

Econometría I

Taller R

26 de agosto de 2021

¿Qué vamos a presentar?

- ▶ R y RStudio
 - ▶ Instalación
 - ▶ RStudio Cloud
 - ▶ Características generales
- ▶ Rproject
- ▶ Librerías y paquetes
 - ▶ tidyverse
- ▶ Comandos básicos
- ▶ Clases de objetos
- ▶ Importar datos
 - ▶ readr
 - ▶ algunas funciones basicas
- ▶ Obtención de ayuda

Descarga e instalación de R y RStudio

- ▶ Descargar R

<https://www.r-project.org/>

- ▶ Descargar RStudio

<https://www.rstudio.com/products/rstudio/download/>

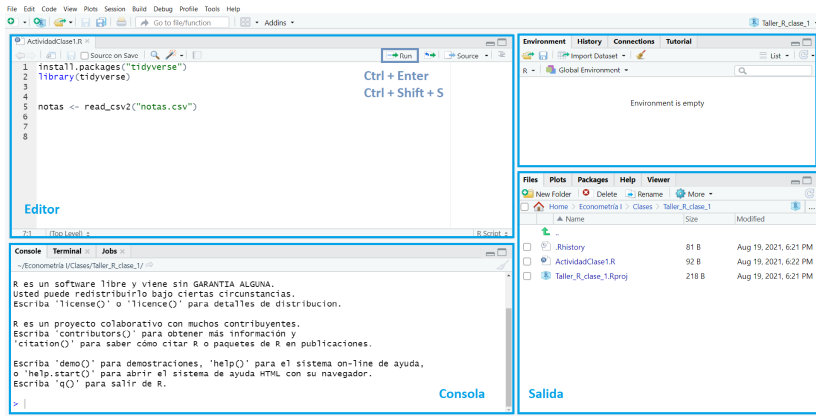
- ▶ VIDEO: Tutorial para instalación

<https://www.youtube.com/watch?v=ZTzbiAznSjc>

- ▶ Si queremos trabajar en la nube:

<https://rstudio.cloud>

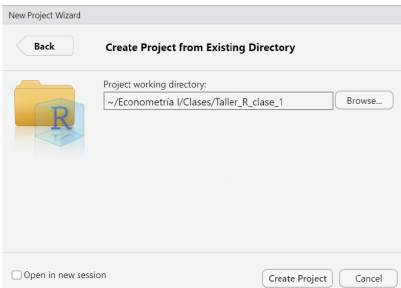
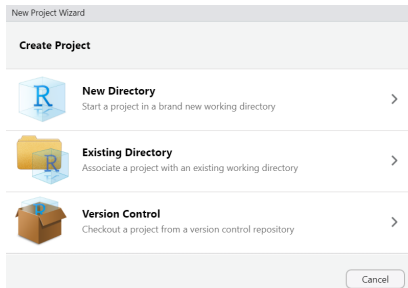
Características generales



¿Qué es un project?

- ▶ *Directorio de trabajo* (working directory): aquí es donde R busca y guarda los archivos.
- ▶ Un *proyecto* permite mantener todos los archivos asociados en un mismo lugar — datos de entrada, scripts, resultados, gráficos.
- ▶ Hagamos un proyecto:

Haz clic en File > New Project, y después:



Librerías y paquetes

- ▶ Un paquete es una colección de funciones, datos y documentación que permite extender las capacidades de R base.
- ▶ *Tidyverse* es un conjunto de paquetes que comparten una misma filosofía de datos y están diseñados para trabajar conjuntamente de forma natural.
- ▶ Instalemos el paquete tidyverse:

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

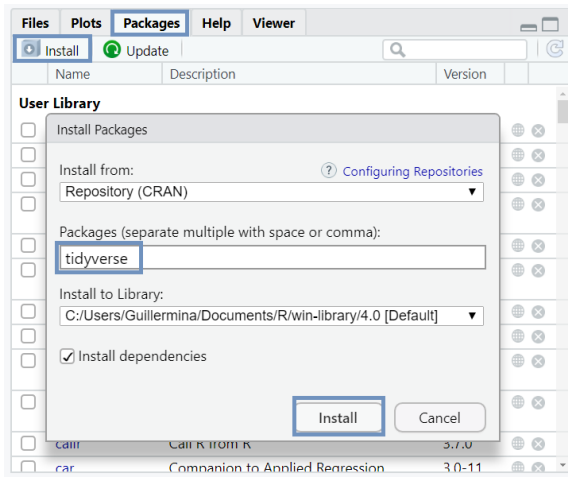
- ▶ Una vez instalado el paquete, cargémoslo:

```
library(tidyverse)
```

► Otra forma de instalar paquetes:

Haz click en Packages > Install, y se abre una ventana emergente.

Escribe el nombre del paquete que quieras instalar, finalmente click en Install.



Comandos básicos

- Puedes usar R como una calculadora:

```
1 / 200 * 30
```

```
## [1] 0.15
```

```
(59 + 73 + 2) / 3
```

```
## [1] 44.66667
```

```
sin(pi / 2)
```

```
## [1] 1
```


- ▶ Puedes crear objetos nuevos usando `<-`:
 - ▶ `nombre_objeto <- valor`
 - ▶ Alt + - (signo menos).
 - ▶ Los nombres de los objetos deben comenzar con una letra y solo pueden contener letras, números, `_` y `..`

```
x <- 3 * 4
```

- ▶ Puedes examinar un objeto escribiendo su nombre:

```
x
```

```
## [1] 12
```

- ▶ Puedes crear un vector con la función `c()`:

```
v1 <- c(1, 2, 3, 4.5, "seis")
```

```
v1
```

```
## [1] "1"      "2"      "3"      "4.5"    "seis"
```

```
v2 <- c(3, 7:10)
```

```
v2
```

```
## [1] 3 7 8 9 10
```

- ▶ Puedes crear matrices.

¿Qué clases de objetos existen?

- ▶ *Vectores*
- ▶ *Matrices y Arrays*: Todos los elementos son numéricos
`matrix()`
- ▶ *Listas*: pueden contener elementos de distinto tipo `list()`
- ▶ *Factores*: se usan para variables categóricas `factor()`
- ▶ *Data Frames*: se usan para almacenar datos en forma de tablas (filas / columnas). Pueden contener elementos de distinto tipo
`data.frame()`
- ▶ *Tibbles*: dataframes con algunas particularidades por ejemplo no convierten por defecto vectores de texto en factores
- ▶ *Funciones*: argumento, cuerpo, resultado

¿Cómo importamos datos?

- ▶ Utilizamos la librería `readr` del paquete `tidyverse`.
 - ▶ `read_csv()` lee archivos delimitados por coma.
 - ▶ `read_csv2()` lee archivos separados por punto y coma.
 - ▶ `read_tsv()` lee archivos delimitados por tabulaciones.
 - ▶ `read_delim()` archivos con cualquier delimitador.
- ▶ El archivo `notas.csv`, incluye una muestra de los resultados de las personas que rindieron la segunda revisión de Econometría 1 para 2017-2018 y lo pueden encontrar aquí:

https:

[//eva.fcea.udelar.edu.uy/course/view.php?id=183§ion=10](https://eva.fcea.udelar.edu.uy/course/view.php?id=183§ion=10)

- Leemos el archivo `notas.csv` que se encuentra guardado en la misma carpeta del proyecto.

```
notas <- read.csv2("Datos/notas.csv")  
notas <- read_delim("Datos/notas.csv", delim = ";")
```

```
## # A tibble: 47 x 4  
##   primera `segunda` year Mujer  
##   <dbl>      <dbl> <dbl> <dbl>  
## 1     415         41  2018     1  
## 2    2725        3575  2018     1  
## 3      20         19  2018     1  
## 4      26        4025  2018     1  
## 5     245        3375  2018     0  
## 6     312         305  2018     0  
## 7     205        1925  2018     0  
## 8     315        3625  2018     1  
## 9    2625         28  2018     1  
## 10    2655        2475  2018     0  
## # ... with 37 more rows
```

