

¡Bienvenidos al taller CDSB 2019!

1. Red Wifi: LCG4 password: aula4LCG+
2. Get the course materials

```
usethis::use_course("ComunidadBioinfo/cdsb2019")
```
3. Verifica que tienes los paquetes instalados corriendo setup.R
4. ¿Estás atorado? ¡Pide ayuda!

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Preliminares

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

HOLA

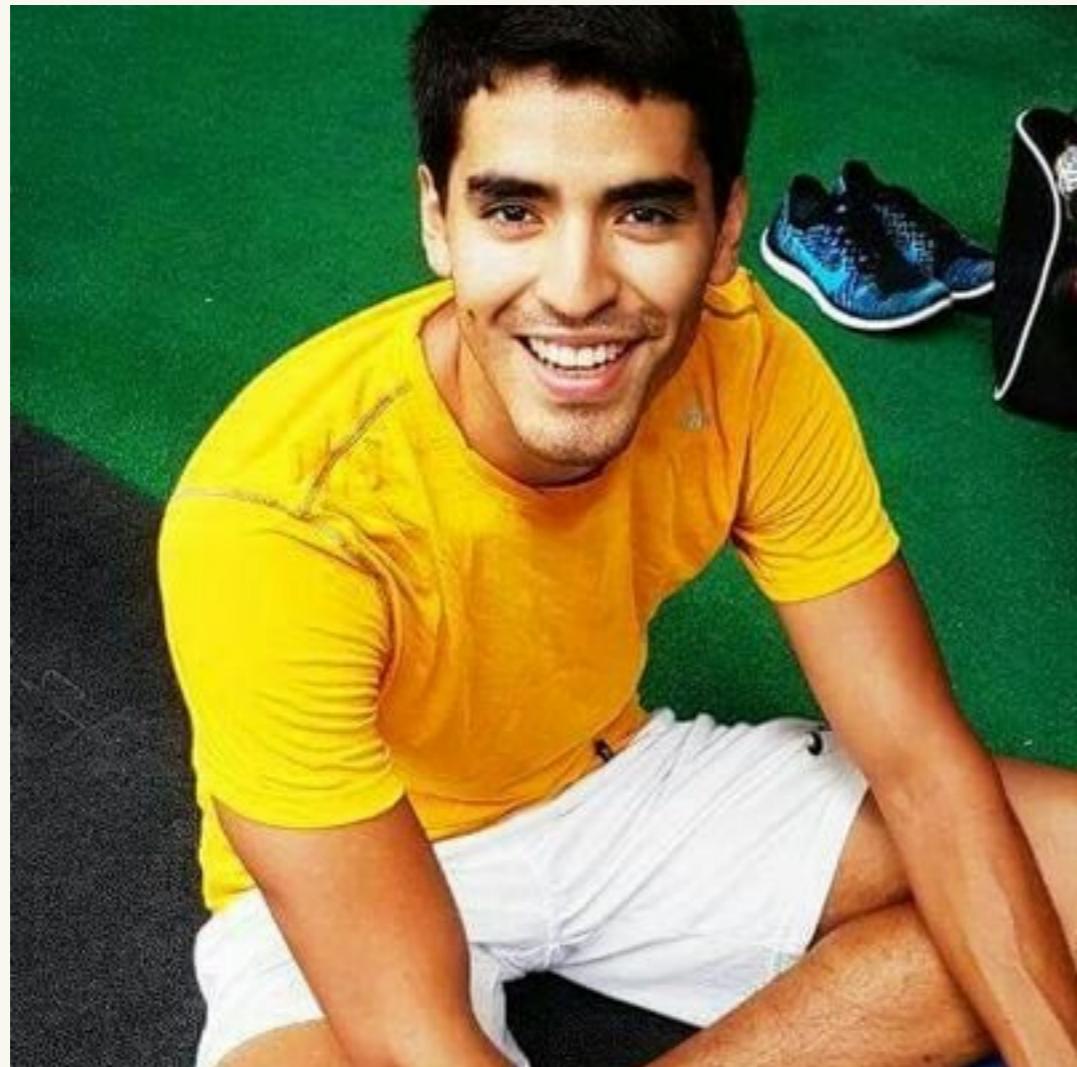
mi nombre es

Leonardo

lcolladtor.github.io

@fellgernon

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Alejandro Reyes

Ph.D., postdoc

Dana-Farber Data
Sciences, Harvard
University

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Alejandra Medina-Rivera

Ph.D., Principal
Investigator
LIIGH-UNAM

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



María Teresa Ortíz
M.A., Data Scientist
& Lecturer
CONABIO and ITAM

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Charlotte Wickham

Ph.D., Oregon State
University

Creadora del taller
"Building Tidy Tools"
de rstudio::conf(2019)

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Hadley Wickham

Ph.D., Chief Scientist
RStudio

Co-creador del taller
"Building Tidy Tools"

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Tu turno

Este curso tu escribirás código, y mientras estamos aquí para ayudarte, la mejor fuente es la persona que está sentada junto a ti.

Preséntate a tus vecinos. ¿Quién eres y para qué usar R?

¡Esto significa que tendrás que trabajar!

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Objetivo: ayudarte a escribir herramientas tidy

Escribir funciones

Funciones individuales para resolver problemas individuales

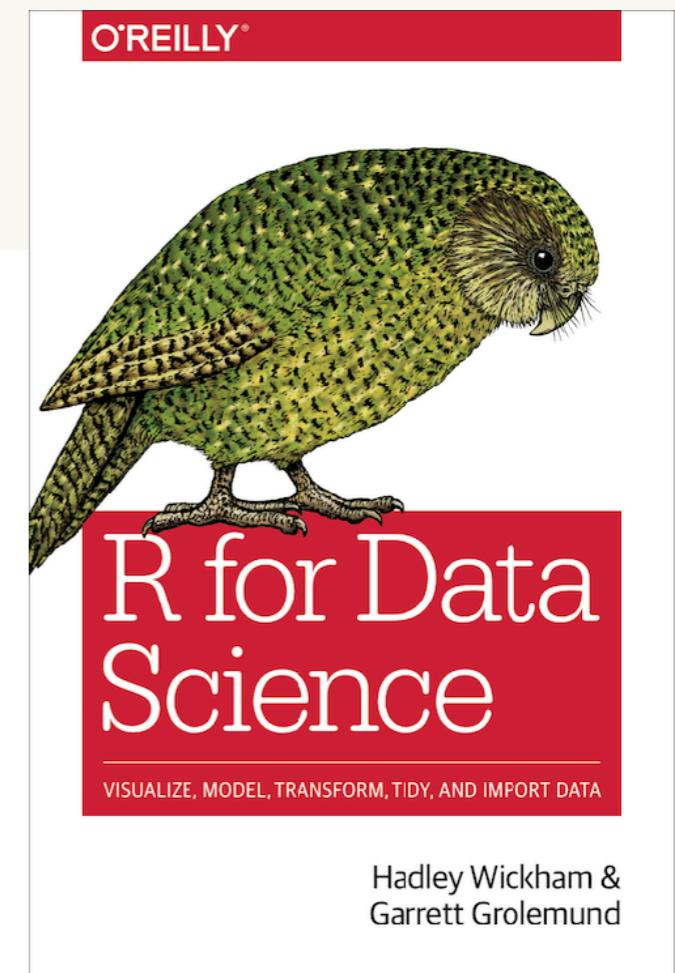
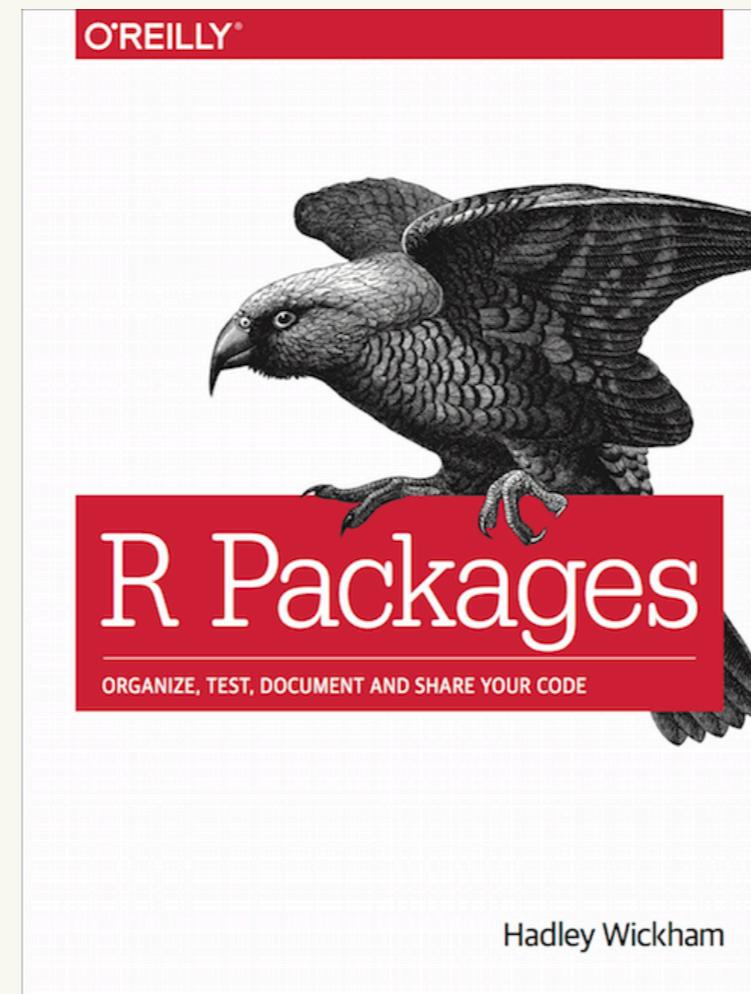
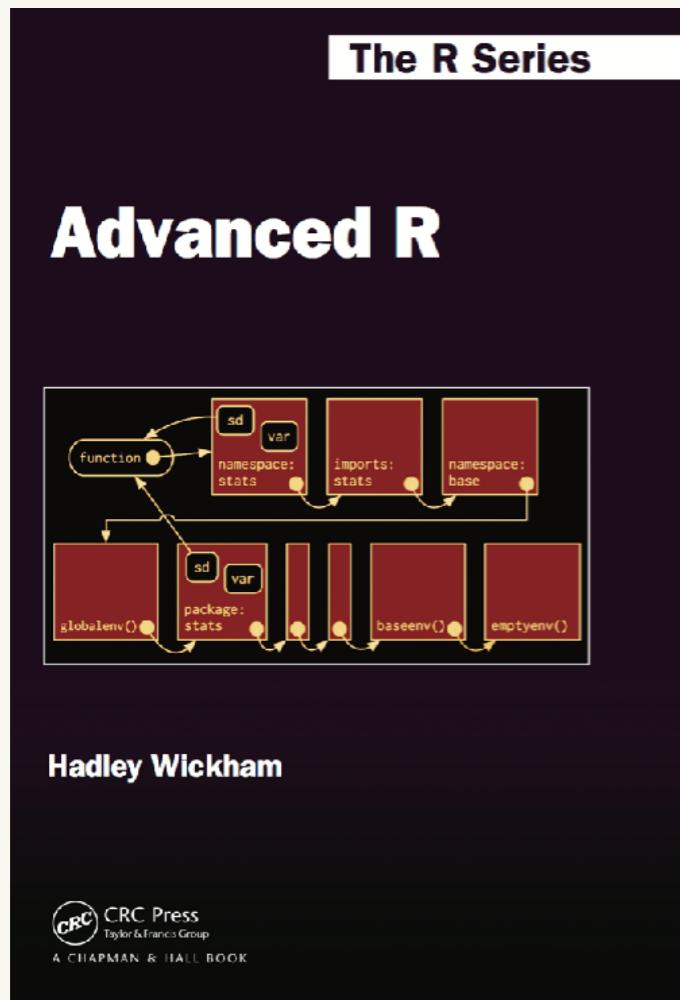


