

Introducción a R y Rstudio

¿Qué vamos a ver hoy?

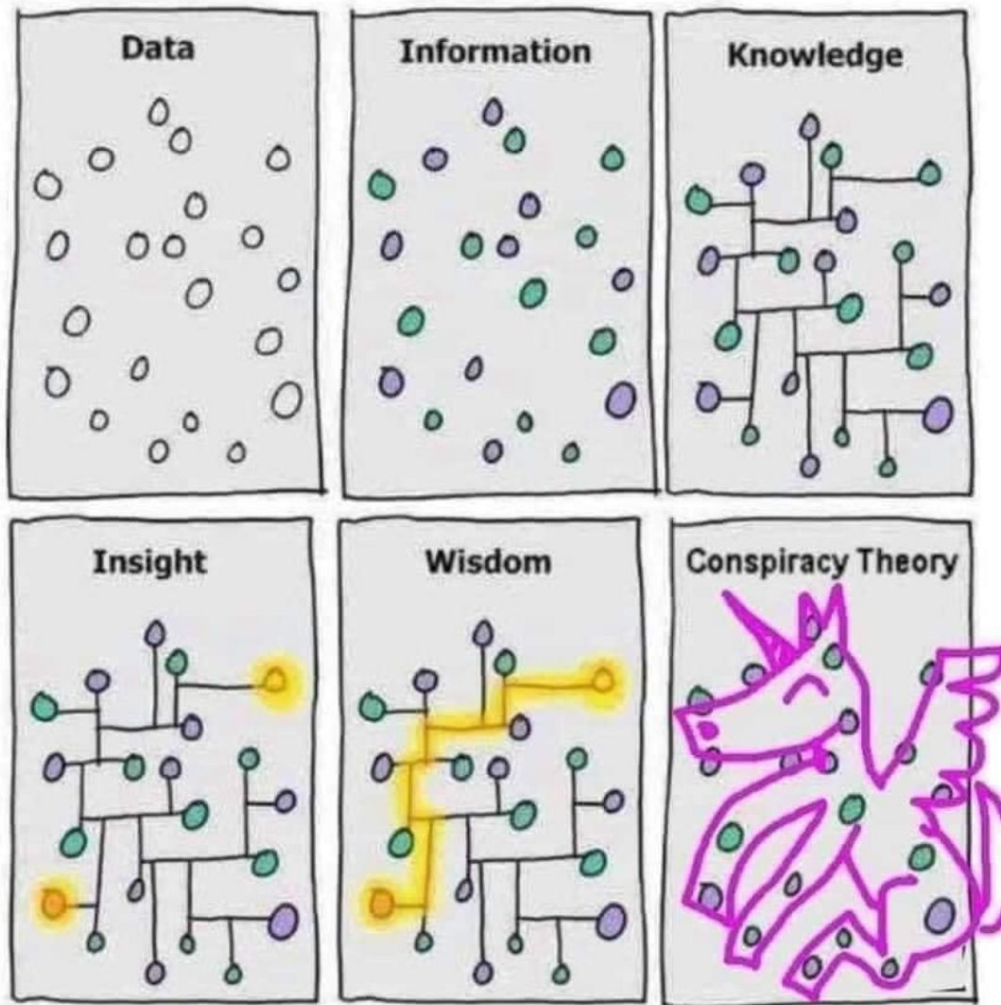
- Descarga de R y Rstudio
- Introducción a R y Rstudio
- Solapas de R
- Rscript
- Aritmética y objetos en R
- Asignación de variables



NO ES TODO

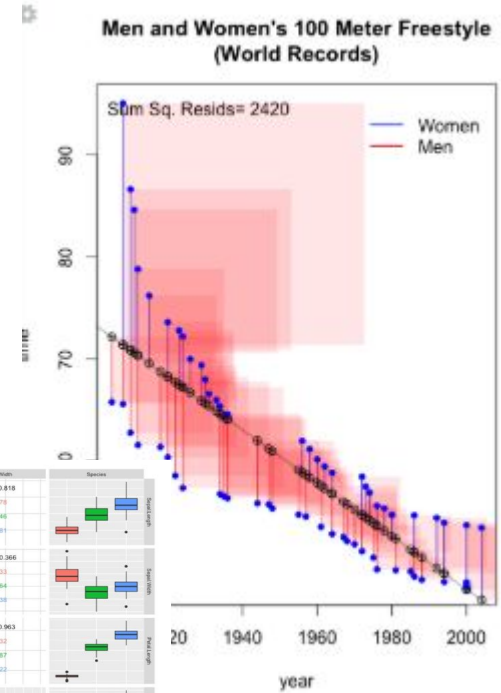
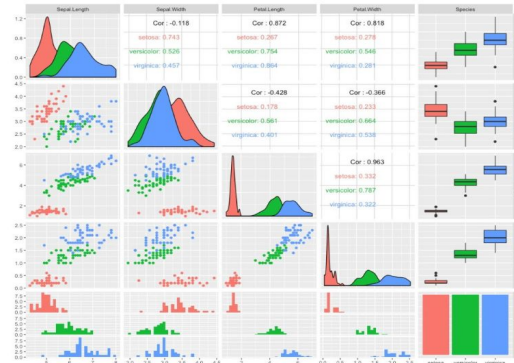
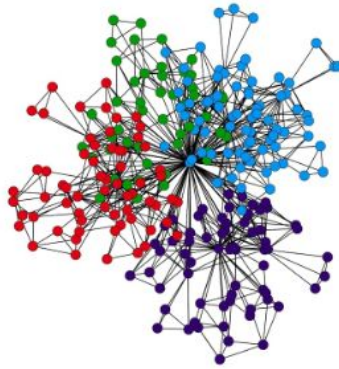


PERO COMO AYUDA



¿Qué es ?

