

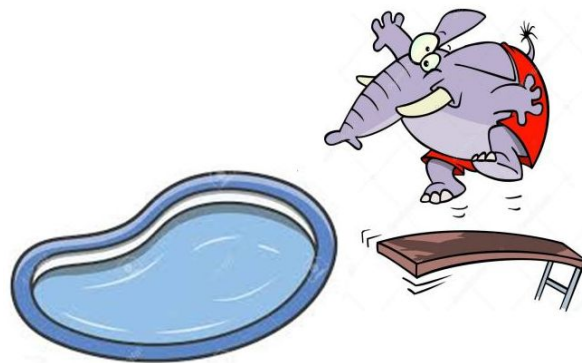
Taller R desde cero con tidyverse



RLadies BRC

10 de Abril 2019

Sofia Ruiz Suarez





¿QUÉ ES?

R es un **lenguaje de programación** orientado a la realización de procesos estadísticos y gráficos.

¿POR QUÉ?

Es **100% gratuito!** y como resultado posee una gran **comunidad**

Es increíblemente **versátil**: Lo podemos usar tanto para el cálculo de estadísticos resumen, para realizar simulaciones complejas o para crear gráficos, mapas o diagramas.

Los códigos resultan fáciles de interpretar y por lo tanto de reproducir y compartir



¿QUÉ ES?

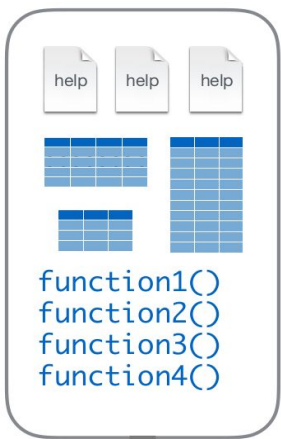
RStudio es un entorno de desarrollo integrado (IDE) para R.

¿POR QUÉ?

Nos permite **combinar y visualizar** nuestro código, gráficos, análisis y escribir texto de manera elegante todo en un mismo lugar.

1. Permite abrir y editar varios archivos a la vez.
2. Muestra el workspace
3. Integra ayuda y documentación
4. Muestra el historial de sentencias ejecutadas

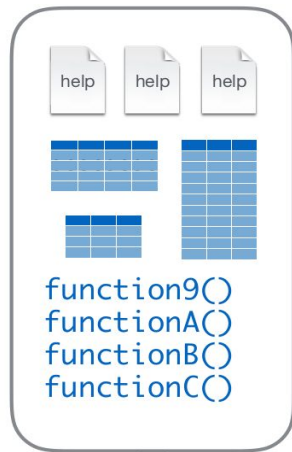
Librerías de R



Cuando instalamos R por primera vez

R Base

```
mean(), sd(), hist() ...
```



Agregar funciones, datos y código escritos por otras personas:

Librerías de R

```
install.packages() library()
```

Bueno pero el tidyverse?...