Diseño de APIs

Familia de funciones que trabajan en conjunto para resolver una familia de problemas

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Mucho de este curso está basado en estos libros



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

Working on 2nd ed

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

<https://amzn.com/1491910399>

<http://r4ds.had.co.nz>

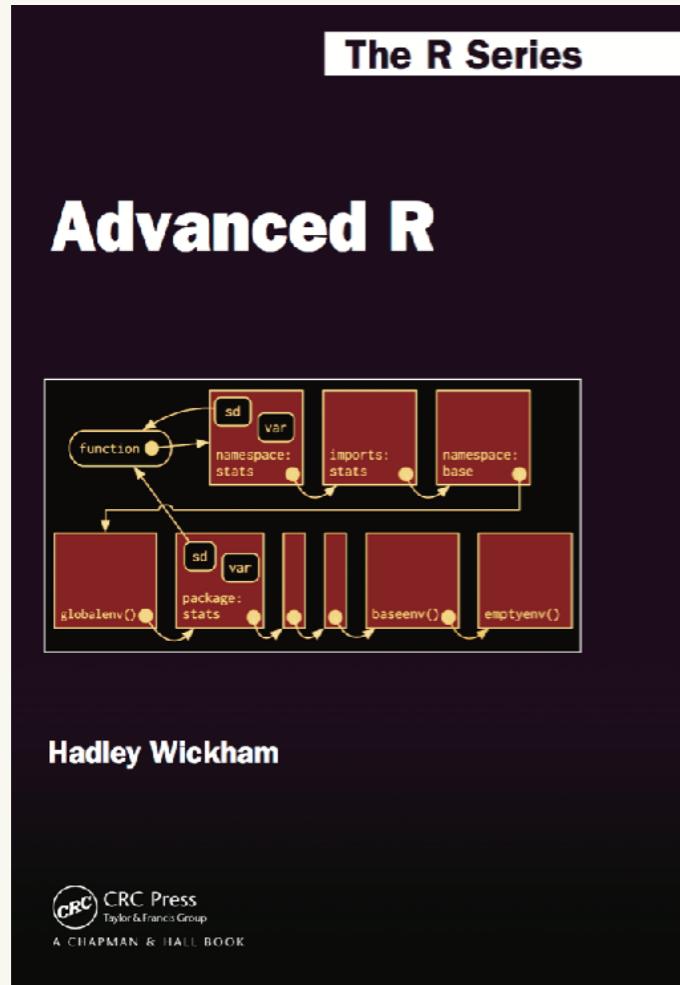
<https://amzn.com/1491910399>

Versión en español:

<https://es.r4ds.hadley.nz/>

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Pero el libro principal todavía no existe



Cómo funciona R

Tidy tools

Hadley Wickham

Como resolver
problemas reales con R

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Horario

Día 1	Día 2	Día 3
Preliminares	Programación funcional	Evaluación "Tidy"
11:00-11:30 Descanso y café	11:00-11:30 Descanso y café	11:00-11:30 Descanso y café
Paquetes y Pruebas	Errores	Documentar y Compartir
14:00-15:30 Comida	14:00-15:30 Comida	14:00-15:30 Comida
Diseño de API	Programación Orientada a Objetos	Introducción a GitHub

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Calentamientos

¡No esperes saber todas las respuestas!

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Tu turno

¿Cuáles son los cuatro tipos de **vectores atómicos**? (Puntos extras para los dos tipos raros)

¿Cuáles son las tres propiedades de un vector?

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Cuatro tipos comunes: logical, integer, double, character

```
typeof(TRUE)
```

```
typeof(1L)
```

```
typeof(1.5)
```

```
typeof("a")
```

Hablaremos de estos después (S3):

```
typeof(factor(1:10))
```

```
typeof(Sys.Date())
```

Cada vector tiene tres propiedades:

```
x <- 1:5
```

```
# 1. Tipo:  
typeof(x)
```

```
# 2. Longitud  
length(x)
```

```
# 3. Atributos  
attributes(x)  
# (¿Tal vez los veremos después?)
```

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Valores faltantes

¿Qué resulta de evaluar `NA == NA`?

¿Por qué?

¿Qué deberías de usar en vez?

No hay un solo valor para representar los valores desconocidos

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

```
is.na(x)
```

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

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

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Tu turno

¿Cuáles son los seis tipos de cosas que podemos meter dentro de []?

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

espacio

incluye todo

integer

+ve: incluye
0: elimina todo
-ve: excluye

logical

retiene los valores TRUEs

character

busca por nombre

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Conjunto de caracteres para búsquedas sencillas

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

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Tu turno

```
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)
#> ???
```