*R es un software libre especializado en **análisis estadístico** y **visualización de datos**.*

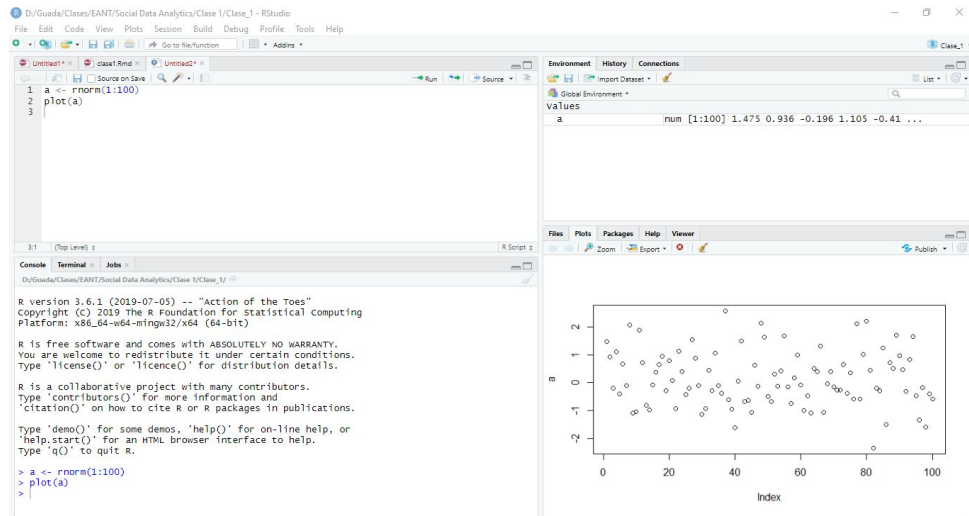
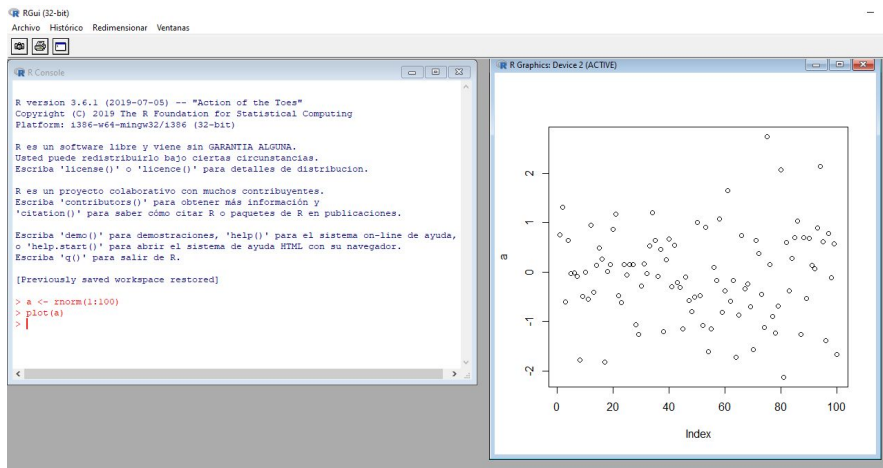


¿Qué es Studio®?

Es un *entorno de desarrollo integrado* (IDE), una **interfaz** para facilitarnos la programación en R.

Tiene ciertas ventajas en su uso:

1. Autocompletado
2. Resaltado de sintaxis
3. Documentación y ayuda integrada
4. Posibilidad de realizar informes con RMarkdown y de trabajo colaborativo



Abramos Rstudio!

RStudio interface showing the Editor de texto, Environment, Console, and Outputs panels.

Editor de texto

```
1 a <- rnorm(1:100)
2 plot(a)
3
```

Environment

Global Environment

Values

a	num [1:100] 1.475 0.936 -0.196 1.105 -0.41 ...

Console

R version 3.6.1 (2019-07-05) -- "Action of the Toes"
Copyright (c) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

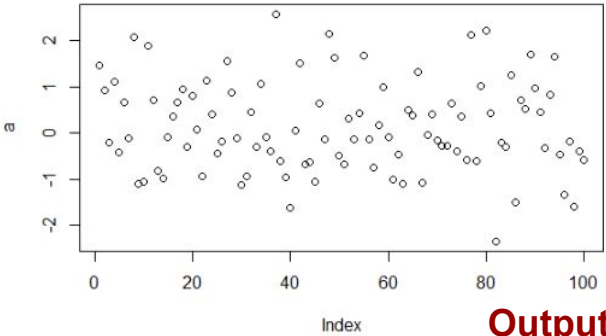
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

```
> a <- rnorm(1:100)
> plot(a)
```