Tidyverse

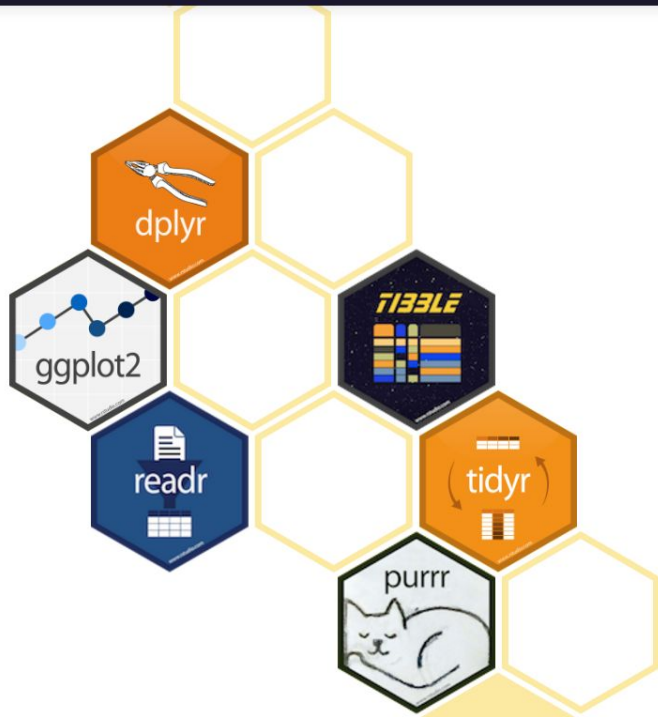
Packages

Articles

Learn

Help

Contribute

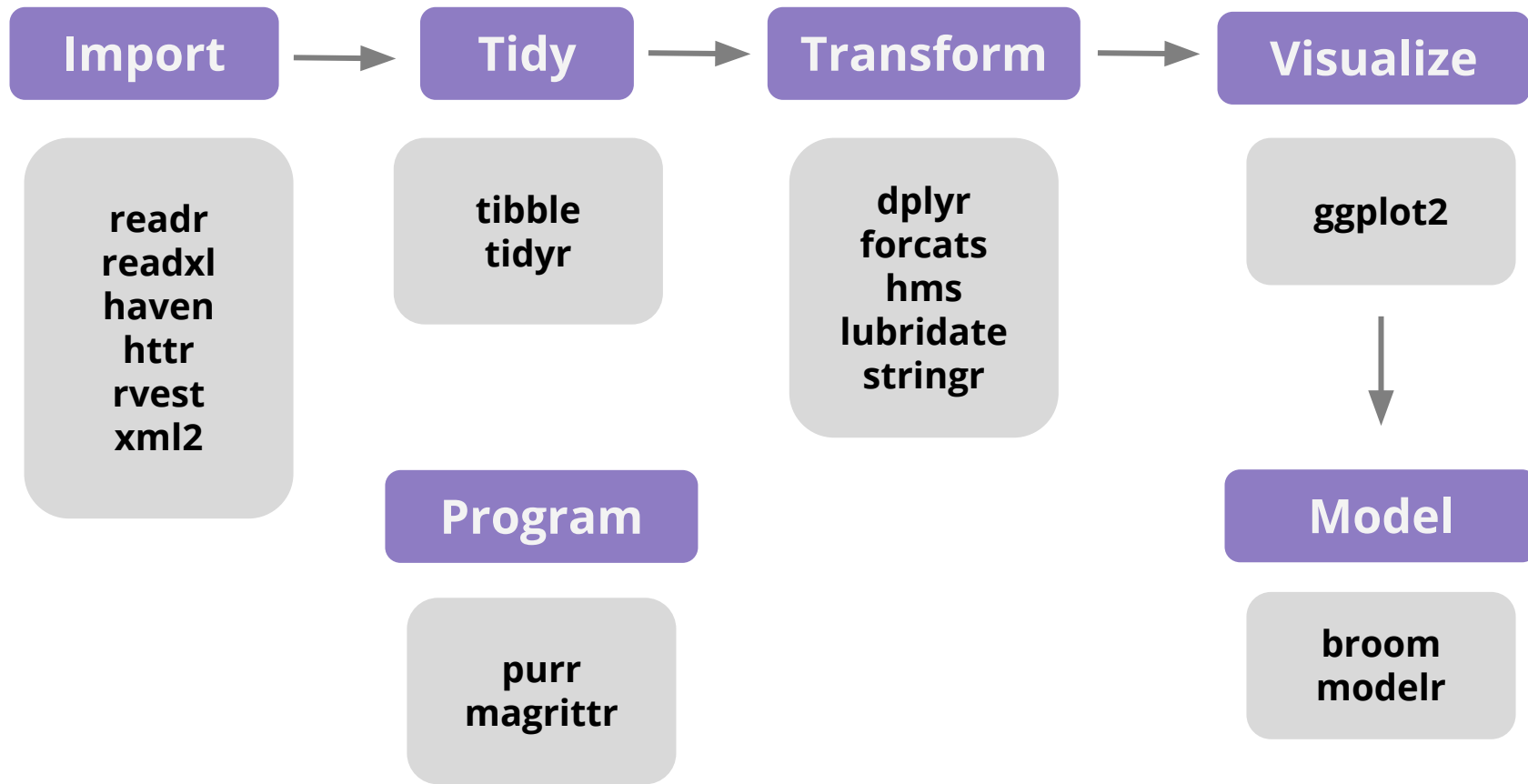


R packages for data science

Es una colección de paquetes de R que comparten una misma filosofía y que están diseñados para usarlos en conjunto

Install the complete tidyverse with:

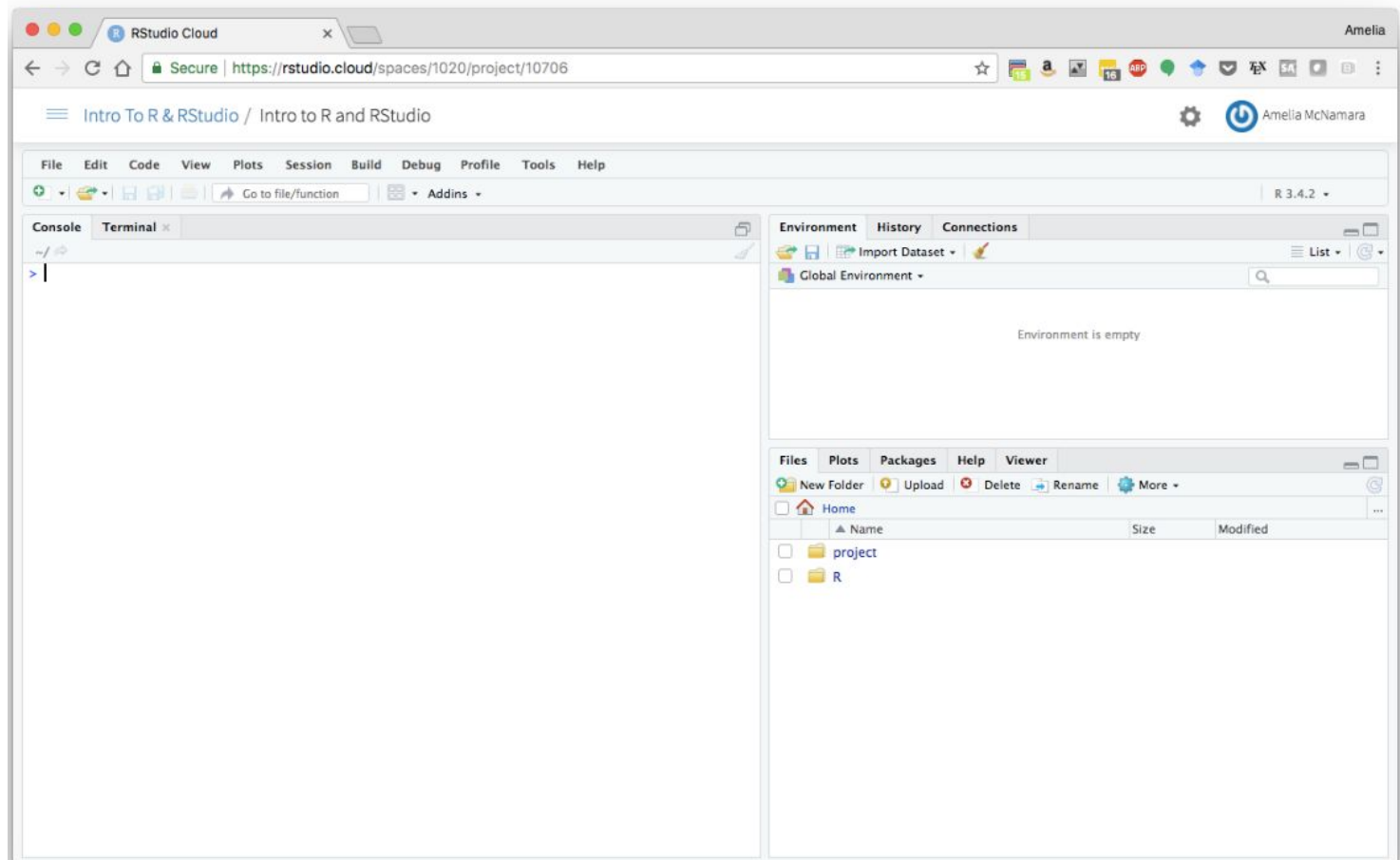
```
install.packages("tidyverse")
```

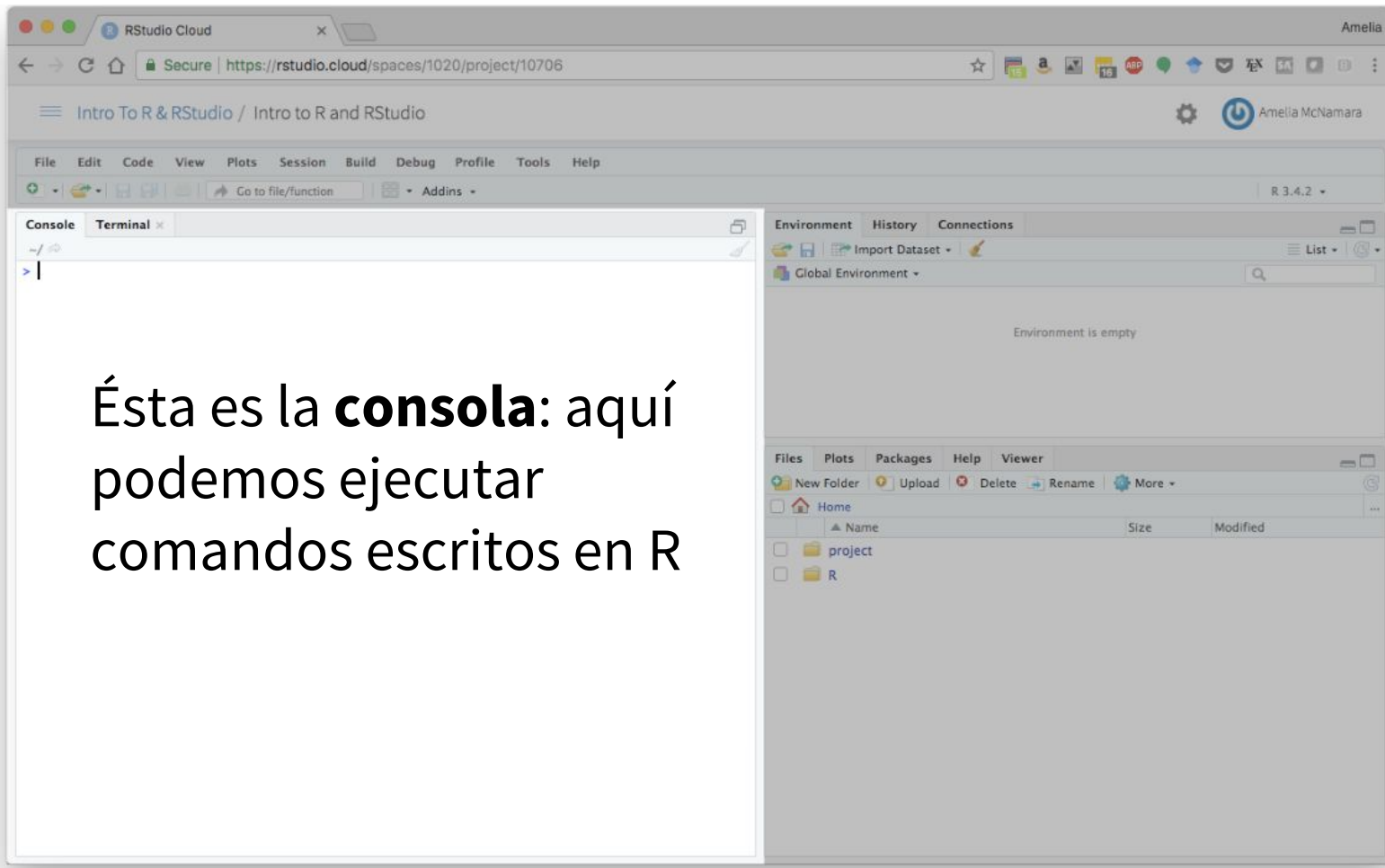


MANOS A LA OBRA !!!!!



Abramos Rstudio!!





Ésta es la **consola**: aquí podemos ejecutar comandos escritos en R

A screenshot of the RStudio Cloud web interface. The browser address bar shows 'https://rstudio.cloud/spaces/1020/project/10706'. The page title is 'Intro To R & RStudio / Intro to R and RStudio'. The user 'Amelia McNamara' is logged in. The interface includes a menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help), a toolbar, a console with a prompt '> |', and a file explorer on the right. The file explorer shows a directory structure with 'Home', 'project', and 'R' folders. A text box is overlaid on the bottom right of the interface.

Carpetas y archivos
del directorio donde
estamos trabajando



RStudio Cloud

Secure | <https://rstudio.cloud/spaces/1020/project/10706>

Intro To R & RStudio / Intro to R and RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Console Terminal

```
> ?ggplot
> |
```