<https://lobstr.r-lib.org/>

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

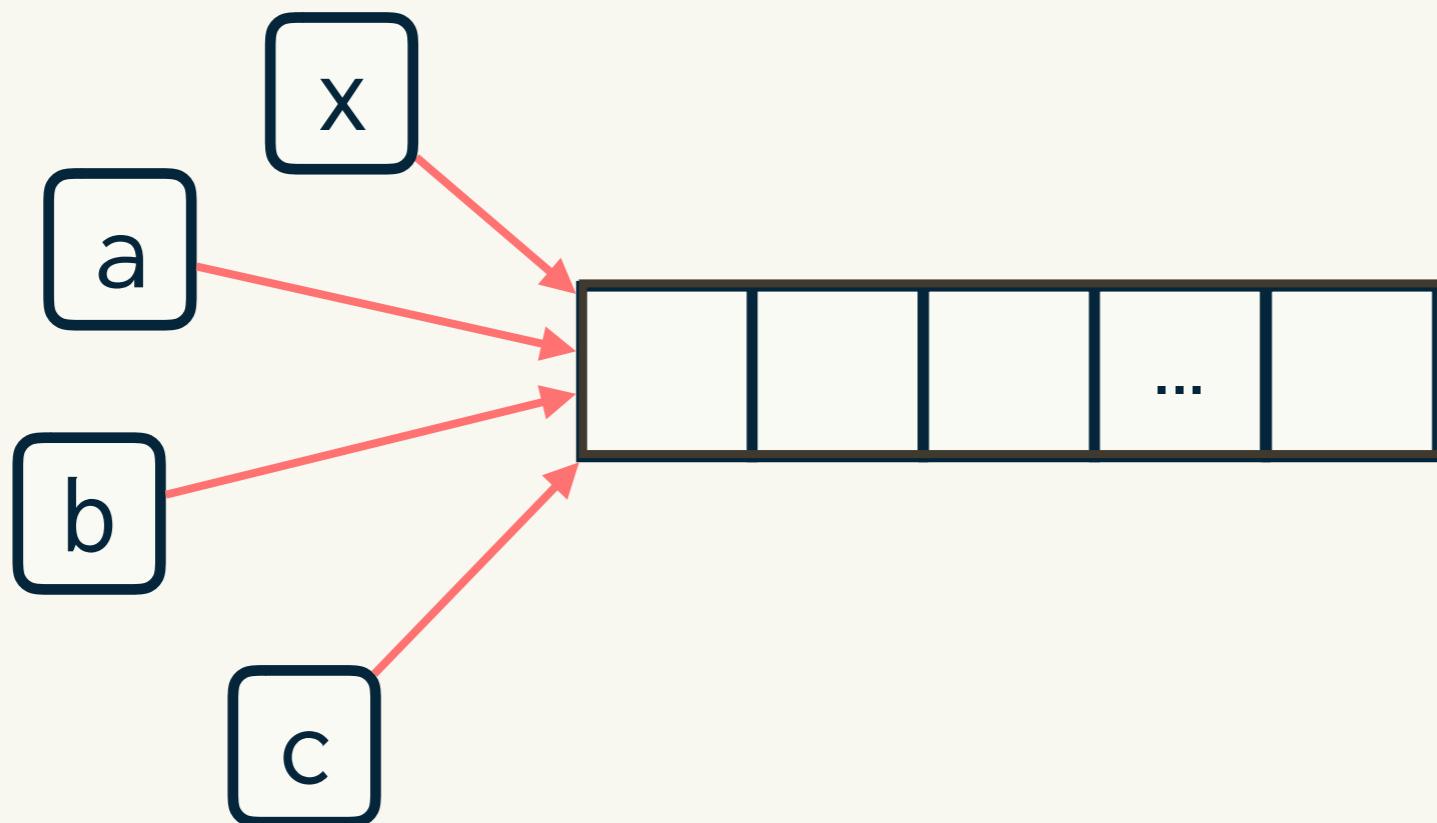
Un nombre es una referencia a algún valor

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



Muchas referencias pueden apuntar a un objeto

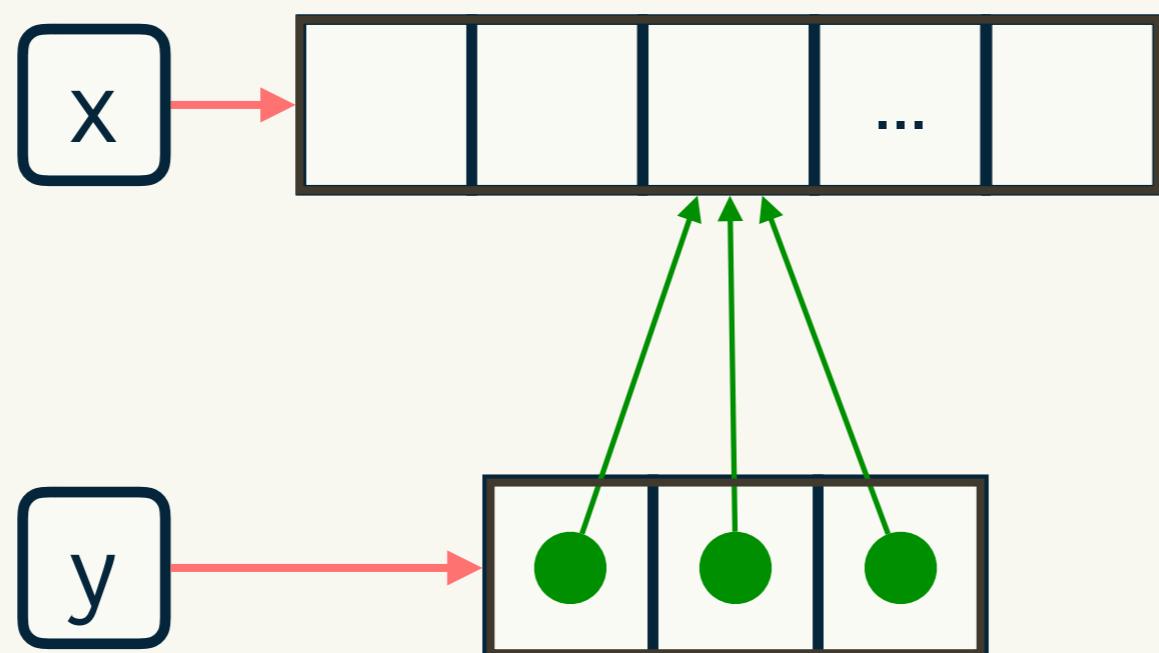
```
a <- b <- c <- x
```



¿Qué tan grande es **a**?
¿Qué tan grande es **b**?
¿Qué tan grande son **a** y **b** juntos?

Elementos de listas también son referencias

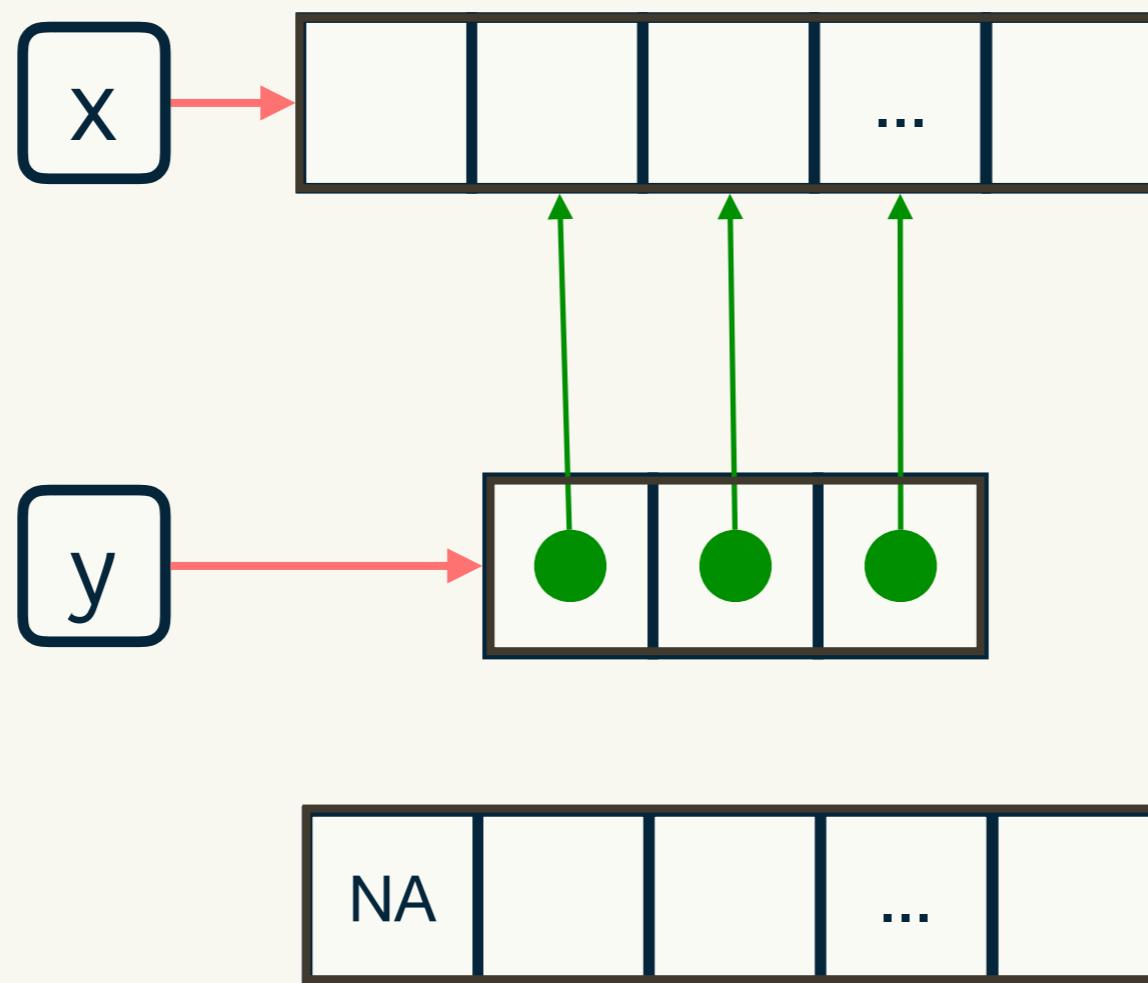
```
y <- list(x, x, x)
```



¿Qué tan grande es **x**?
¿Qué tan grande es **y**?
¿Qué tan grandes son **x** y **y** juntos?