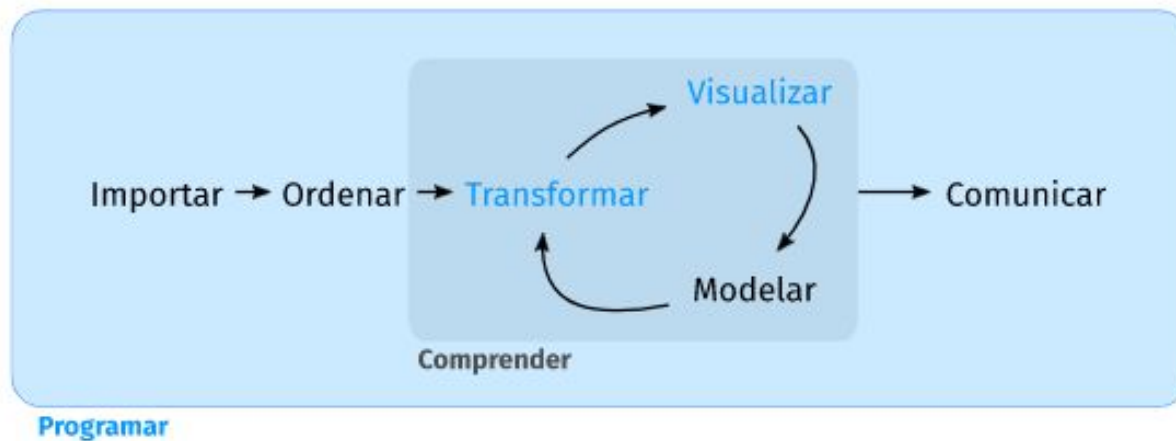
Outputs



Interfaz de Rstudio

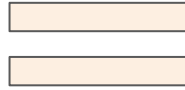
- **Editor de texto (Source):** En este panel vas a escribir/editar/ver los R script y los datasets.
- **Consola/Terminal:** Acá es donde se ejecutan los comandos redactados en el editor (Editor).
- **Environment:** En este panel Rstudio te muestra qué datasets y qué objetos (variables) tenes creadas en la memoria. La solapa *"History"* contiene el historial de los comandos ejecutados en R.
- **Outputs:** Este es un panel multiproposito que devuelve información solicitada. En *"Files"* vos podés navegar en tu escritorio y seleccionar archivos que deseés. En *"Plots"* te va a mostrar los gráficos que realices, y si son interactivos van a aparecer en *"Viewer"*. En *"Packages"* puedes ver cuáles son los paquetes instalados o disponibles. Por último en *"Help"* vas a poder consultar dudas sobre funciones o paquetes.

¿Cómo es el proceso de análisis?



Fuente: [R4DS en español](#)

Script



Código



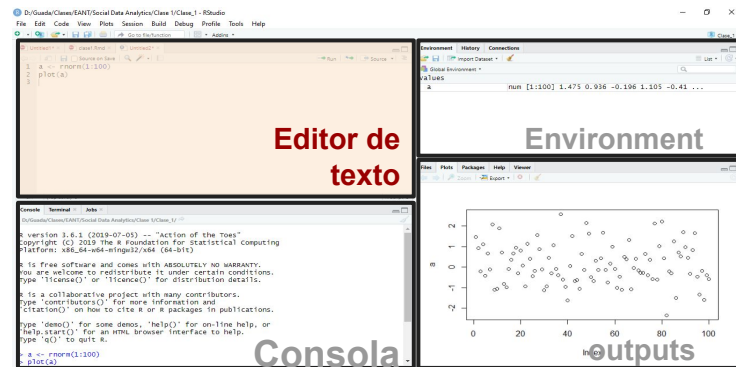
*Conjunto de
instrucciones*

R script

Formato para guardar y editar nuestros códigos al mismo tiempo que podemos ejecutarlos (correrlos)

Ventajas de utilización de Scripts:

- Ordena las instrucciones utilizadas
- Permite guardar los diferentes comandos
- En caso de algún cambio se puede editar una parte del código
- Permite replicar el código a utilizar



Script

*Formato para guardar y editar nuestros códigos al mismo tiempo que podemos ejecutarlos (correrlos) apretando **"Ctrl+Enter"***

- **Para ejecutar parte del código:**

1. Presionar **"Ctrl + Enter"**
2. En caso estar seleccionando parte del código, se va a ejecutar la selección, sino, la línea del script → todo en la consola

- **Para ejecutar todo del código:**

1. Apretamos **"Ctrl + Shift + Enter"** (esto ejecutará el código fuera de la consola y sólo va a mostrar los resultados que explícitamente se impriman)

R como calculadora: operaciones

Aritméticas

- **Suma:** $2+1$
- **Resta:** $3-2$
- **Producto:** $7*8$
- **División:** $7/3$
- **Potencia:** 2^8
- **División Entera:** $7\%/\%2$
- **Resto:** $7\%\%2$

Comparación

- **Menor que:** $a < b$
- **Menor o igual que:** $a \leq b$
- **Mayor que:** $a > b$
- **Mayor o igual que:** $a \geq b$
- **Igual:** $a == b$
- **Diferente:** $a != b$

