

Paquetes y lectura de archivos

Big Data e Inteligencia Territorial

Paquetes

Paquetes

¿Qué son?

Paquetes

¿Qué son?

✓ Los paquetes son, principalmente, funciones que no están en r base, pero también bases de datos y objetos de R en general.

Paquetes

¿Qué son?

✓ Los paquetes son, principalmente, funciones que no están en r base, pero también bases de datos y objetos de R en general.

¿Para qué nos sirven?

✓ Ayudan a resolver problemas que con r base no se podía, o que era muy difícil hacerlo.

R =



R base =



Paquete =



Paquete

Funciones =



R

Funciones =

Paquetes =



Paquetes

INSTALACIÓN

- Podemos descargarlos e instalarlos con el siguiente comando:

```
install.packages("nombre_del_paquete")
```

- Sólo es necesario instalar el paquete una vez por computadora.

Paquetes

INVOCACIÓN

- Una vez instalado, cada vez que abramos una nueva sesión de R y querramos utilizar el paquete debemos **cargarlo al ambiente de trabajo** mediante la función:

```
library(nombre_del_paquete)
```



Tidyverse

Tidyverse es una colección de paquetes de R, pensados para trabajar en "ciencia de datos".



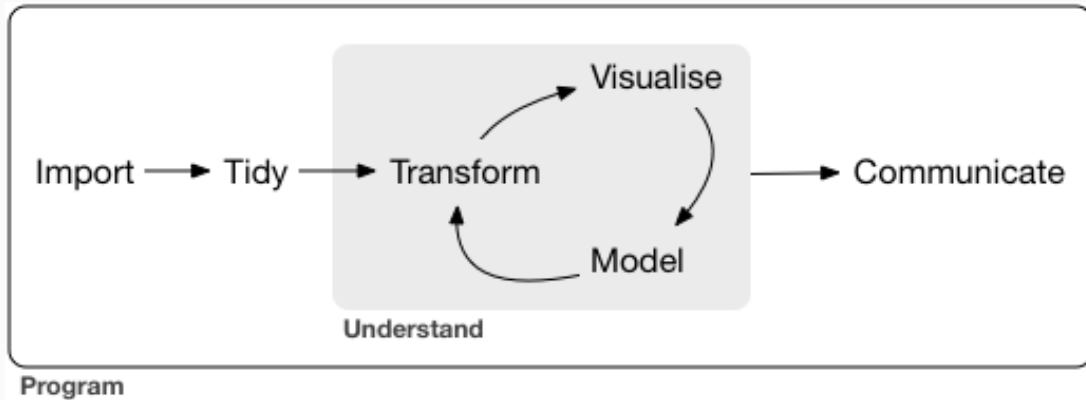
¿Por qué tidyverse?

¿Por qué tidyverse?