Environment History Connections

Import Dataset

Global Environment

R 3.4.2

Panel de ayuda

Files Plots Packages Help Viewer

R: Create a new ggplot Find in Topic

ggplot {ggplot2} R Documentation

Create a new ggplot

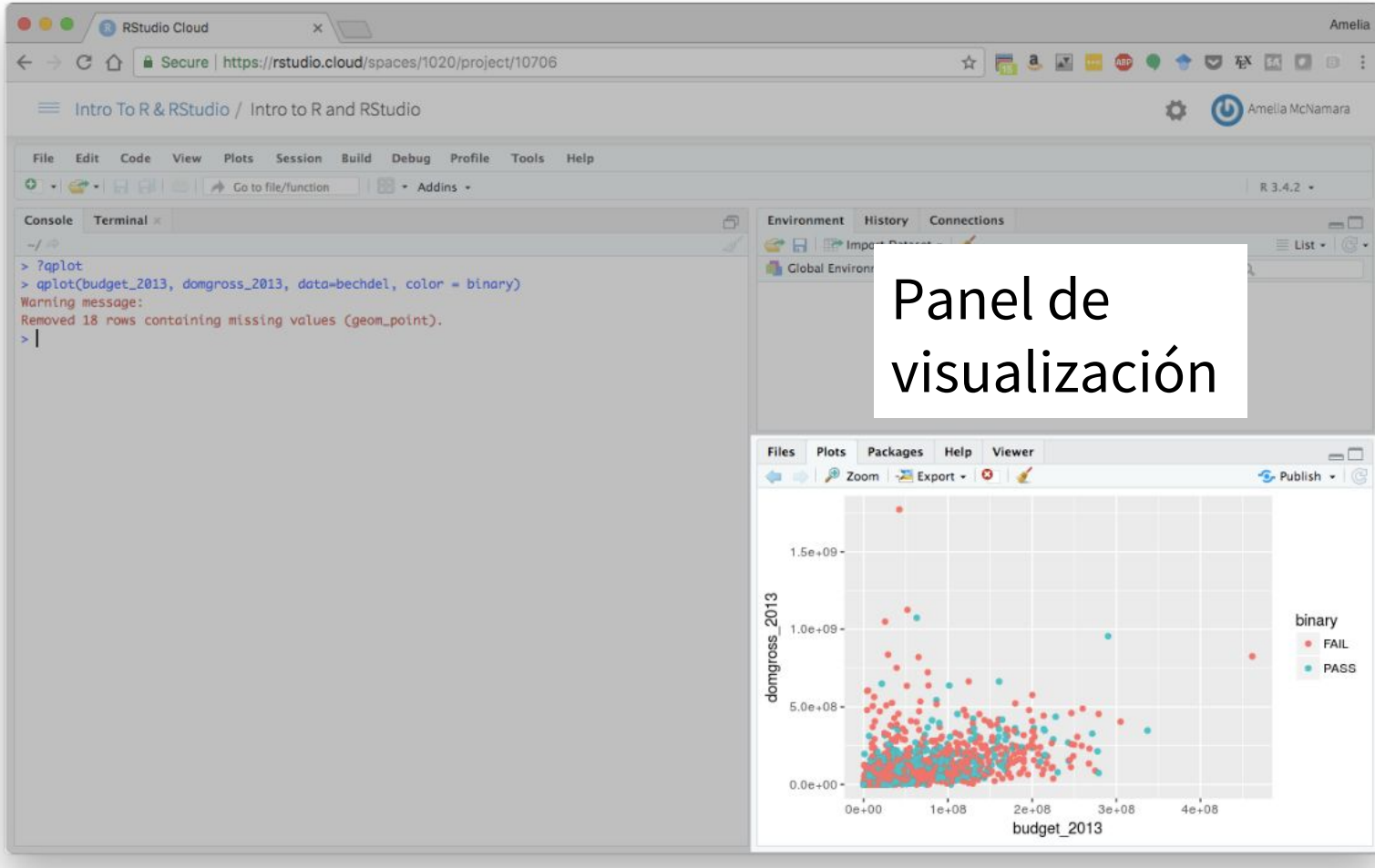
Description