El resultado de la comparación es FALSE o TRUE

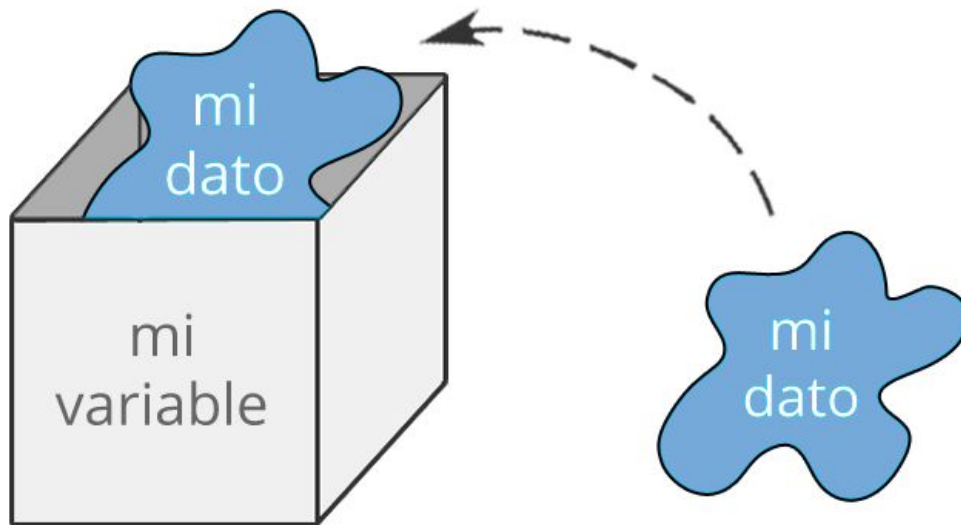
Variables

Son ***espacios reservados en la memoria*** que, como su nombre indica, pueden cambiar de contenido a lo largo de la ejecución de un programa

Pueden ser

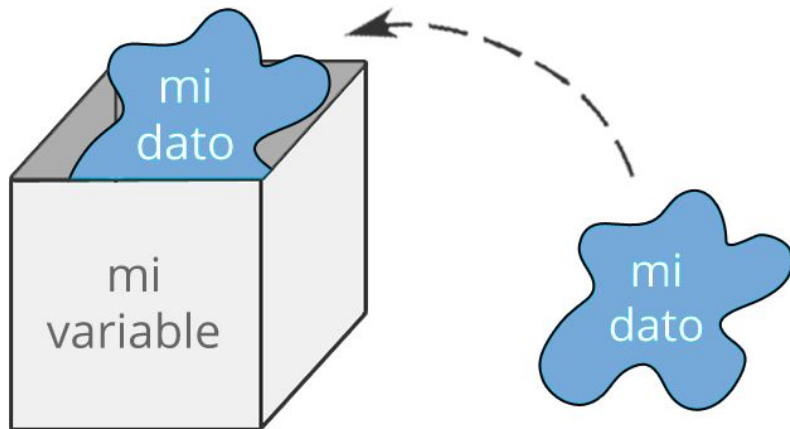
- Números
- Cadenas de texto
- Elementos complejos
- Funciones

Asignación de variables



```
##{r}  
#mi variable "x" tiene el valor 1  
x <- 1  
x  
##
```

Asignación de variables



```
```{r}
#mi variable "x" tiene el valor 1
x <- 1
x
```
```

1

nombre <- objeto

2

nombre = objeto

3

objeto -> nombre

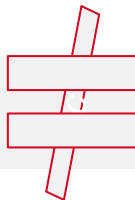
Atención! I



1

R es case sensitive (distingue mayúsculas y minúsculas)

HOLA



hola

Atención! II



2

Si volvemos a asignar a una variable, esta nueva pisa y elimina la anterior

Documentar bien el código:

1. Realizá comentarios (#)

```
#esto es un comentario
```

2. Generá variables con nombres descriptivos:

```
V1 <- 2+1
```

```
Suma <- 2+1
```

Documentar bien el código:

1. Realizá comentarios (#)

```
#esto es un comentario
```

2. Generá variables con nombres descriptivos:

```
V1 <- 2+1 ❌
```

```
Suma <- 2+1 ✅
```

Tipo de variables

| Tipo | Ejemplo |
|---------------------------------|---------------------------------|
| numeric | 3.14 |
| character | "Hola, cómo estas?" |
| factor | "Femenino", "Masculino", "Otr@" |
| logical | TRUE / FALSE |
| missing o no disponibles | NA |
| nulos | Valores indefinidos |

Si quieres saber más clickea [acá...](#)

Operadores con R

| Lógico | Descripción |
|--------|-------------|
| & | AND |
| | OR |
| ! | NOT |

| Condicional | Descripción |
|-------------|---------------|
| == | IGUAL |
| != | DIFERENTE |
| > | MAYOR |
| < | MENOR |
| >= | MAYOR O IGUAL |
| <= | MENOR O IGUAL |

