

# Herramientas Cuantitativas para el Análisis Político

[CP44] Maestría en Ciencia Política

Juan Pablo Ruiz Nicolini

Universidad Torcuato Di Tella

22/09/2020

# SESIÓN 2

## Introducción: Conociendo R + RStudio

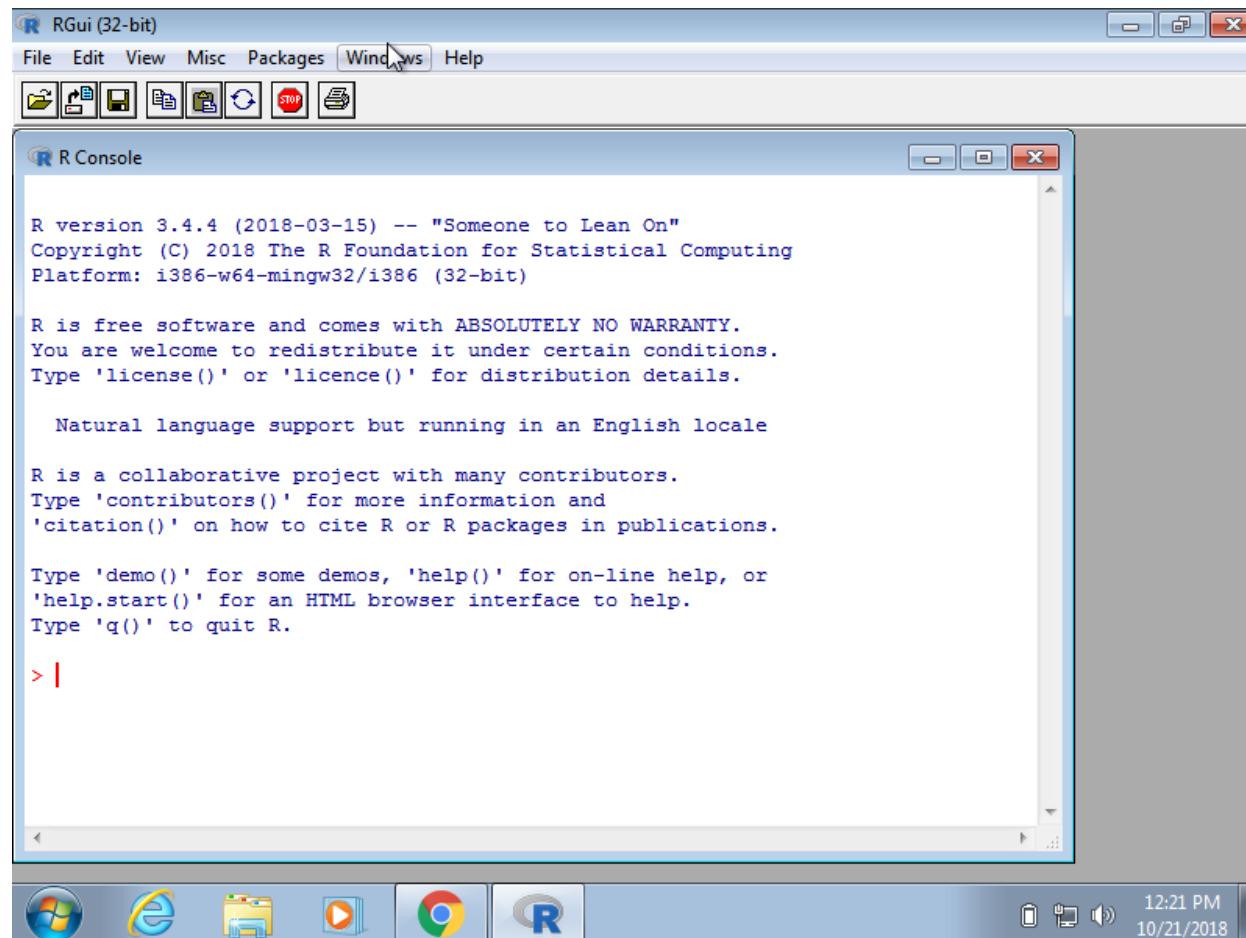
/MetodosCiPol/

/MetodosCiPol/

# Referencias

1. Comenzando con R y RStudio, en "Introducción a la Ciencia de Datos" (Irizarry, 2020).
2. Prerequisitos, en "R para Ciencia de Datos" (Wickham y Grolemund).
3. R Básico en Urdinez y Cruz; Conociendo RStudio en Montané; y Una presentación a toda marcha en Vázquez Brust.

