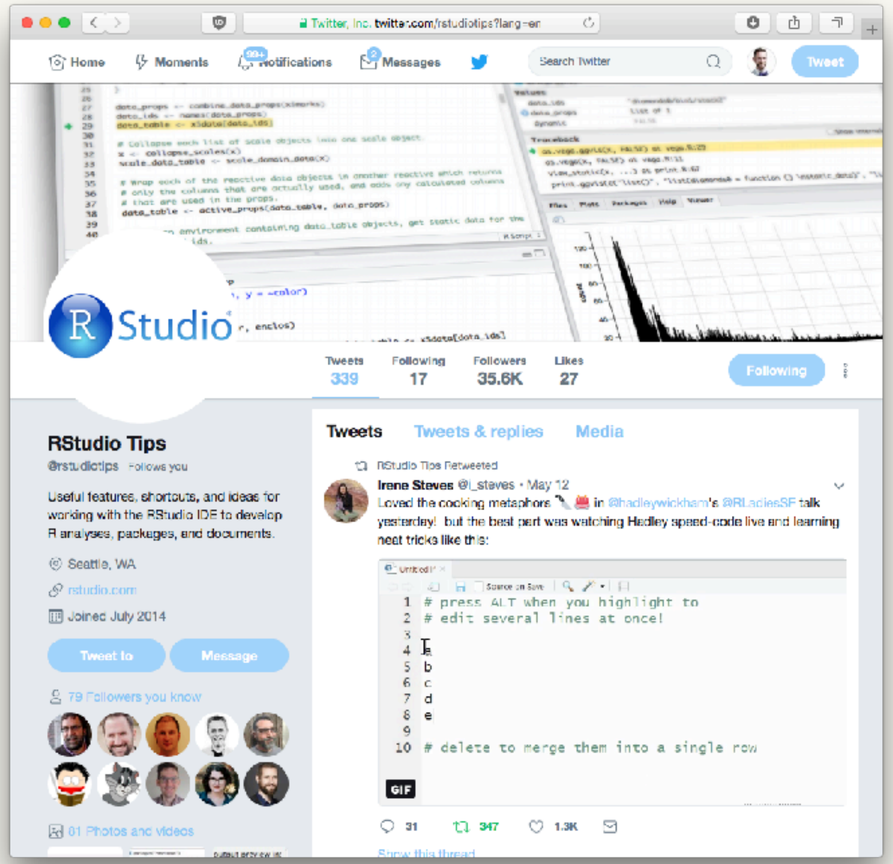

RStudio
Go to file/function
Addins
Project: (None)
Keyboard Shortcut Quick Reference
See All Shortcuts...
Tabs
^⇧. Switch to Tab...
^→⇥ Next Tab
^⇧→⇥ Previous Tab
^⇧F11 First Tab
^⇧F12 Last Tab
Panes
^1 Move Focus to Source
^2 Move Focus to Console
⇧⌥T Move Focus to Terminal
^3 Move Focus to Help
^4 Show History
^5 Show Files
^6 Show Plots
^7 Show Packages
^8 Show Environment
^9 Show Viewer
^F1 Show Vcs
^F2 Show Build
^F5 Show Connections
^⇧0 Show All Panes
Add Shift to zoom (maximize) pane.
Files
Source Navigation
⌘F9 Back
⌘F10 Forward
⌥⌘U Find Usages
⌘E Use Selection for Find
⌘F Find...
^G Find Next
⇧⌘G Find Previous
⇧⌘J Replace and Find
^. Go To File/Function...
⇧⌥⌘G Go to Line...
⇧⌥⌘J Jump To...
^P Jump To Matching
^⇧E Expand To Matching
⇧⌘O Show Document Outline
^⌥↑ Add Cursor Above Current Cursor
^⌥↓ Add Cursor Below Current Cursor
⇧⌥⌘↑ Expand Selection
⇧⌥⌘↓ Shrink Selection
⌘PgDn Go to Next Section
⌘PgUp Go to Previous Section
^⌥A Split Into Lines
^⇧⌥↑ Move active cursor up
^⇧⌥↓ Move active cursor down
Source Editor
⌥⌘I Insert Chunk
⇧⌘R Insert Section...
⌥⌘X Extract Function
⌥⌘V Extract Variable
⇧⌘C Comment/Uncomment Lines
⌘I Reindent Lines
⇧⌘/ Reflow Comment
⇧⌘A Reformat Code
⇧⌥⌘D Show Diagnostics (Project)
⌥⌘L Collapse Fold
⇧⌥⌘L Expand Fold
⌥⌘O Collapse All Folds
⇧⌥⌘O Expand All Folds
^K Delete to Line End
⌥↑ Move Lines Up
⌥↓ Move Lines Down
⌘D Delete Line
^U Yank Line Up to Cursor
^K Yank Line After Cursor
^Y Insert Yanked Text
^T Transpose Letters
⌥- Insert Assignment Operator
⇧⌘M Insert Pipe Operator
⇧⌥⌘M Rename in Scope
⇧⌥⌘R Insert Roxygen Skeleton
Source Control
^⌥D Diff Files
^⌥M Commit...
Build
⇧⌘K Compile PDF
⇧⌘K Preview
⇧⌘K Knit Document
⇧⌘B Install and Restart
⇧⌘L Load All
⇧⌘E Check Package
⇧⌘T Test Package
⇧⌘D Document
Console
^L Clear Console
⌘↑ Popup Command History
Terminal
⇧⌥R New Terminal
^⌥F11 Previous Terminal
^⌥F12 Next Terminal
Environment
History
Connections
Import Dataset
List
Global Environment
Console
Terminal
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Name
Size
Modified
.gitignore
192 B
Feb 28, 2012, 5:4
.Rhistory
771 B
Apr 21, 2018, 3:5
.Rprofile
11.6 KB
May 9, 2018, 1:5
Applications
Archive
Jan 25, 2018, 1:0

Your turn

What's the shortcut for <- (assignment)?

What about %>% (pipe)?

How can you quickly comment a block of lines?



@rstudiotips