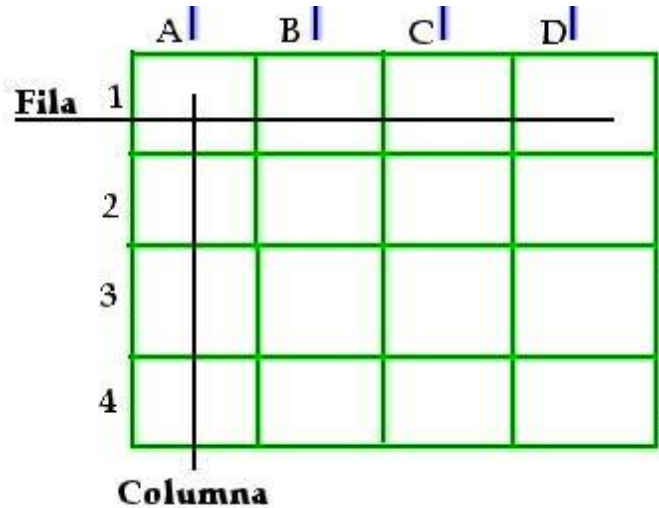
Vectores

Serie de elementos de un mismo tipo



Data Frame

Estructuras tabulares de dos dimensiones: filas y columnas



A diagram illustrating a two-dimensional tabular structure (Data Frame) with 4 rows and 4 columns. The grid is defined by green lines. The columns are labeled A, B, C, and D at the top, with a vertical line pointing to the first column labeled 'Columna'. The rows are labeled 1, 2, 3, and 4 on the left, with a horizontal line pointing to the first row labeled 'Fila'.

| | A | B | C | D |
|---|---|---|---|---|
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |
| 4 | | | | |

¿Qué es un algoritmo?

Es un conjunto de instrucciones o reglas definidas y no-ambiguas, ordenadas y finitas que permite, típicamente, solucionar un problema.

Pastel
1- 250 g de harina
1/2 cucharadita Royal
100 g manteca

2- 100 g azúcar
1 huevo
6 cucharadas leche

Relleno
1/2 taza maizena
1 taza azúcar
corteza rallada 1 limón
1/2 taza zumo limón
4 yemas de huevo
1/8 cucharadita sal
1 cucharada manteca
1 1/2 taza agua

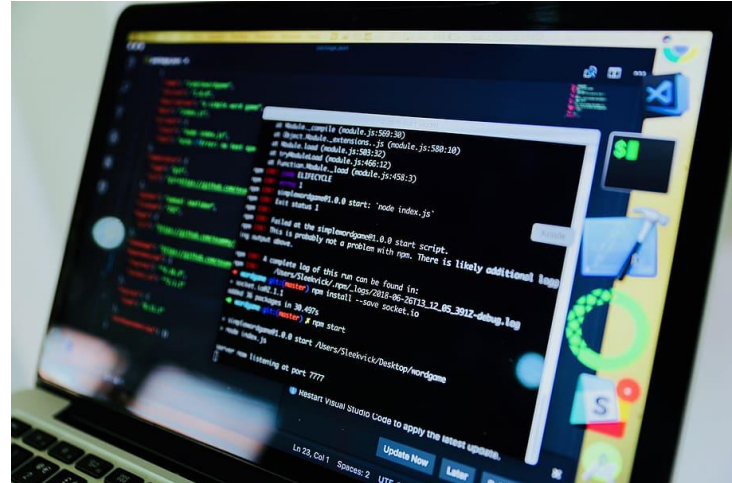
Cubierta
4 claras
1/2 taza azúcar

Le
e
m
o
n

Masa:
* Hacer un hueco en 1 y agregar 2.
* Mezclar bien y hacer un bollo; ponerlo 15' en la nevera.
* Hornear 15' en horno moderado en un molde enmantecado y enharinado.

Resto:
* Cocer el relleno a fuego lento.
* Colocar sobre el pastel horneado.
* Agregar las claras a punto de nieve.
* Dar un golpecito de horno para que se dore el merengue.