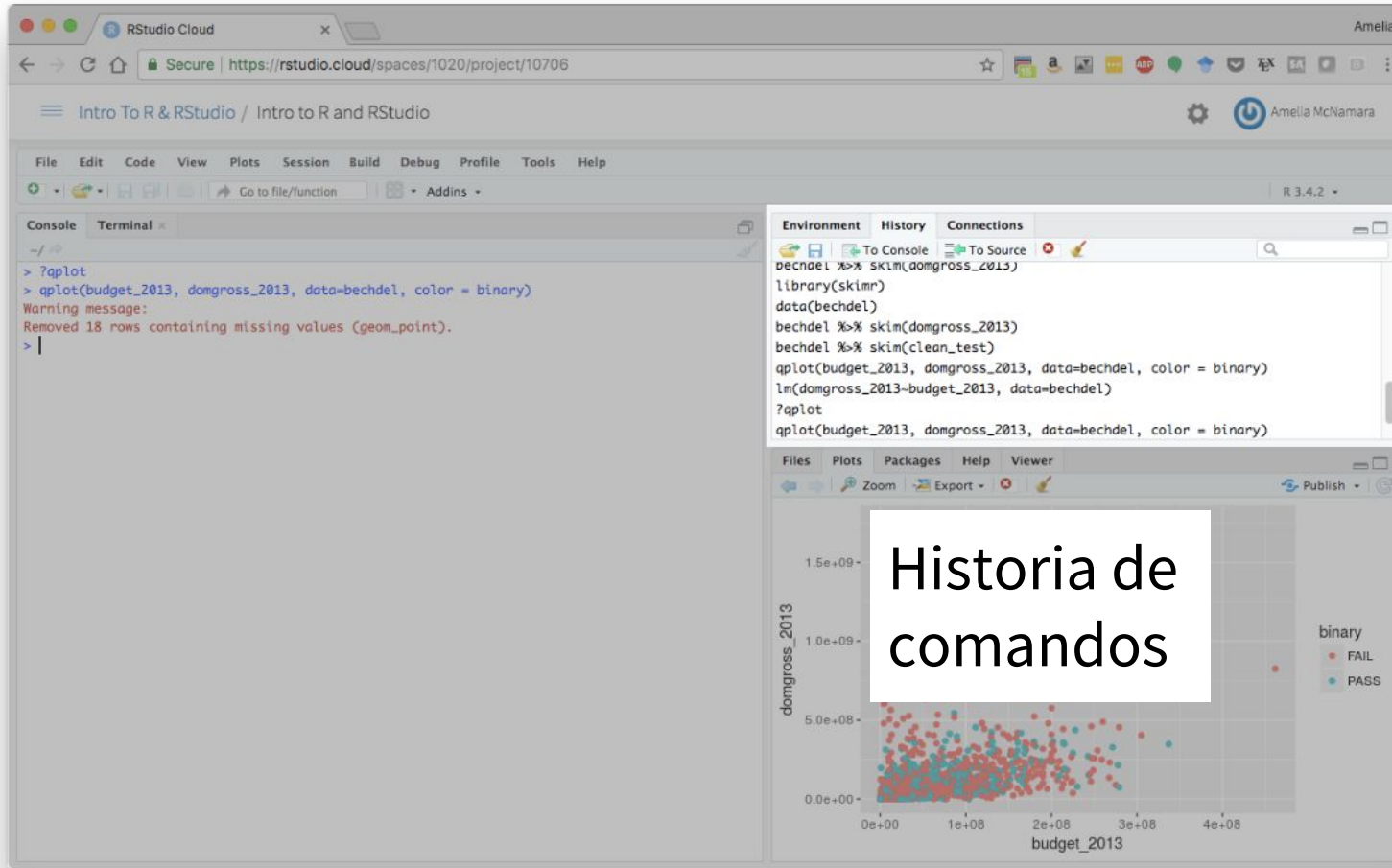
`ggplot()` initializes a ggplot object. It can be used to declare the input data frame for a graphic and to specify the set of plot aesthetics intended to be common throughout all subsequent layers unless specifically overridden.

