

Armando un proyecto de trabajo

Big Data e Inteligencia Territorial

Hoja de ruta

- ✓ Estructura de carpetas
- ✓ Proyectos de trabajo (*R project*)

Proyectos y rutas de trabajo

Proyectos y rutas de trabajo

Intentar correr el siguiente script:

```
library(tidyverse)
```

```
b_inmuebles <- read_csv("C:\Users\pablo\Documents\Pablo\Cursos R\R+FLACSO\R+BigData\r_bigdataEit
```

Proyectos y rutas de trabajo

Intentar correr el siguiente script:

```
library(tidyverse)
```

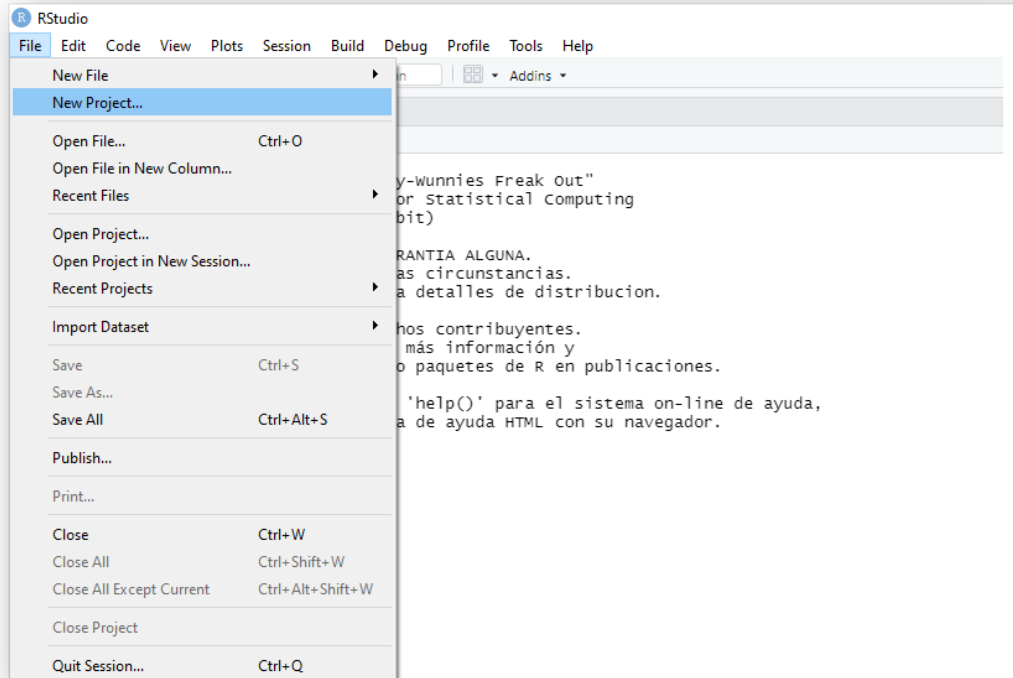
```
b_inmuebles ← read_csv("C:\Users\pablo\Documents\Pablo\Cursos R\R+FLACSO\R+BigData\r_bigdataEit
```

A tener en cuenta respecto a la sentencia para importar el archivo (nuestra base):

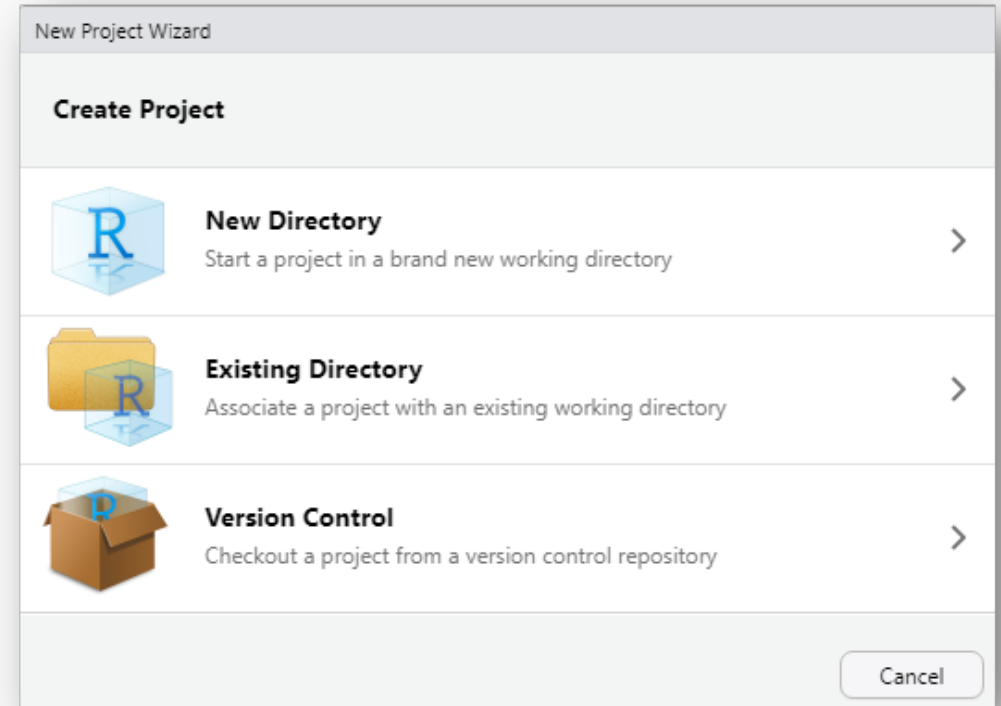
- Si se la compartimos a otra persona, **se rompe**
- Si cambiamos de computadora, **se rompe**
- Si lo cambiamos de lugar, **se rompe**