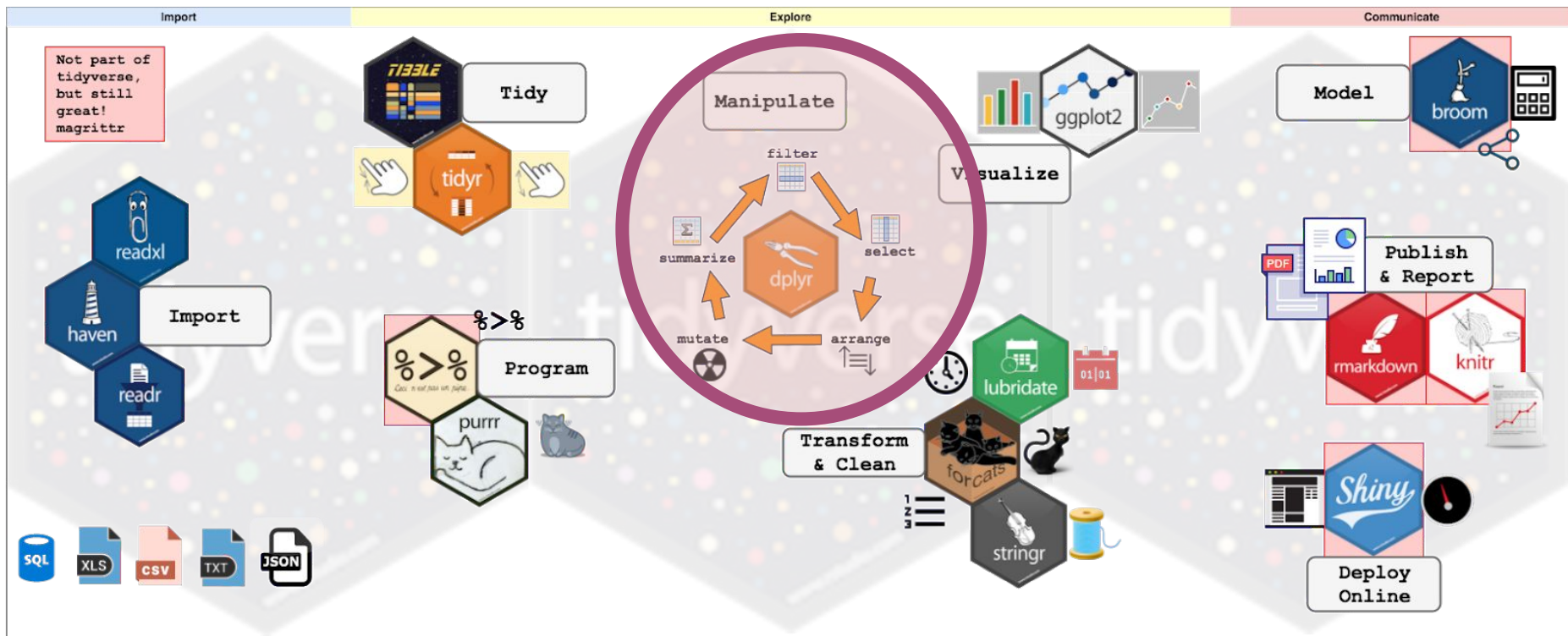
• **pie**

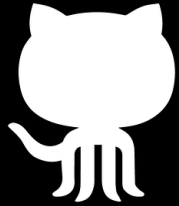


COVID-19

Coronavirus Disease 2019

**¿Cómo llegaron los sistemas sanitarios
del país a la pandemia?**

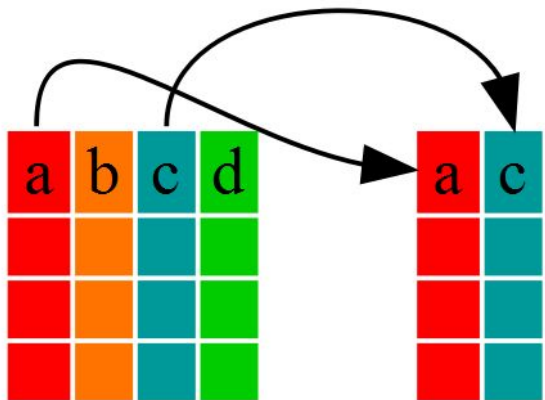




<https://github.com/labpoliticassuba>

Select

```
select(data.frame,a,c)
```



Filter

```
filter(dataframe , [logical conditions])
```

↑
The dataframe
that you want
to filter

↑
Logical conditions
specifying the rows
you want to return

df

| color | value |
|-------|-------|
| blue | 1 |
| black | 2 |
| blue | 3 |
| blue | 4 |
| black | 5 |

→

| color | value |
|-------|-------|
| blue | 1 |
| blue | 4 |

```
filter(df, value %in% c(1, 4))
```


Arrange

df

| color | value |
|-------|-------|
| 4 | 1 |
| 1 | 2 |
| 5 | 3 |
| 3 | 4 |
| 2 | 5 |

→

| color | value |
|-------|-------|
| 5 | 3 |
| 4 | 1 |
| 3 | 4 |
| 2 | 5 |
| 1 | 2 |

```
arrange(df, desc(color))
```

Mutate

mutate()

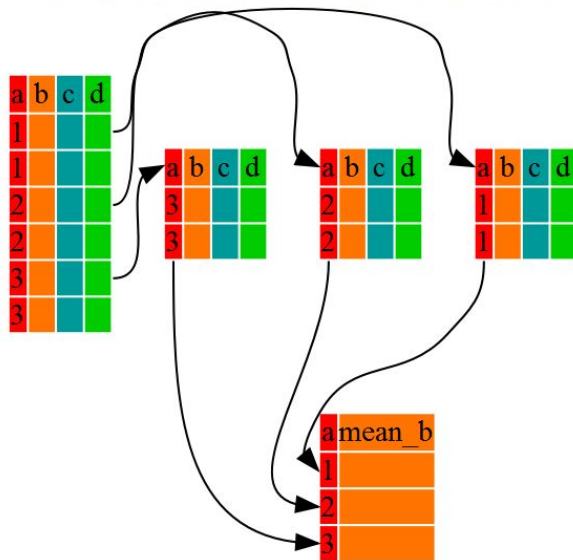
| storm | wind | pressure | date |
|---------|------|----------|------------|
| Alberto | 110 | 1007 | 2000-08-12 |
| Alex | 45 | 1009 | 1998-07-30 |
| Allison | 65 | 1005 | 1995-06-04 |
| Ana | 40 | 1013 | 1997-07-01 |
| Arlene | 50 | 1010 | 1999-06-13 |
| Arthur | 45 | 1010 | 1996-06-21 |



| storm | wind | pressure | date | ratio |
|---------|------|----------|------------|-------|
| Alberto | 110 | 1007 | 2000-08-12 | 9.15 |
| Alex | 45 | 1009 | 1998-07-30 | 22.42 |
| Allison | 65 | 1005 | 1995-06-04 | 15.46 |
| Ana | 40 | 1013 | 1997-07-01 | 25.32 |
| Arlene | 50 | 1010 | 1999-06-13 | 20.20 |
| Arthur | 45 | 1010 | 1996-06-21 | 22.44 |