- Orientado a ser leído y escrito por y para seres humanos

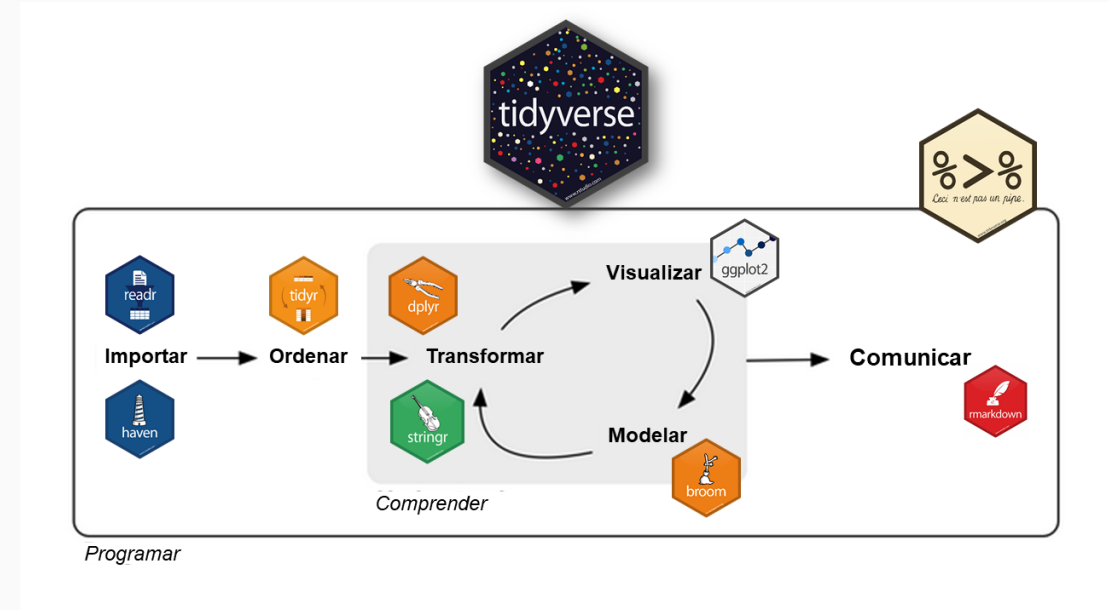
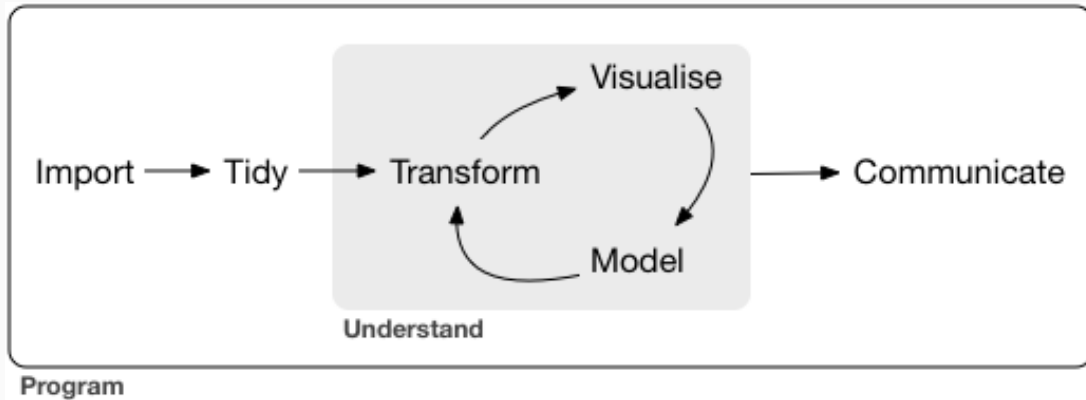
¿Por qué tidyverse?

- Orientado a ser leído y escrito por y para seres humanos
- Funciones no pensadas para una tarea específica sino para un proceso de trabajo



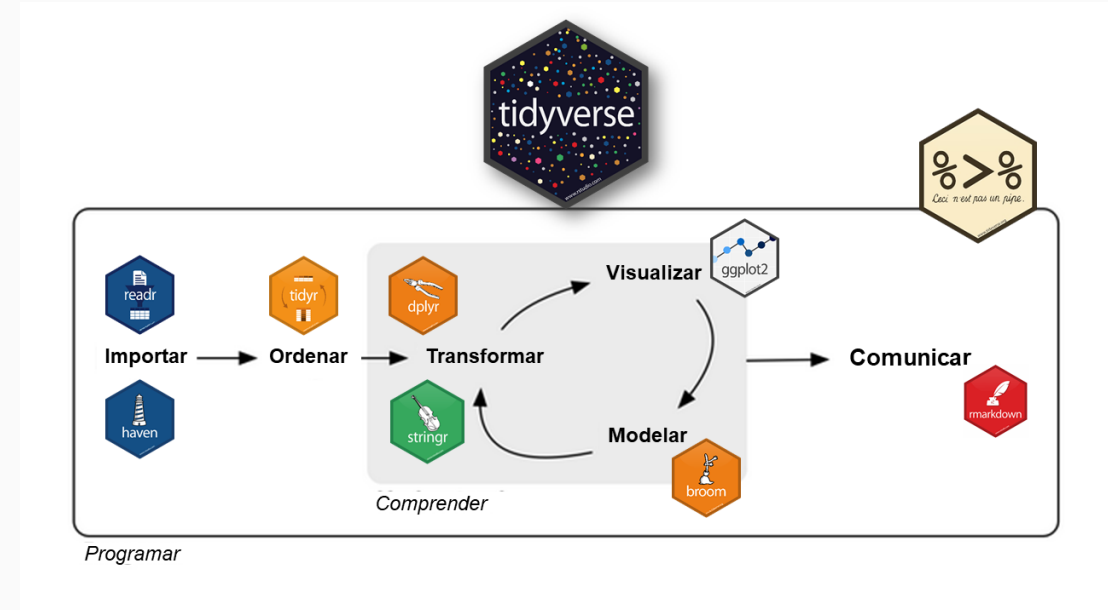
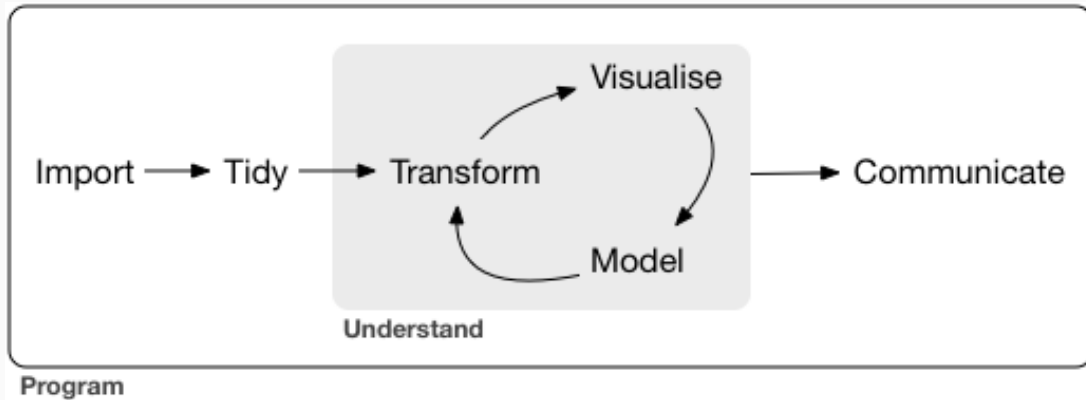
¿Por qué tidyverse?