Modificar un objeto crea una copia

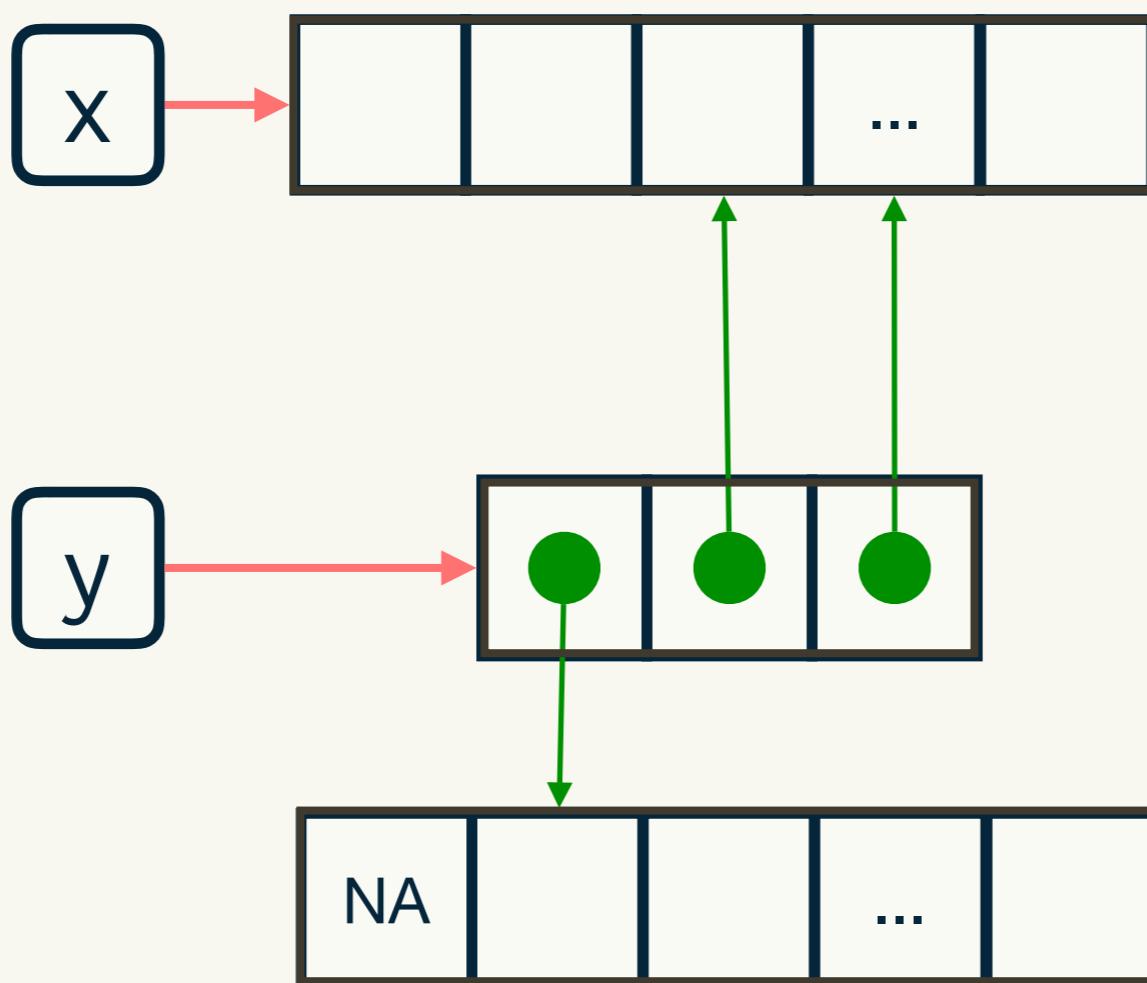
```
y[[1]][[1]] <- NA
```



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Modificar un objeto crea una copia

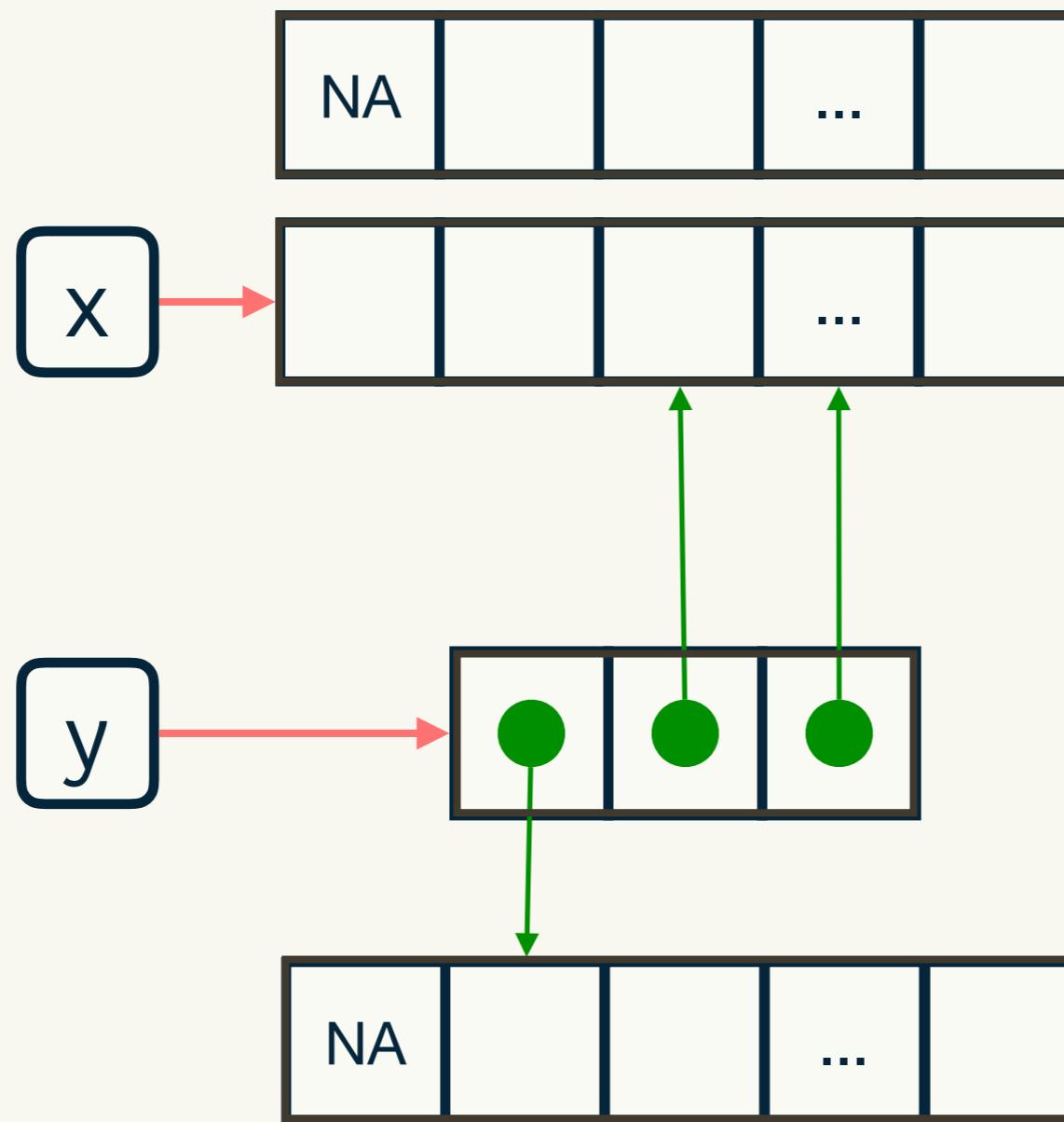
```
y[[1]][[1]] <- NA
```



¿Qué tan grande es **x**?
¿Qué tan grande es **y**?
¿Qué tan grandes son **x** y **y** juntos?