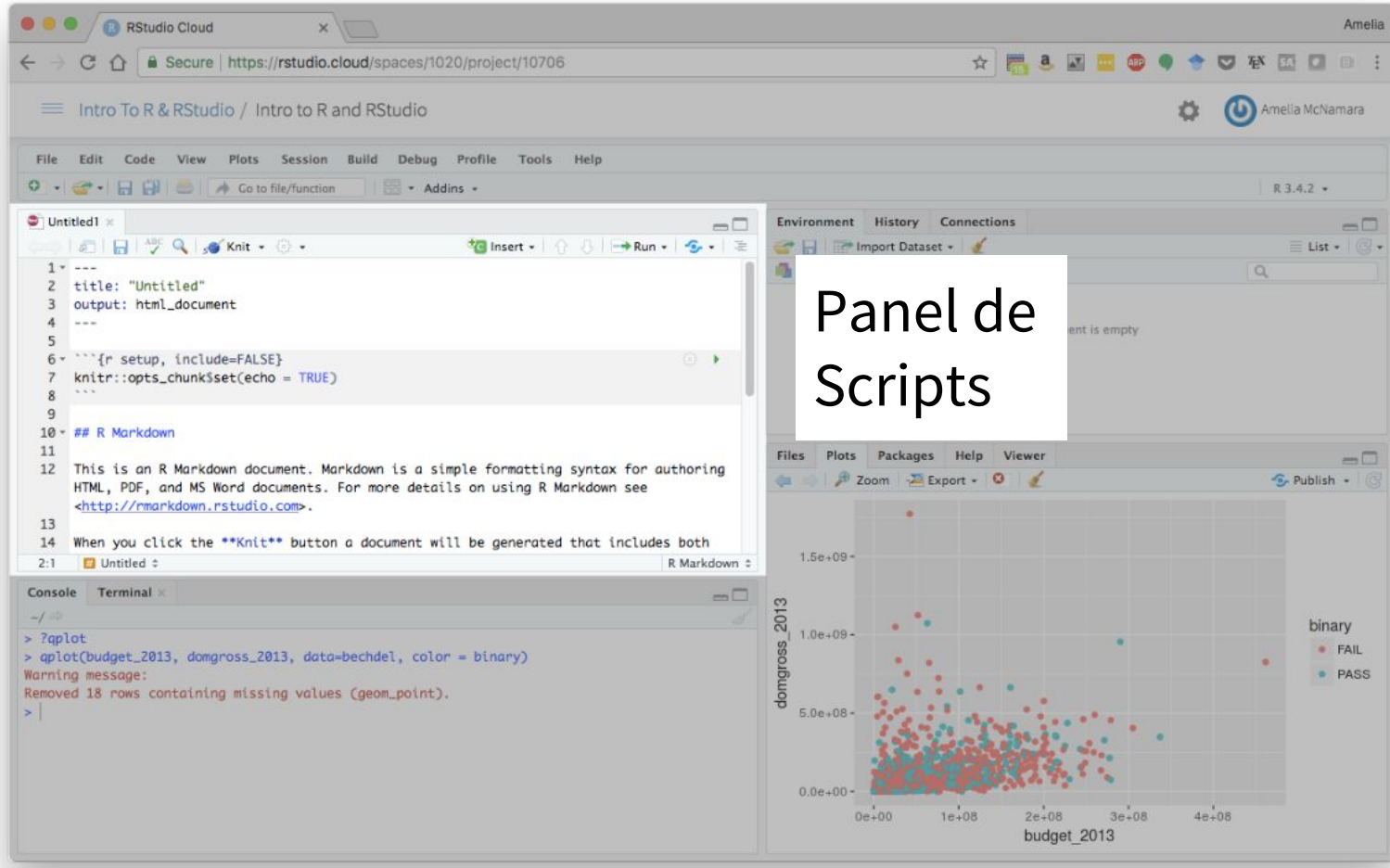
Usage

```
ggplot(data = NULL, mapping = aes(), ..., environment = parent.frame())
```

Arguments





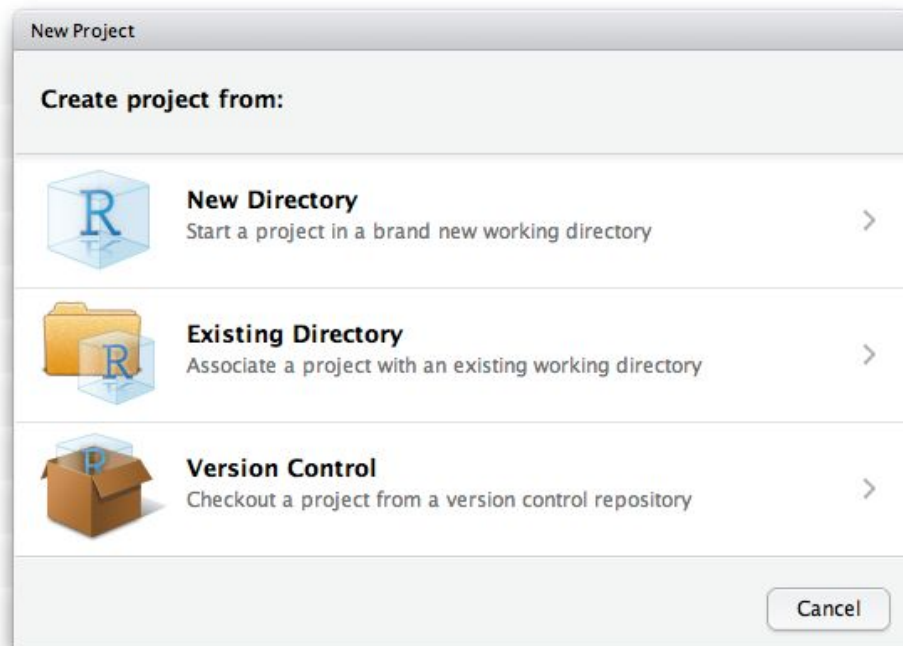


Proyectos de Rstudio

Usar *proyectos de Rstudio* nos permite **dividir nuestro trabajo en múltiples contextos**, cada uno en su propio directorio de trabajo, workspace y documentos fuentes.