1. Los datos pueden no tener nombres de columna. Se utiliza `col_names = FALSE`.

```
read_csv2("Datos/notas.csv", col_names = FALSE)
```

Alternativamente,

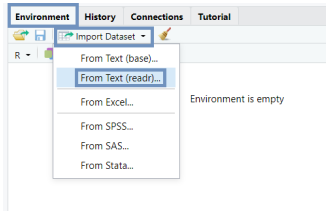
```
read_csv2("Datos/notas.csv", col_names=c("P", "S", "A", "M"))
```

2. `na` del inglés, “not available”, especifica el valor que se utiliza para representar los valores faltantes del archivo:

```
read_csv2("Datos/notas.csv", na = ".")
```

- Otra forma de importar datos:

Haz click en Environment > Import Dataset.



Se abre una ventana emergente:

Import Text Data

File/URL:
~/Econometria I/Clases/notas.csv Browse...

Data Preview:

primera (double)	segunda (double)	year (double)	Mujer (double)
415	41	2018	1
2725	3375	2018	1
20	19	2018	1
26	4025	2018	1
245	3375	2018	0
312	305	2018	0
205	1925	2018	0
315	3625	2018	1
2625	28	2018	1
2655	2475	2018	0
2525	1975	2018	1
2725	3125	2018	0

Previewing first 50 entries.

Import Options:

Name: ☒ First Row as Names Delimiter: Semicolon Escape: None
Skip: ☒ Trim Spaces Quotes: Default Comment: Default
☒ Open Data Viewer Locale: Configure... NA: Default

Code Preview:

```
library(readr)
notas <- read_delim("~/Econometria I/Clases/notas.csv",
  delim = ";", escape_double = FALSE, trim_ws = TRUE)
View(notas)
```

Import Cancel

? Reading rectangular data using readr

Algunas funciones básicas

- ▶ Una vez importados los datos podemos hacer estas funciones:
 - ▶ `view()` visualiza los datos como una “tabla”.
 - ▶ `names()` muestra o establece los nombres de un objeto.
 - ▶ `head()` devuelve las primeras 6 partes de un objeto.
 - ▶ `tail()` devuelve las ultimas 6 partes de un objeto.
 - ▶ `dim()` establece la dimensión de un objeto.
 - ▶ `str()` muestra de forma compacta la estructura interna de un objeto.
 - ▶ `summary()` produce resúmenes de los datos o de los resultados de funciones de ajuste de modelos.

```
names(notas)
```

```
## [1] "primera" "segunda" "year" "Mujer"
```

```
head(notas)
```

```
## # A tibble: 6 x 4
```

```
## primera `segunda` year Mujer
```

```
## <dbl> <dbl> <dbl> <dbl>
```

```
## 1 415 41 2018 1
```

```
## 2 2725 3575 2018 1
```

```
## 3 20 19 2018 1
```

```
## 4 26 4025 2018 1
```

```
## 5 245 3375 2018 0
```

```
## 6 312 305 2018 0
```

```
dim(notas)
```

```
## [1] 47  4
```

```
str(notas)
```

```
## spec_tbl_df [47 x 4] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ primera  : num [1:47] 415 2725 20 26 245 ...
## $ segunda  : num [1:47] 41 3575 19 4025 3375 ...
## $ year      : num [1:47] 2018 2018 2018 2018 2018 ...
## $ Mujer     : num [1:47] 1 1 1 1 0 0 0 1 1 0 ...
## - attr(*, "spec")=
## .. cols(
## ..   primera = col_number(),
## ..   ` segunda ` = col_number(),
## ..   year = col_double(),
## ..   Mujer = col_double()
## .. )
## - attr(*, "problems")=<externalptr>
```

```
summary(notas)
```

##	primera	segunda	year	Mu
##	Min. : 20.0	Min. : 13.0	Min. :2017	Min.
##	1st Qu.: 27.0	1st Qu.: 27.5	1st Qu.:2017	1st Qu.
##	Median : 225.0	Median : 35.0	Median :2017	Median
##	Mean : 424.4	Mean : 653.6	Mean :2017	Mean
##	3rd Qu.: 315.0	3rd Qu.: 295.0	3rd Qu.:2018	3rd Qu.
##	Max. :2725.0	Max. :4025.0	Max. :2018	Max.