Modificar un objeto crea una copia

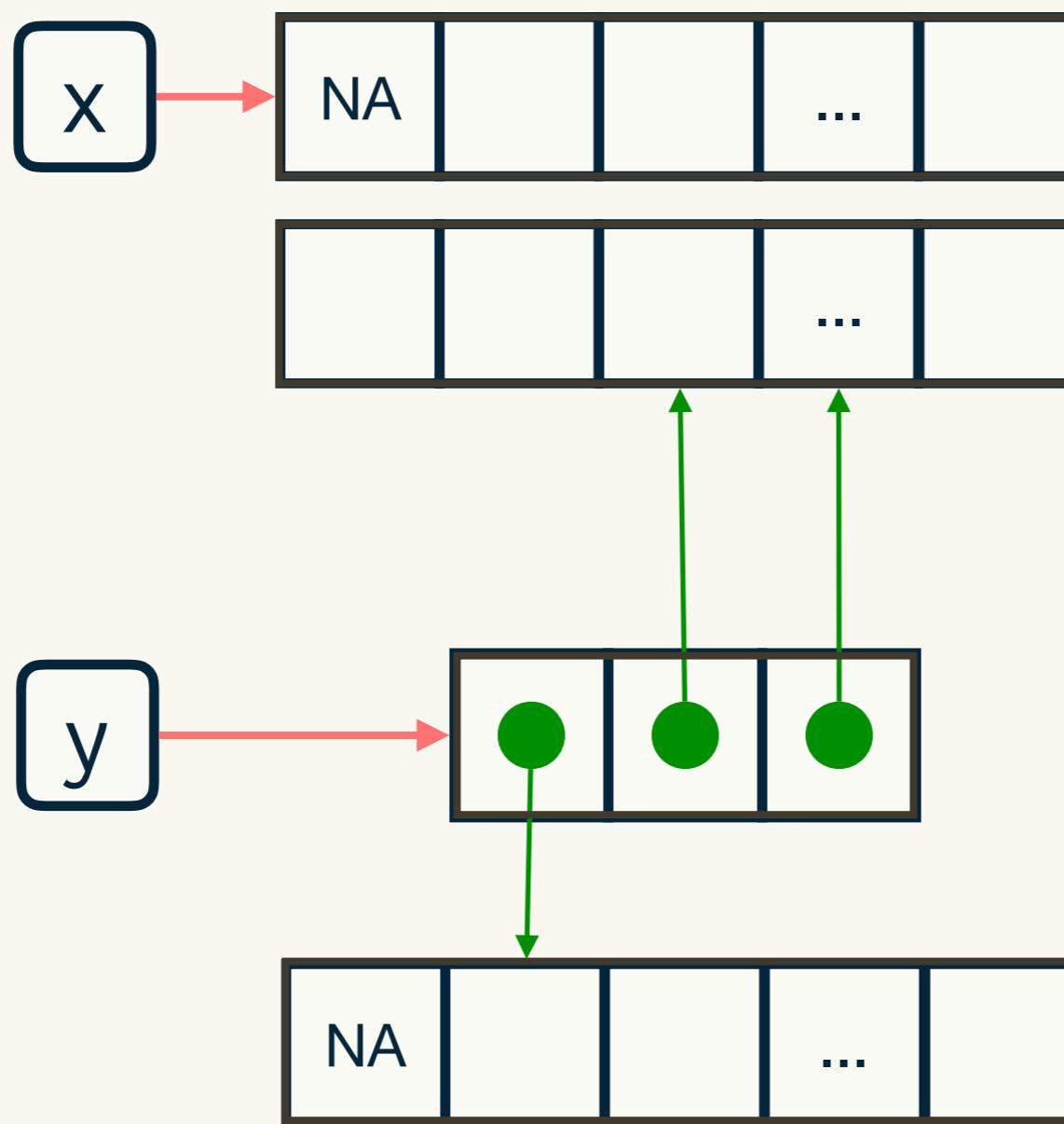
```
x[[1]] <- NA
```



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Modificar un objeto crea una copia