Mantiene todos los archivos asociados (datos, scripts, resultados analíticos, figuras)

Es una muy buena práctica acostumbrarse a su uso!



**Veamos cómo lo
hacemos**

VAMOS A RSTUDIO

Cargamos datos

Vamos a usar las funcione del paquete **readr**

Hay varias y se usan de formas distintas.

La más común es la función `read_csv()`

Es más fácil cargarlos directamente usando el Rstudio

Veamos cómo!

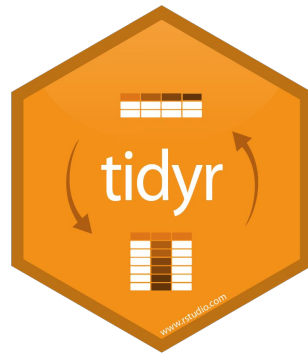


**Veamos cómo lo
hacemos**

VAMOS AL RSTUDIO

Los hacemos “tidy”

Vamos a ver como organizar nuestros datos en R de una manera **simple, sencilla y ordenada** y vamos a ver cómo esto nos ahorra tiempo y nuestros códigos quedan mucho más prolijos y entendibles



country	year	cases	pop
Albania	1999	2745	1929771
Albania	2000	2822	2009721
Albania	2001	2899	2089721
Albania	2002	2976	2169721
Albania	2003	3053	2249721
Albania	2004	3130	2329721

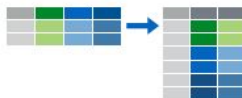
Queremos

1. Cada **variable** en cada **columna**
2. Cada **observación** por cada **fila**
3. Cada **valor** en cada **celda**

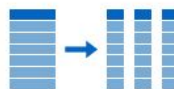
Puede pasar que la base de datos con la que estamos trabajando no esté en su versión “tidy”.



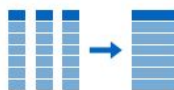
cuando no tenemos una observación por fila: **spread()**



cuando tenemos las observaciones en más de una fila
gather()



cuando tenemos información de dos variables en una
sola columna: **separate()**



cuando tenemos la misma variable en más de una
columna: **unite()**

**Veamos cómo lo
hacemos**

VAMOS AL SCRIPT

AHORA USTEDES!

Ejercicio

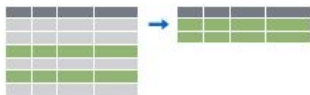
1. Cargar la base de datos “bikedata1.csv” y “bikedata2.csv”.
2. Cuáles son las variables y sus valores?
3. Están estas bases en su versión tidy? Si no lo están , por qué no? y hacerlas tidy



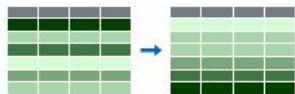
Si tenemos la base de datos organizada de esta manera hay varias funciones de **dplyr** que nos van a ayudar a hacer nuestros análisis de manera sencilla y con pocas líneas de código



Extraer variables con **select()**



Extraer casos con **filter()**



Cambiar el orden de las filas con **arrange()**

Reducir múltiples valores en un valor resumen con **summarize()**

Agregar variables nuevas que sean función de las anteriores con **mutate()**

**Veamos cómo lo
hacemos**

VAMOS AL SCRIPT

AHORA USTEDES!

Ejercicio

1. A partir de los datos “bikedata” tomar aquellos casos del puente Tilikum cuyos casos de accidentes sean mayores a 450.
2. Calcular el promedio y el desvío de la cantidad de casos
3. Calcular el mínimo y el máximo de los casos sobre cada uno de los puentes
4. Modificar el valor de la cantidad de casos que corresponden a los viernes en el puente de Haewthorne de 876 a 935.
5. Crear una nueva variable que sea 2 veces la cantidad de casos y agregarla al data.frame con el nombre “dos.casos”



Libros altamente recomendados