Armar un nuevo proyecto

Paso 1:

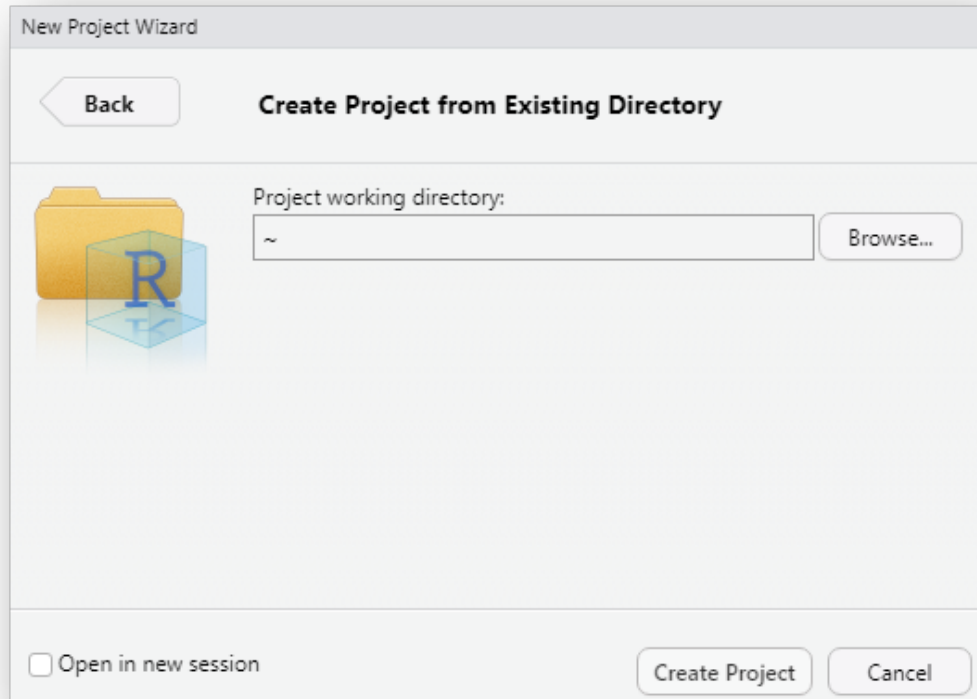


Paso 2:

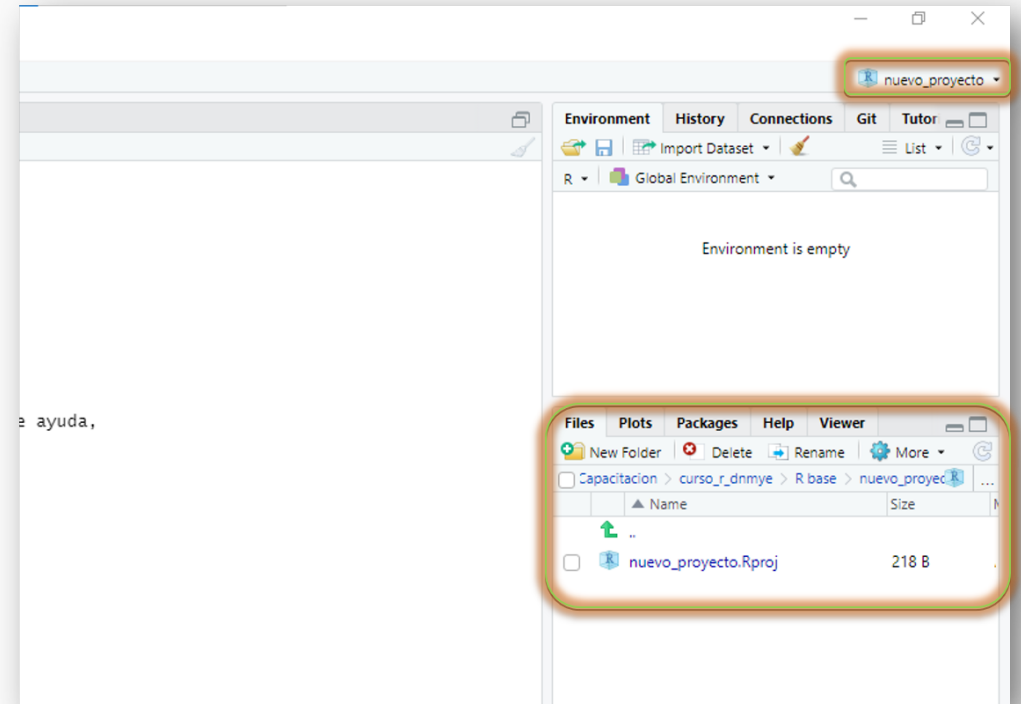


Armar un nuevo proyecto

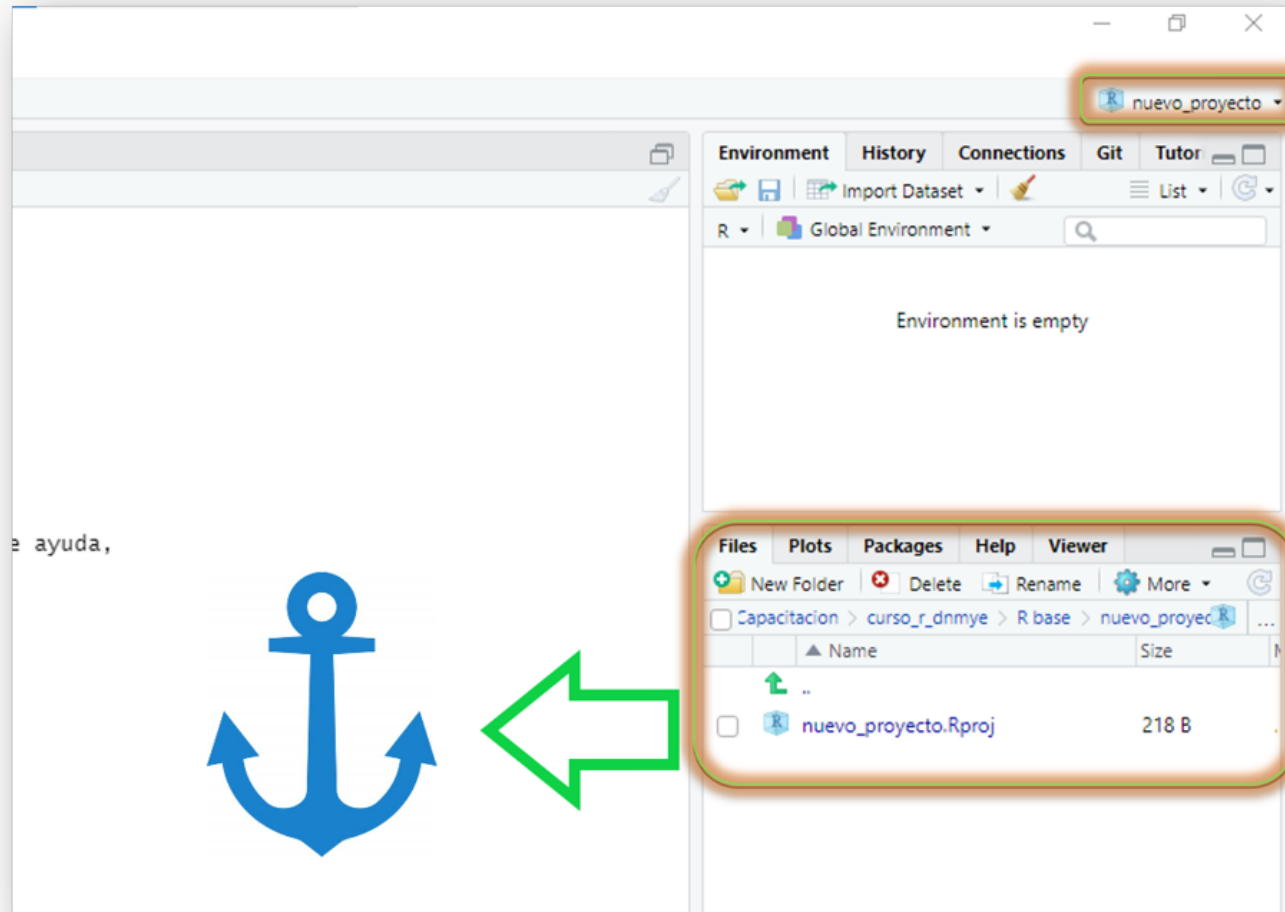
Paso 3:



Resultado:



Armar un nuevo proyecto



Proyectos y rutas de trabajo

Armando un proyecto con Rstudio, al ubicar la base de datos en la carpeta, la función se escribiría de esta forma:

```
library(tidyverse)
```

```
b_inmuebles ← read_csv("inmuebles-estado-nacional.csv")
```

Proyectos y rutas de trabajo

Armando un proyecto con Rstudio, al ubicar la base de datos en la carpeta, la función se escribiría de esta forma:

```
library(tidyverse)  
  
b_inmuebles ← read_csv("inmuebles-estado-nacional.csv")
```

Ventajas:

- R pone un ancla, un punto de partida desde donde recorre las carpetas de la computadora.
- A = orden de carpeta → = código para encontrar el archivo.

Flujo de trabajo




Supongamos que queremos comenzar un nuevo proyecto para analizar el conjunto de inmuebles del Estado Nacional Argentino.

Flujo de trabajo

Supongamos que queremos comenzar un nuevo proyecto para analizar el conjunto de inmuebles del Estado Nacional Argentino.







El flujo de trabajo sería el siguiente:

1) Armo la carpeta del proyecto, por ejemplo con la estructura siguiente:

-  inmuebles_estado
 -  datos
 -  resultados
 -  scripts
 -  docs








Proyecto de trabajo

2) Ubico la base de datos en la carpeta correspondiente (**datos**):

-  inmuebles_estado
 -  datos
 -  inmuebles-estado-nacional.csv
 -  resultados
 -  scripts
 -  docs_metodologicos


Proyecto de trabajo

3) Abro Rstudio y creo un proyecto, ubicado en la carpeta principal (**inmuebles_estado**)

-  inmuebles_estado
 -  *proyecto_inmuebles_estado.Rproj*
 -  datos
 -  *inmuebles-estado-nacional.csv*
 -  resultados
 -  scripts
 -  docs

Proyecto de trabajo

4) Creo un script llamado *levantar_datos.R* para trabajar en R y lo guardo en la carpeta **scripts**

-  inmuebles_estado
 -  *inmubles_estado.Rproj*
 -  datos
 -  *inmuebles-estado-nacional.csv*
 -  resultados
 -  scripts
 -  *1_levantar_datos.R*
 -  docs

Proyecto de trabajo

5) En el script, cargo los paquetes cuyas funciones voy a utilizar. En este caso, el paquete `tidyverse` para poder importar bases de datos con la función `read_csv()`:

```
library(tidyverse)
```


Proyecto de trabajo

5) En el script, cargo los paquetes cuyas funciones voy a utilizar. En este caso, el paquete `tidyverse` para poder importar bases de datos con la función `read_csv()`:

```
library(tidyverse)
```

6) Creo un objeto en el cual voy a importar la base de datos y utilizo la función `read_csv()` para hacerlo.

Recordar: *Estamos en un proyecto, por lo que la ruta comienza en la carpeta en la que se encuentra el archivo .Rproj*

```
b_inmuebles ← read_csv("datos/inmuebles-estado-nacional.csv")
```

Funciones para una primera mirada a la base de datos:

Función	Acción
<code>dim()</code>	<i>Nos dice cuántas filas y columnas tiene la base</i>
<code>colnames()</code>	<i>Devuelve todos los nombres de las variables de la base</i>
<code>str()</code>	<i>Devuelve una vista preliminar de la estructura de las variables</i>
<code>summary()</code>	<i>Devuelve un resumen de medidas centrales para variables numéricas y características para variables de texto</i>
<code>class()</code>	<i>Chequeo el tipo de objeto</i>
<code>length()</code>	<i>Devuelve la cantidad de elementos de un objeto</i>
<code>table()</code>	<i>Frecuencia simple o bivariada de una variable</i>
<code>is.na()</code>	<i>Verifica la presencia de valores NA (missinh) en la base/variable</i>
<code>mean()</code>	<i>Calcula el valor promedio de una variable numérica</i>

PRÁCTICA

Práctica

1. Crear una carpeta donde alojar el proyecto de trabajo
2. Crear la estructura de carpetas propuesta
3. En la carpeta **datos** (o como la hayas nombrado) incorporar una base de datos, preferentemente que usemos cotidianamente.
4. Crear un script de trabajo y alojarlo en la carpeta **scripts** (o como la hayas nombrado)
5. Importar la base de datos y aplicar alguna de las funciones descritas previamente

10:00