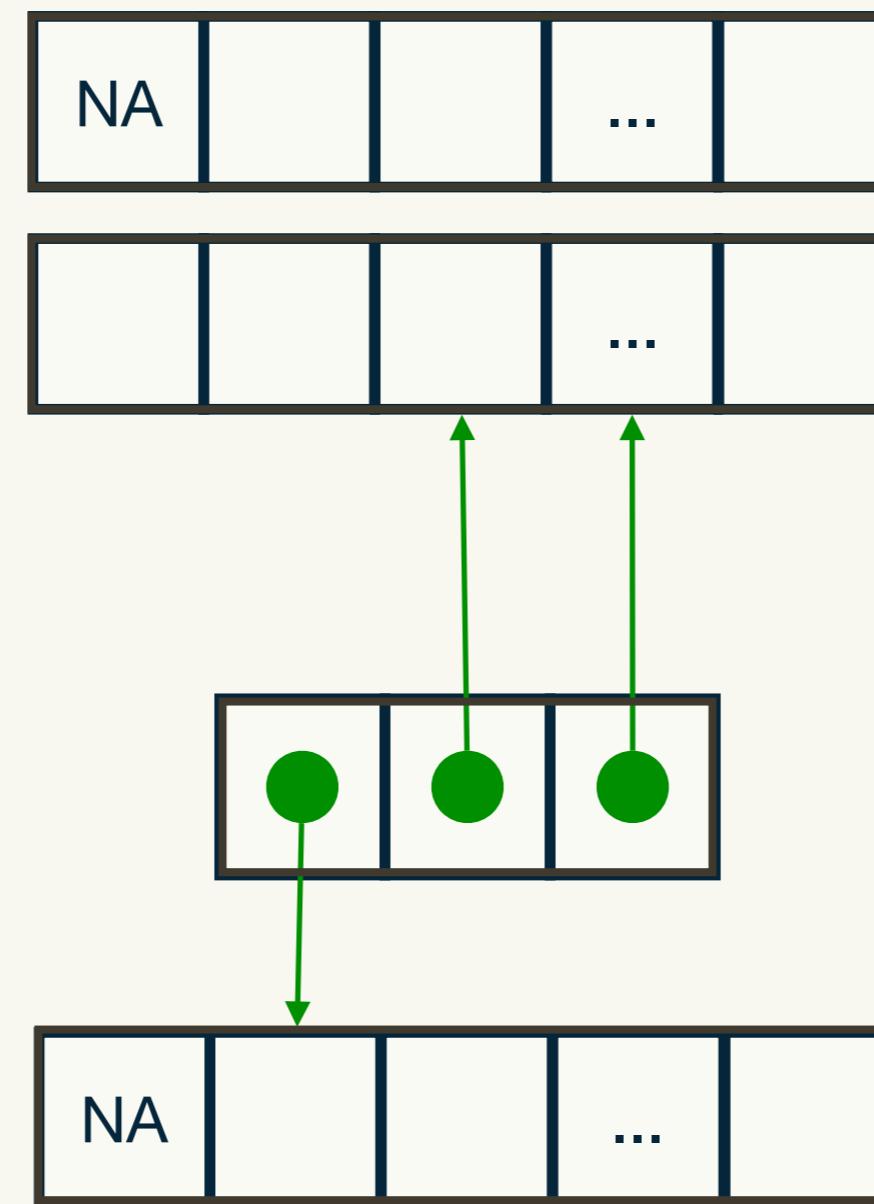
```
x[[1]] <- NA
```



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

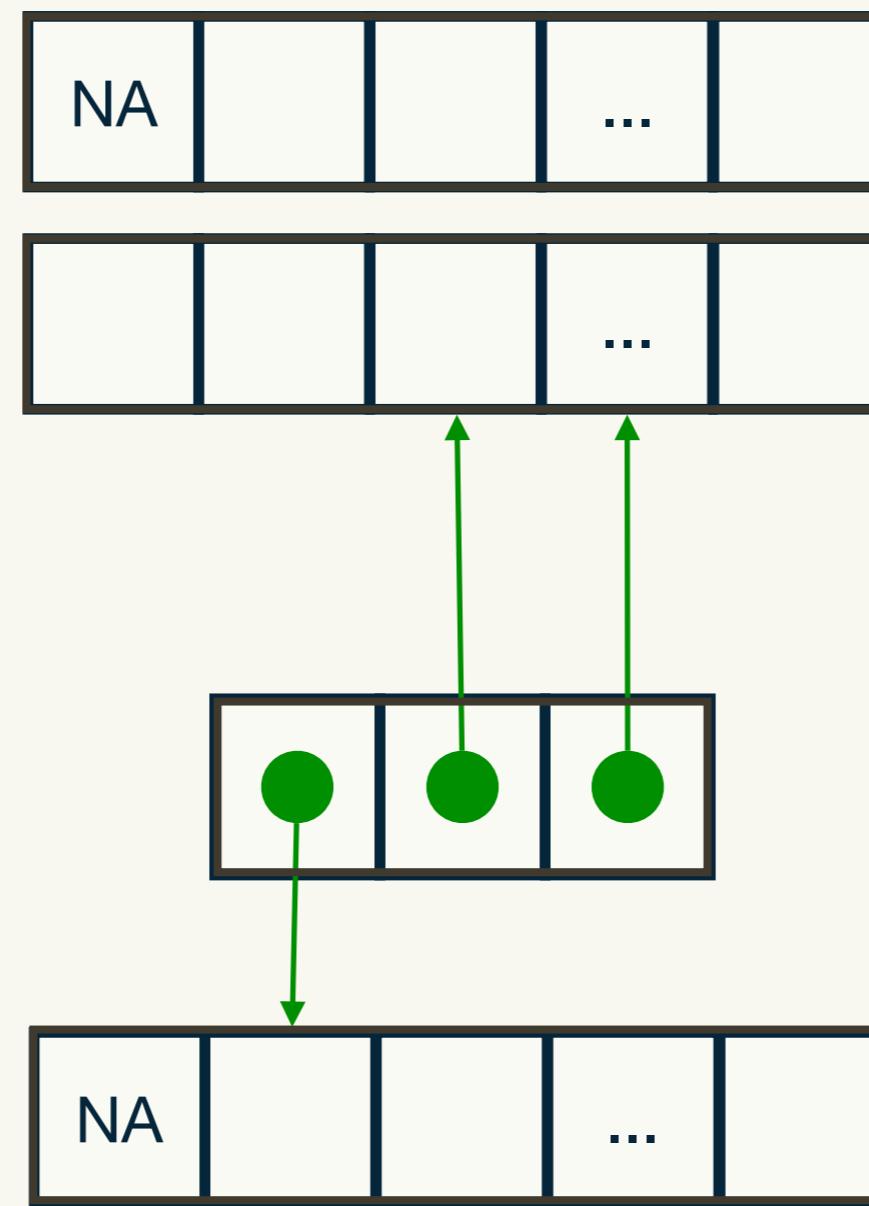
`rm()` elimina referencias

`rm(x, y)`



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

El colector de basura elimina referencias



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

El colector de basura elimina referencias

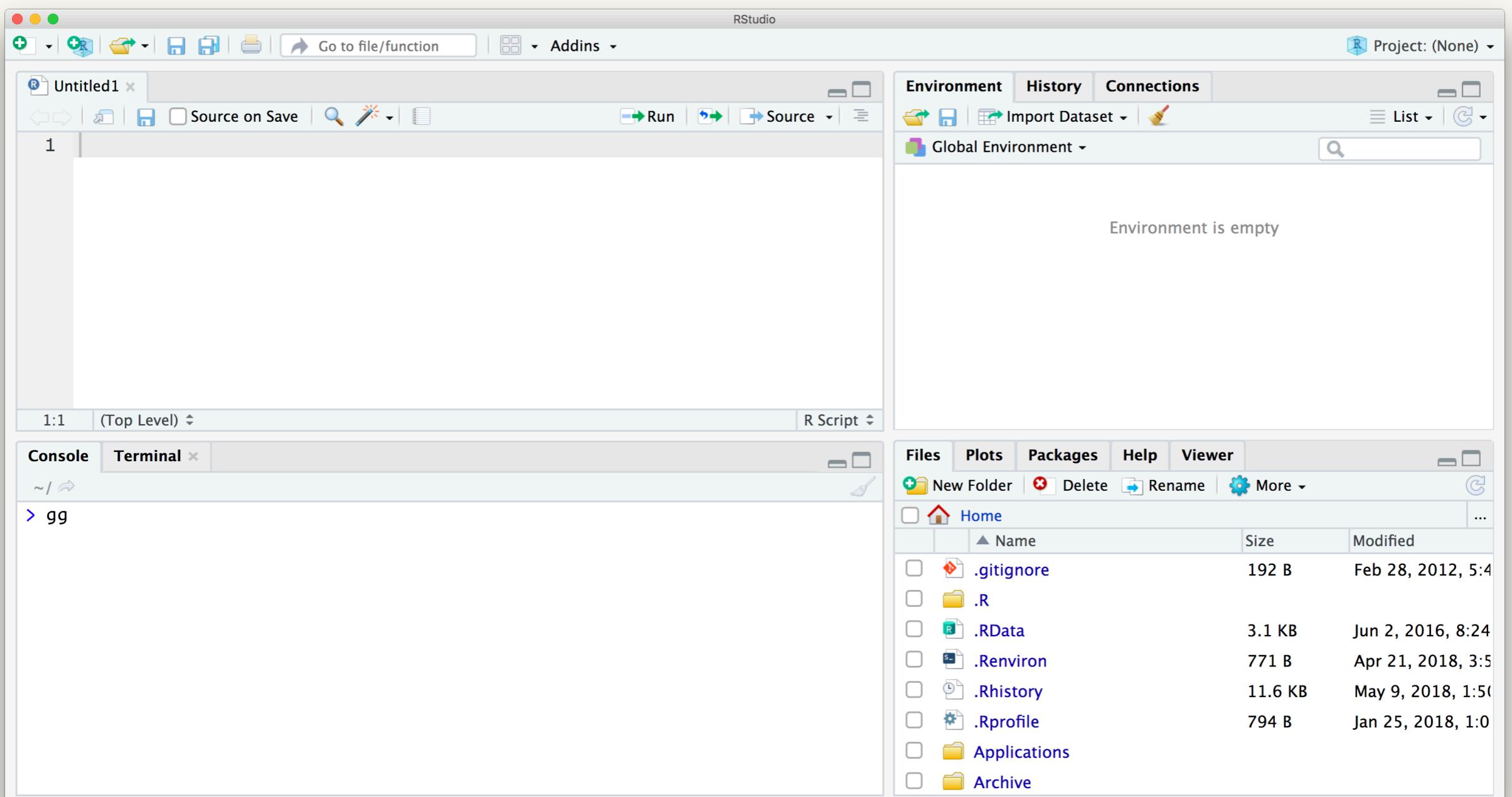


Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

RStudio

No necesitas usar RStudio,
pero si lo usas, ¡intenta dominarlo!

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

The screenshot shows the RStudio interface. The left pane contains a script editor with the following R code:

```
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_distributome()
ggplot(fake_df, aes(x=x,y=y)) + geom_fan() + theme_bw() + scale_fill_gradient()
ggplot(fake_df, aes(x=x, y=y)) + geom_fan(intervals=c(50,80,95)/100) +
  geom_fan(intervals=c(50,80,95)/100)
ggplot(fake_df_sex, aes(x=x,y=y)) + geom_fan() + theme_bw() + scale_fill_grey()
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_fan()
```

The right pane shows the Environment tab with the message "Environment is empty".

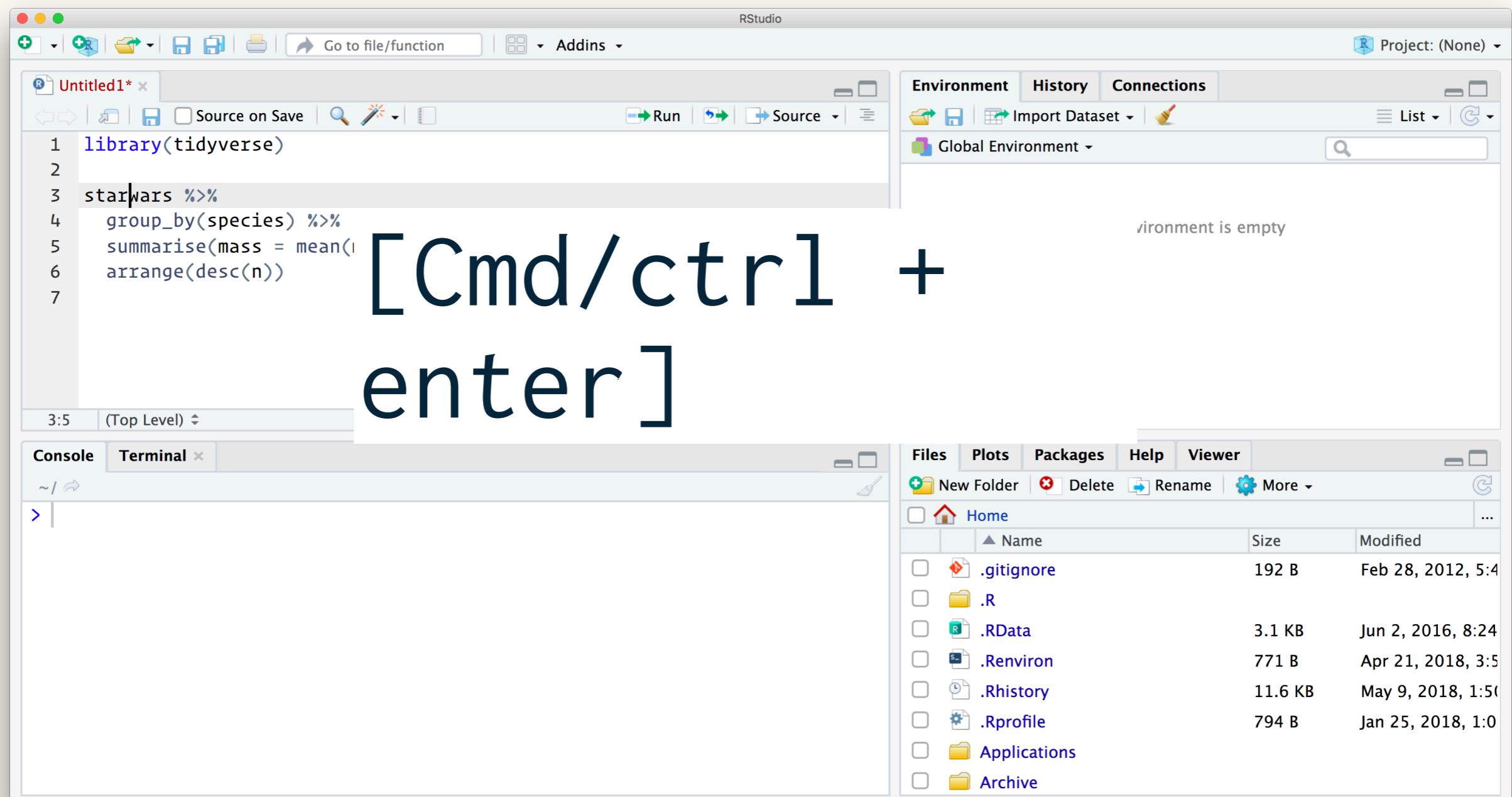
The bottom-right pane shows the Files tab of the file browser with the following contents:

Name	Size	Modified
.gitignore	192 B	Feb 28, 2012, 5:44 PM
.R	3.1 KB	Jun 2, 2016, 8:24 AM
.RData	771 B	Apr 21, 2018, 3:50 PM
.Renviron	11.6 KB	May 9, 2018, 1:50 PM
.Rhistory	794 B	Jan 25, 2018, 1:00 PM
.Rprofile		
Applications		
Archive		

[Cmd/Ctrl]

+ ↑]

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

The screenshot shows the RStudio interface. In the top-left code editor, there is a large text block containing R code. To the right of the code editor, a large, semi-transparent watermark displays the text "[Cmd/ctrl] + enter". The code in the editor is:

```
library(tidyverse)
starwars %>%
  group_by(species) %>%
  summarise(mass = mean(mass, na.rm = TRUE), n = n()) %>%
  arrange(desc(n))
```

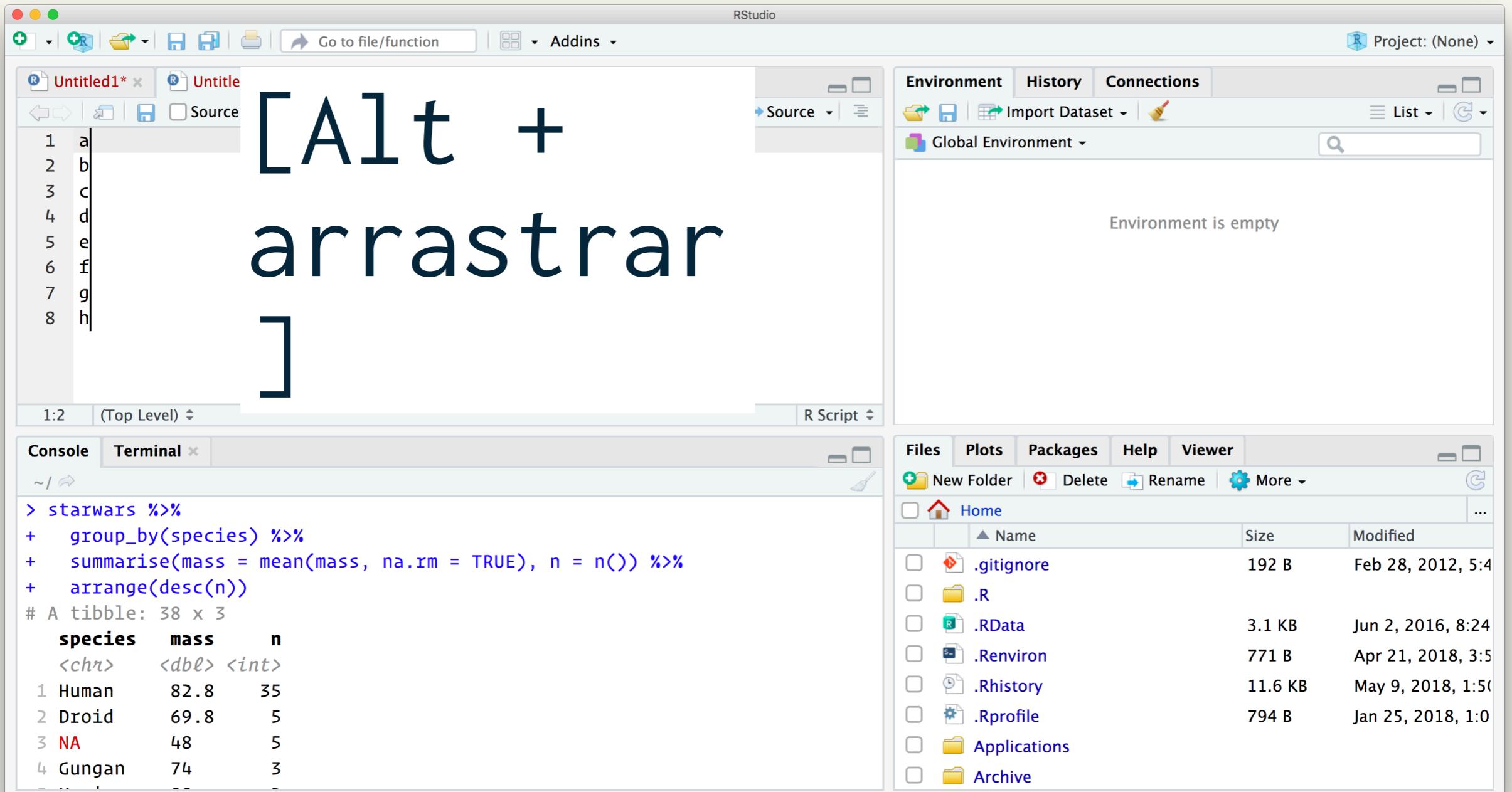
The code is run in the console, which shows the output:

```
> 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
5 Wookiee     100      2
6 Ewok          35      2
7 Dewback      100      1
8 Rancor       100      1
9 Tauntaun     100      1
10 Bantha      100      1
# ... with 33 more rows
```

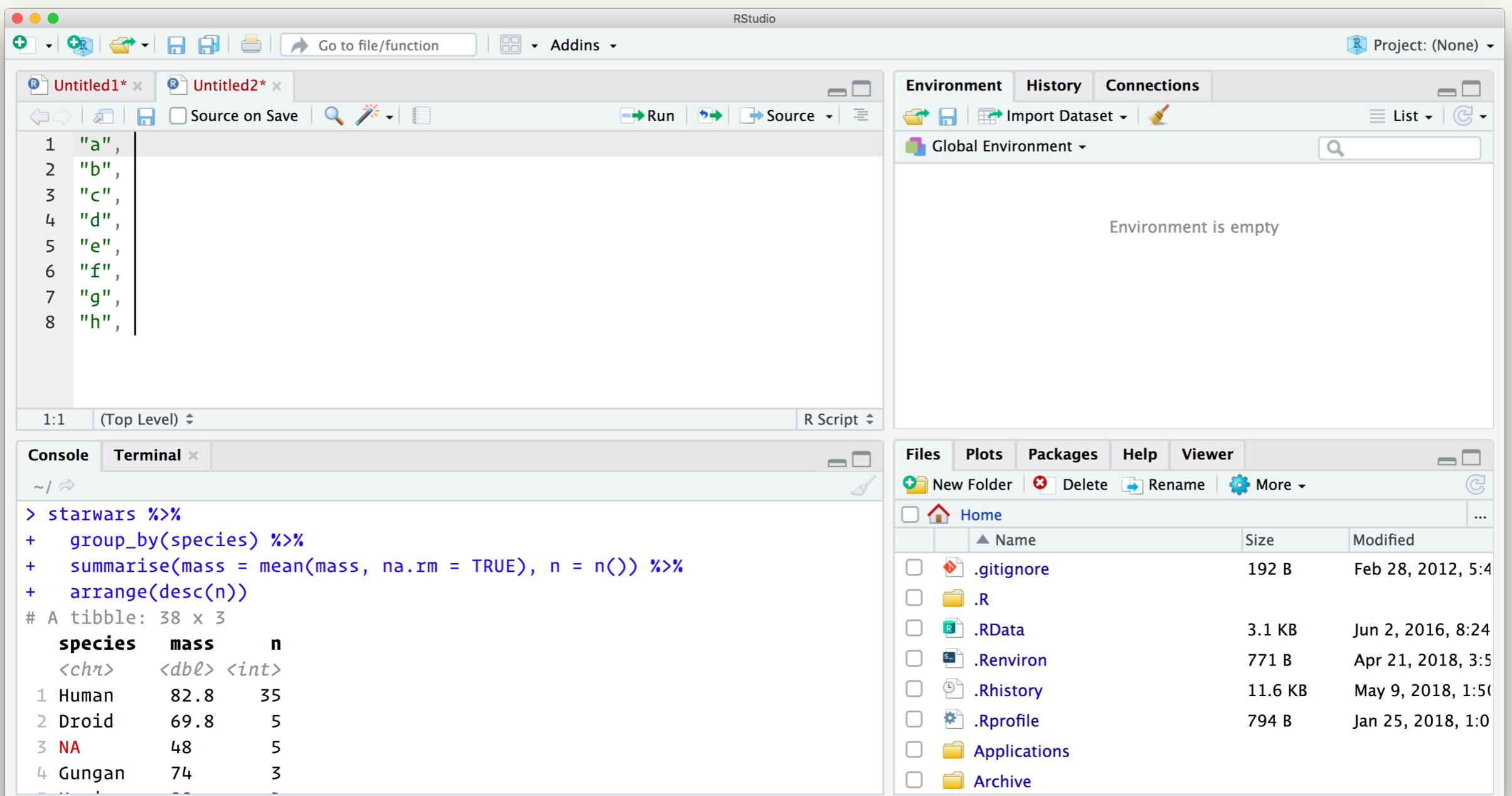
The environment tab in the top-right panel shows "Global Environment" with the message "Environment is empty". The file browser panel shows the contents of the current directory:

Name	Size	Modified
.gitignore	192 B	Feb 28, 2012, 5:4
.R	3.1 KB	Jun 2, 2016, 8:24
.RData	771 B	Apr 21, 2018, 3:5
.Renviron	11.6 KB	May 9, 2018, 1:50
.Rhistory	794 B	Jan 25, 2018, 1:0
Applications		
Archive		

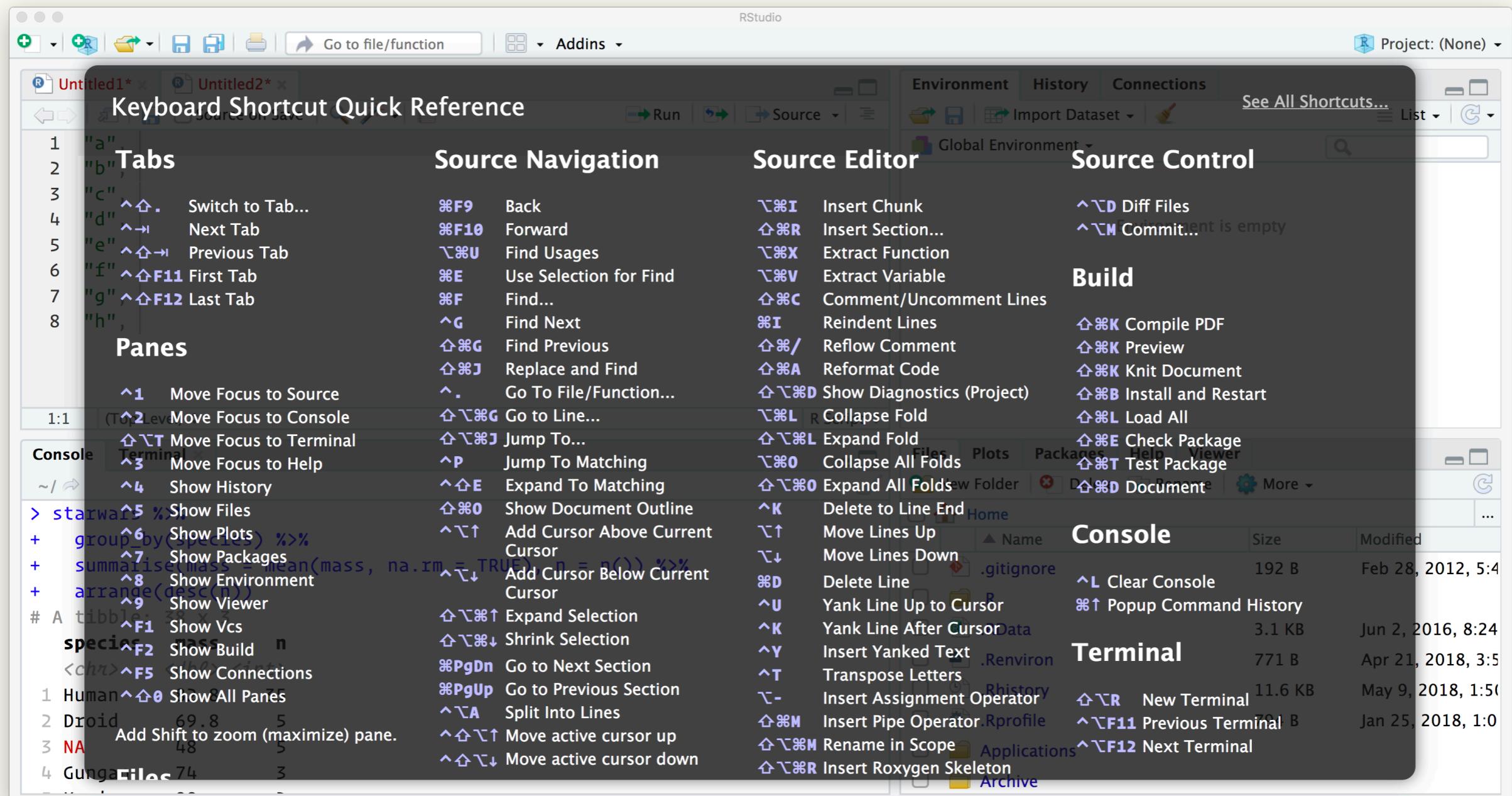
Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')



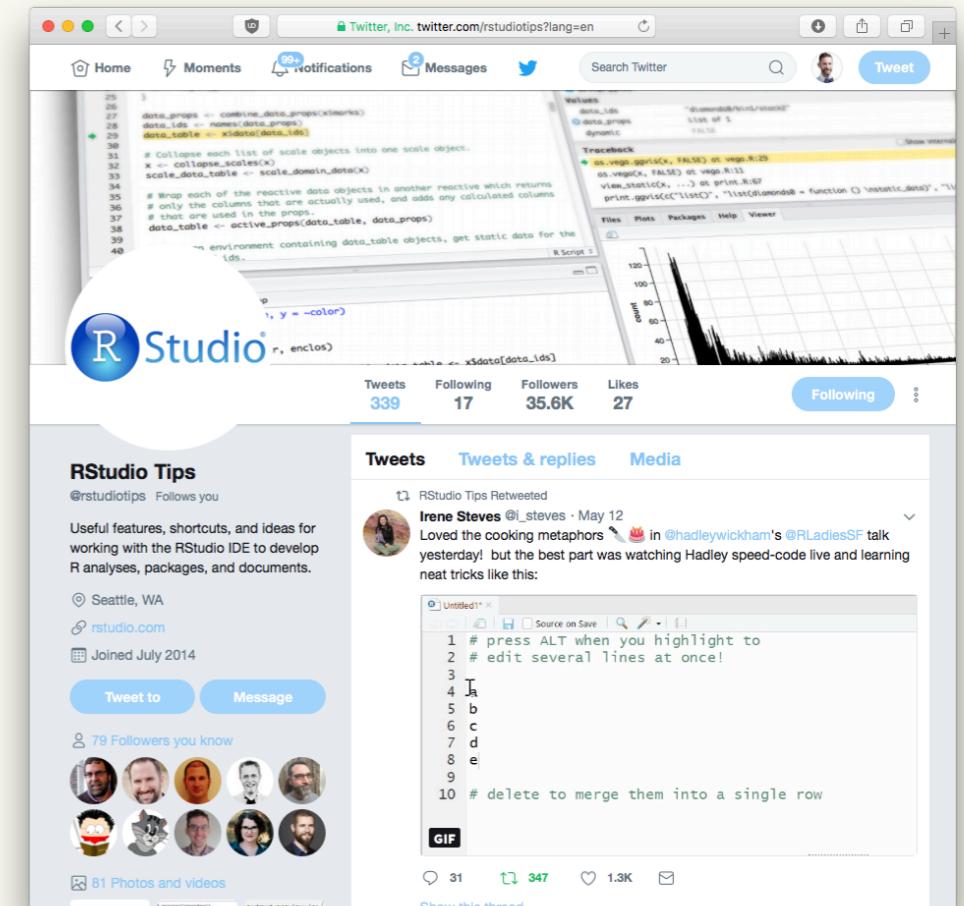
Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Tu turno

¿Cómo escribes <- rápido
(asignar valores)?

¿Y como es %>% (pipe)?

¿Cómo puedes comentar
un bloque de líneas
rápidamente?



@rstudiotips

Materiales: usethis::use_course('ComunidadBioinfo/cdsb2019')

Adapted from *Tidy Tools* by Hadley Wickham

This work is licensed as
Creative Commons
Attribution-ShareAlike 4.0
International

To view a copy of this license, visit
[https://creativecommons.org/
licenses/by-sa/4.0/](https://creativecommons.org/licenses/by-sa/4.0/)