- Orientado a ser leído y escrito por y para seres humanos
- Funciones no pensadas para una tarea específica sino para un proceso de trabajo



¿Por qué tidyverse?

- Orientado a ser leído y escrito por y para seres humanos
- Funciones no pensadas para una tarea específica sino para un proceso de trabajo



- Su comunidad, basada en los principios del código abierto y trabajo colaborativo

Instalación y uso

- Sólo una vez (por computadora):

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

Instalación y uso

- Sólo una vez (por computadora):

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

- En cada inicio de sesión de R o Rstudio:

```
library(tidyverse)
```

Instalación y uso

- Sólo una vez (por computadora):

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

- En cada inicio de sesión de R o Rstudio:

```
library(tidyverse)
```

No es necesario esto:

```
install.packages("dplyr")  
install.packages("tidyr")  
install.packages("ggplot2")
```

Ni esto:

```
library(dplyr)  
library(tidyr)  
library(ggplot2)
```

Hoja de ruta

Presentación de los paquetes `dplyr` y `tidyr`

✓ lectura / escritura

{*readr*}: ✓ `read_csv()` ✓ `read_table()`

{*haven*}: ✓ `read_sav()` ✓ `read_dat()`

✓ `read_sas()`

✓ magrittr

✓ `%>%`

Hoja de ruta

✓ dplyr

✓ `select()` ✓ `filter()`

✓ `mutate()` ✓ `rename()`

✓ `arrange()` ✓ `summarise()`

✓ `group_by()`

✓ tidyr

✓ `pivot_longer()` ✓ `pivot_wider()`

Lectura / escritura de archivos

Lectura / escritura de archivos

Hay varios paquetes que ayudan a la lectura de archivos, los principales son:

Paquete `{readr}`

- `read_csv()` --> `.CSV`
- `read_table()` --> `.txt`
- `read_rds()` --> `.rds`

Paquete `{haven}`

- `read_sav()` --> SPSS
- `read_dta()` --> STATA
- `read_sas()` --> SAS

Base de trabajo

En el material de la clase se puso a disposición una base de ejercicio. La información con la que se cuenta es aquella referida a los **Inmuebles Propios del Estado Nacional y Alquilados**, provista por la Dirección Nacional del Registro de Bienes Inmuebles - Agencia de Administración de Bienes del Estado - Jefatura de Gabinete de Ministros.

 [Descarga de la base de datos](#)

Base de trabajo

En el material de la clase se puso a disposición una base de ejercicio. La información con la que se cuenta es aquella referida a los **Inmuebles Propios del Estado Nacional y Alquilados**, provista por la Dirección Nacional del Registro de Bienes Inmuebles - Agencia de Administración de Bienes del Estado - Jefatura de Gabinete de Ministros.

[Descarga de la base de datos](#)

Con la siguiente sentencia puedo importarla a R:

```
base_inmuebles <- read_csv(file = "la_ruta_del_archivo/el_nombre_del_archivo.csv",  
                           col_names = TRUE)
```

PRÁCTICA

Práctica

1. Armar una carpeta de trabajo en la computadora.
2. Descargar la base de datos de inmuebles del Estado Nacional y ubicar el archivo en la carpeta de trabajo. Link para descarga [aquí](#)
3. Desde un script de Rstudio, importar la base de datos con la función acorde a la extensión del archivo.

Pistas:

Del paquete *{readr}*

- .csv --> *read_csv()*
- .txt --> *read_delim()*
- .xlsx (excel) --> Paquete *{readxl}* --> *read_xlsx()* o *read_excel()*