Group by y summarise

```
gapminder %>% group_by(a) %>% summarize(mean_b=mean(b))
```



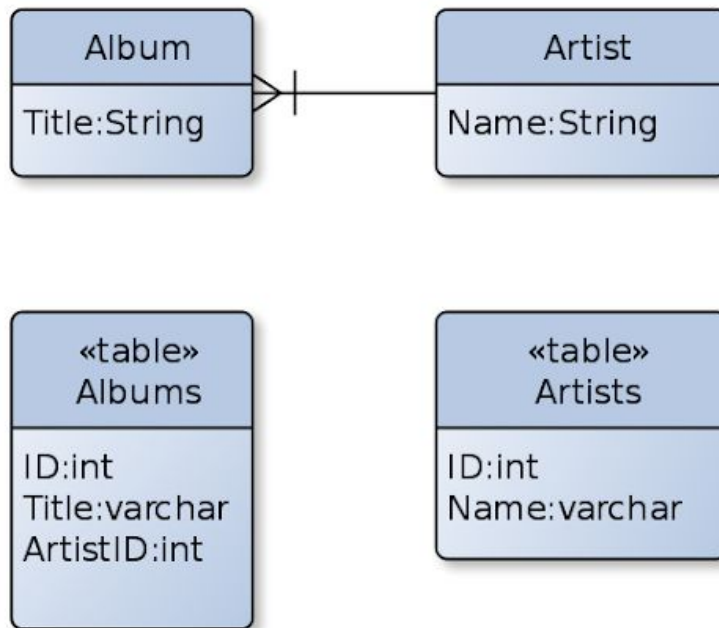
Pipe

`%>%`

Es un ***operador*** para concatenar funciones en tidyverse.

```
yo %>%  
  me_despierto() %>%  
  salgo_de_la_cama() %>%  
  me_visto() %>%  
  preparo_el_mate() %>%  
  me_voy_a_trabajar()
```

Primary Key y Foreign Key



Joins

Combine Data Sets

| a | | b | |
|----|----|----|----|
| x1 | x2 | x1 | x3 |
| A | 1 | A | T |
| B | 2 | B | F |
| C | 3 | D | T |

+

=

Mutating Joins

| x1 | x2 | x3 |
|----|----|----|
| A | 1 | T |
| B | 2 | F |
| C | 3 | NA |

dplyr::left_join(a, b, by = "x1")

Join matching rows from b to a.

| x1 | x3 | x2 |
|----|----|----|
| A | T | 1 |
| B | F | 2 |
| D | T | NA |

dplyr::right_join(a, b, by = "x1")

Join matching rows from a to b.

| x1 | x2 | x3 |
|----|----|----|
| A | 1 | T |
| B | 2 | F |

dplyr::inner_join(a, b, by = "x1")

Join data. Retain only rows in both sets.

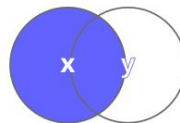
| x1 | x2 | x3 |
|----|----|----|
| A | 1 | T |
| B | 2 | F |
| C | 3 | NA |
| D | NA | T |

dplyr::full_join(a, b, by = "x1")

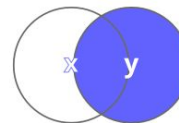
Join data. Retain all values, all rows.

dplyr joins

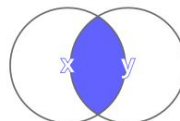
left_join(x, y)



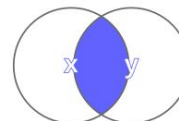
right_join(x, y)



inner_join(x, y)

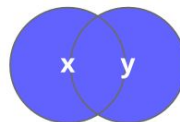


semi_join(x, y)