# *El motor de la maquina*



# Tablero de control (IDE)



The image shows a screenshot of the RStudio IDE interface. It consists of four main panels:

- Script Editor (Top Left):** Shows a script named "Untitled1" with the following code:

```
1 # Este es un script de prueba
2
3 objeto_1 <- c(1:3)
4
```

A large black rectangular overlay with the word "script" in white is centered in this panel.
- Environment (Top Right):** Shows the global environment with one object defined:

```
Values
objeto_1      int [1:3] 1 2 3
```

A large black rectangular overlay with the word "objetos" in white is centered in this panel.
- Console (Bottom Left):** Shows the R console output:

```
4:1 (Top Level) 
Console Terminal Jobs 
~/ 
Natural language support will run using the English locale.

R is a collaborative project with many contributors.
Type 'contributors()' or 'citation()' for more information.
Type 'demo()' to see examples; 'help.start()' to get help.
Type 'q()' to quit R.

> objeto_1 <- c(1:3)
>
```

A large black rectangular overlay with the word "consola" in white is centered in this panel.
- Help/Viewer (Bottom Right):** Shows the help viewer with the following text:

```
Files Plots Packages Help Viewer 
Home 
archivos / gráficos /
paquetes / ayuda
```

A large black rectangular overlay with the text "archivos / gráficos / paquetes / ayuda" in white is centered in this panel.

# Empecemos *codeando* !



# R base vs *Tidyverse*



```
curso <- data.frame(nombre= c("Juan", "Pedro", "María", "José",
                               "Enzo", "Ariel", "Eva", "Domingo"),
                      edad= c(25, 32, 21, 40,
                             30, 28, 37, 25),
                      nacim= c(1993, 1986, 1997, 1978,
                             1988, 1990, 1981, 1993),
                      software.primario= c("spss", "stata", "stata", "excel",
                                           "R", "stata", "spss", "stata"),
                      nivel= c(3, 5, 7, 6,
                             2, 6, 8, 6)
                     )
```

# R base vs *Tidyverse*



## Edad promedio del curso > BASE

```
# BASE  
sum(curso$edad)/length(curso$edad)  
## [1] 29.75
```



# R base vs *Tidyverse*

## Edad promedio del curso > *Tidy*

```
# Tidy / {dplyr}  
  
library(tidyverse) # cargo paquetes  
  
curso %>%  
  summarise(promedio = sum(edad)/length(edad))  
##   promedio  
## 1    29.75
```

# Juntos por el Cambio 2019:



## Parámetros de la elección?

```
library(polAr)

show_available_elections(viewer = F)
## # A tibble: 425 x 5
##   district category round  year    NOMBRE
##   <chr>     <chr>   <chr> <chr> <chr>
## 1 arg       presi    balota 2015  ARGENTINA
## 2 arg       presi    gral   2003  ARGENTINA
## 3 arg       presi    gral   2007  ARGENTINA
## 4 arg       presi    gral   2011  ARGENTINA
## 5 arg       presi    gral   2015  ARGENTINA
## 6 arg       presi    gral   2019  ARGENTINA
## 7 arg       presi    paso   2011  ARGENTINA
## 8 arg       presi    paso   2015  ARGENTINA
## 9 arg       presi    paso   2019  ARGENTINA
## 10 caba    dip     gral   2005  CABA
## # ... with 415 more rows
```