Busquemos ayuda

Puedes encontrar material de ayuda en:

- ▶ Sitio web de la versión en español de “R for Data Science”, de Hadley Wickham y Garrett Grolemund.

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

- ▶ “Machetes” (cheatsheets) de Rstudio.

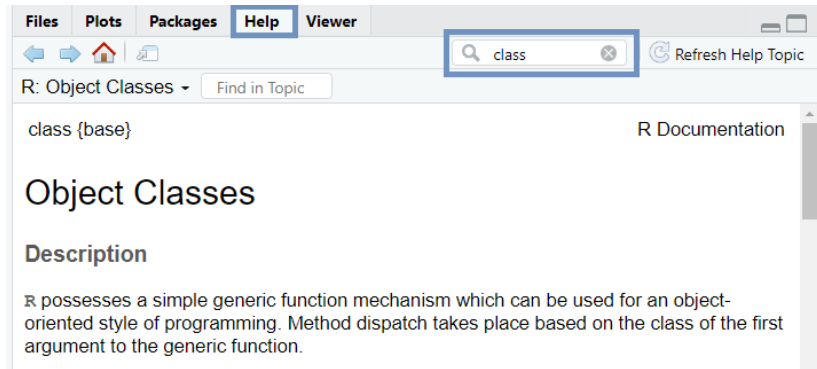
<https://www.rstudio.com/resources/cheatsheets/>

- ▶ En particular el cheatsheet de importación de datos.

https://eva.fcea.udelar.edu.uy/pluginfile.php/286031/mod_folder/content/0/data-import.pdf?forcedownload=1

Busquemos ayuda en R

- ▶ en la consola escribe ? + nombre de la función
- ▶ en la consola escribe ?? + algo parecido al nombre de la función
- ▶ en la pestaña Help de la salida



The screenshot shows the R Help viewer interface. At the top, there are tabs for 'Files', 'Plots', 'Packages', 'Help' (which is selected and highlighted with a blue border), and 'Viewer'. Below the tabs is a search bar containing the text 'class' and a magnifying glass icon. To the right of the search bar is a 'Refresh Help Topic' button. Below the search bar, the text 'R: Object Classes' is displayed, followed by a 'Find in Topic' button. The main content area shows the title 'class {base}' and 'R Documentation'. Below this, the heading 'Object Classes' is displayed. Underneath, the heading 'Description' is shown, followed by a paragraph of text: 'R possesses a simple generic function mechanism which can be used for an object-oriented style of programming. Method dispatch takes place based on the class of the first argument to the generic function.'

Actividad

1. Descargar e instalar R y RStudio. Abrir RStudio.
2. Descargar la base de datos `Notas.csv` del EVA y guardarla en una carpeta llamada `Taller_R_Econometria`.
3. Crear un proyecto en el directorio ya existente `Taller_R_Econometria`.
4. Abrir un nuevo script y nombrarlo `ActividadClase1`.
5. Descargar el paquete e instalar la librería `tidyverse`.
6. Importar el archivo `Notas.csv`.

7. Observar los nombres de las variables, las primeras 6 líneas , la dimensión y un resumen de cada variable.
8. Relizar un gráfico de dispersión entre dos variables y un histograma de otra. (Ayuda: investiga sobre la función `plot()`)
9. Crear un vector, con el nombre `vector` con los valores 4.3456, 2.8888, -2.25 y -4.
10. Hallar el valor absoluto de los valores de `vector`. (Ayuda: investiga sobre la función `abs()`)
11. Redondea los valores de `vector` con una cifra despues de la coma. (Ayuda: investiga sobre la función `round()`).