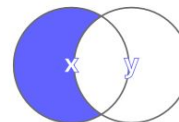


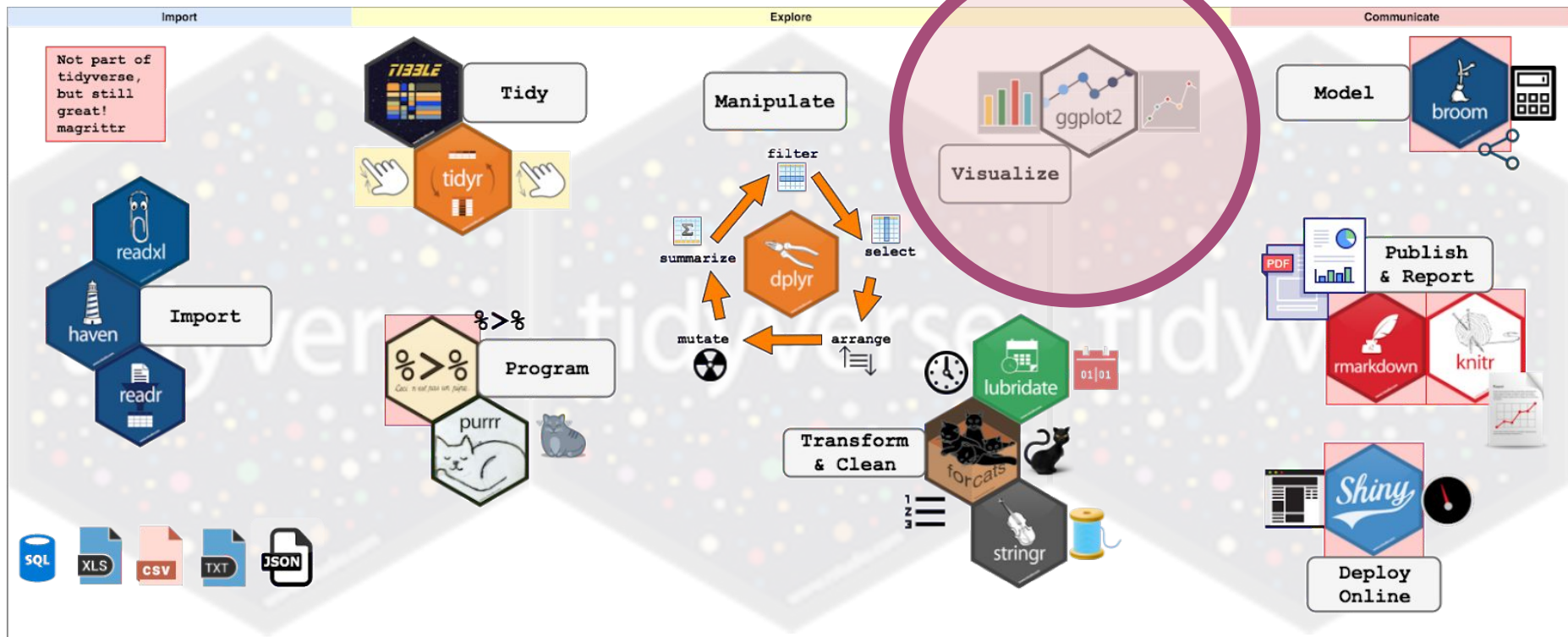
(never duplicate rows of x)

full_join(x, y)

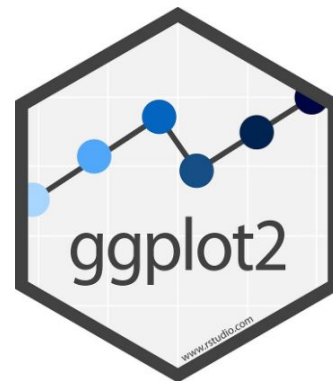


anti_join(x, y)

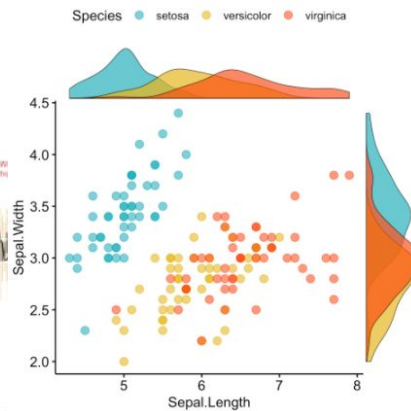
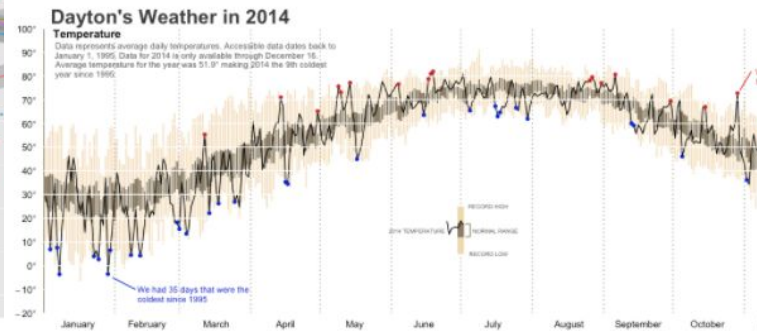
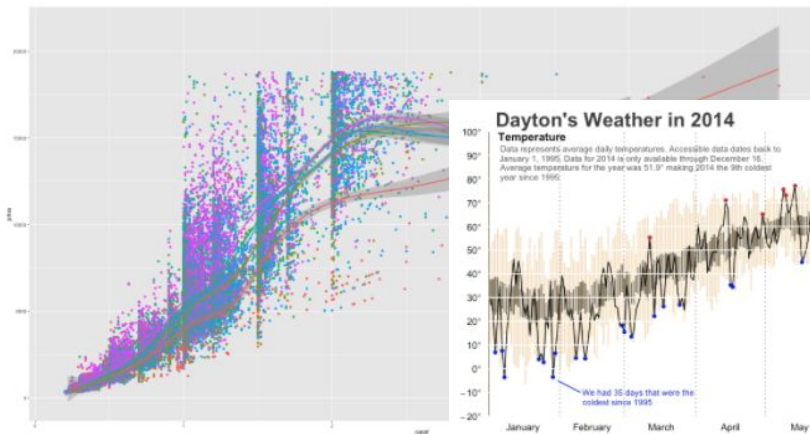




“ggplot2”



Es una librería que nos permite generar gráficos en R con una diversidad de **posibilidades**.





<https://www.data-to-viz.com>



¿Qué vimos hoy?

- Introducción a R y Rstudio
- Solapas de R
- Rscript
- Aritmética y objetos en R
- Asignación de variables
- Tipos de objetos y operadores
- Vectores, dataframes
- Algoritmos
- Tidyverse:
 - select
 - filter
 - mutate
 - group by
 - summarise
 - arrange
- Ggplot2