# Descarga resultados



## Presidente 2019

```
get_election_data(district = 'arg',
                   category = 'presi',
                   round = 'gral',
                   year = 2019)

## # A tibble: 192 x 8
## # Groups:   codprov [24]
##   category round  year codprov name_prov    electores listas     votos
##   <chr>     <chr> <dbl> <chr>   <chr>      <dbl> <chr>      <dbl>
## 1 presi     gral    2019  01     CABA        0 00024  12879
## 2 presi     gral    2019  01     CABA        0 00036  5801
## 3 presi     gral    2019  01     CABA        0 00037  7071
## 4 presi     gral    2019  01     CABA        0 00039  1351
## 5 presi     gral    2019  01     CABA        0 00050  10397
## 6 presi     gral    2019  01     CABA        0 00108  3740
## 7 presi     gral    2019  01     CABA        0 blancos  305
## 8 presi     gral    2019  01     CABA        0 nulos    0
## 9 presi     gral    2019  02 BUENOS AIRES  0 00024  6100
## 10 presi    gral    2019  02 BUENOS AIRES  0 00036  2605
## # ... with 182 more rows
```

# Agrega nombre de listas



```
gral2019 <- get_election_data(district = 'arg',
                                category = 'presi',
                                round = 'gral',
                                year = 2019) %>%
  get_names() # AGREGO NOMBRE DE LISTAS AL DF
```

```
gral2019
## # A tibble: 192 x 9
## # Groups: codprov [24]
##   category round year codprov name_prov electores listas votos no
##   <chr>     <chr> <dbl> <chr>    <chr>      <dbl> <chr>  <dbl> <ch
## 1 presi     gral    2019  01     CABA        0 00024  1.29e5 Co
## 2 presi     gral    2019  01     CABA        0 00036  5.80e4 Fi
## 3 presi     gral    2019  01     CABA        0 00037  7.07e5 Fi
## 4 presi     gral    2019  01     CABA        0 00039  1.36e4 Fi
## 5 presi     gral    2019  01     CABA        0 00050  1.04e6 Ju
## 6 presi     gral    2019  01     CABA        0 00108  3.74e4 Ul
## 7 presi     gral    2019  01     CABA        0 blanc~  3.05e4 b
## 8 presi     gral    2019  01     CABA        0 nulos   0.       n
## 9 presi     gral    2019  02     BUENOS AI~  0 00024  6.10e5 Co
## 10 presi    gral    2019  02     BUENOS AI~  0 00036  2.61e5 Fi
## # ... with 182 more rows
```

# Dónde cosechó menos votos Juntos por el Cambio?



Provincia con votos `min()`

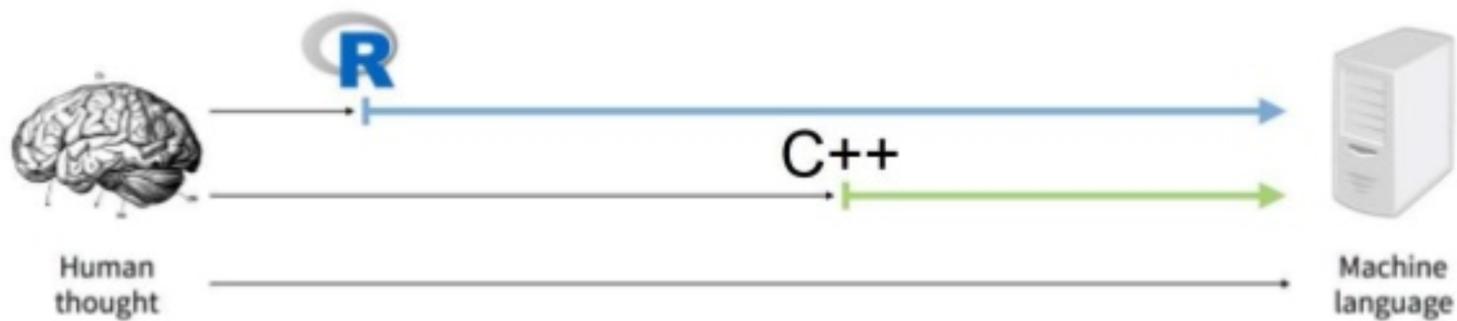
```
jxc <- gral2019[gral2019$nombre_lista == 'JUNTOS POR EL CAMBIO', ] # > l
jxc <- as.data.frame(jxc)

jxc[jxc$votos == min(jxc$votos),]
##   category round year codprov      name_prov electores listas votos no
## 24     presi    gral 2019 24 TIERRA DEL FUEGO       0 00050 256.
##   nombre_lista
## 24 JUNTOS POR EL CAMBIO
```

```
gral2019 %>% # > TIDY
  filter(nombre_lista == 'JUNTOS POR EL CAMBIO') %>%
  ungroup() %>%
  slice_min(votos)
## # A tibble: 1 x 9
##   category round year codprov name_prov   electores listas votos no
##   <chr>     <chr> <dbl> <chr>    <chr>       <dbl> <chr>  <dbl> <ci
## 1 presi     gral    2019 24    TIERRA DEL~       0 00050 25673 JU
```



## A computer language for scientists



# Tidyverse



1. Reutilizar estructuras de datos existentes
2. Armar funciones simples con la *pipa* (`%<%`)
3. Adoptar la programación funcional
4. Diseñado para humanos!

 [The Manifesto](#)

# El núcleo *Tidyverse*



Number of core tidyverse packages downloaded Sept 10, 2018

Each icon represents 4000 downloads, rounding to nearest 4000

Data: cran-logs.rstudio.com



Fuente:

[https://evamaerey.github.io/little\\_flipbooks\\_library/ggtextures/ggtextures#1](https://evamaerey.github.io/little_flipbooks_library/ggtextures/ggtextures#1)

```
# Install from CRAN  
install.packages("tidyverse")
```



# Hastá acá

(1) Codeo

`base` vs *Tidyverse*

# Lo que falta

(2) Organización de proyectos

→ `.Rproj` y `here::`

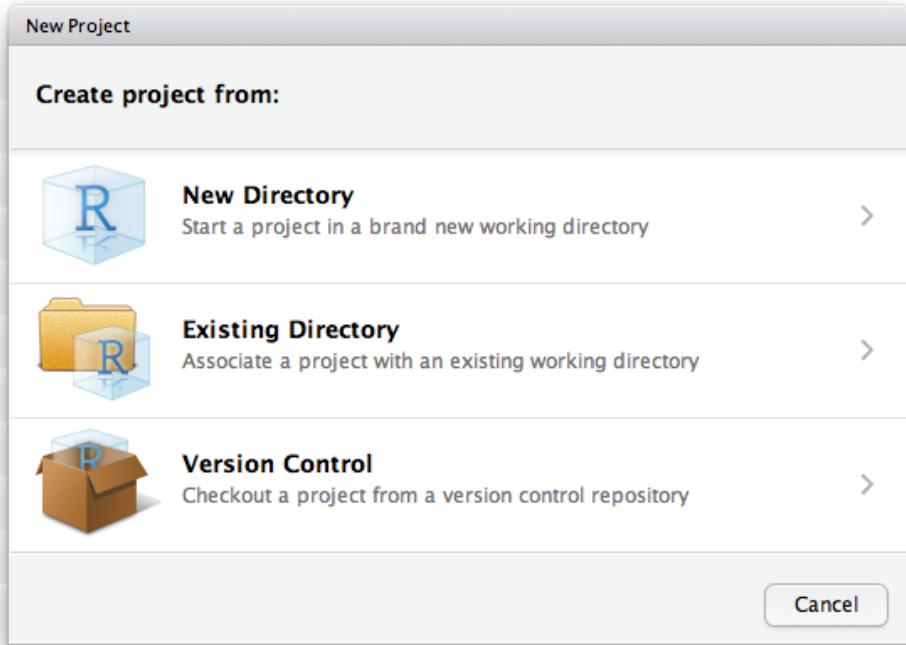
(3) Reproducibilidad y Comunicación

→ `R` y `Rmd`

R

(Extras)

# .Rproj



## ⌚ Using Projects

# {here}



# ⌚ {here}

# R Markdown

