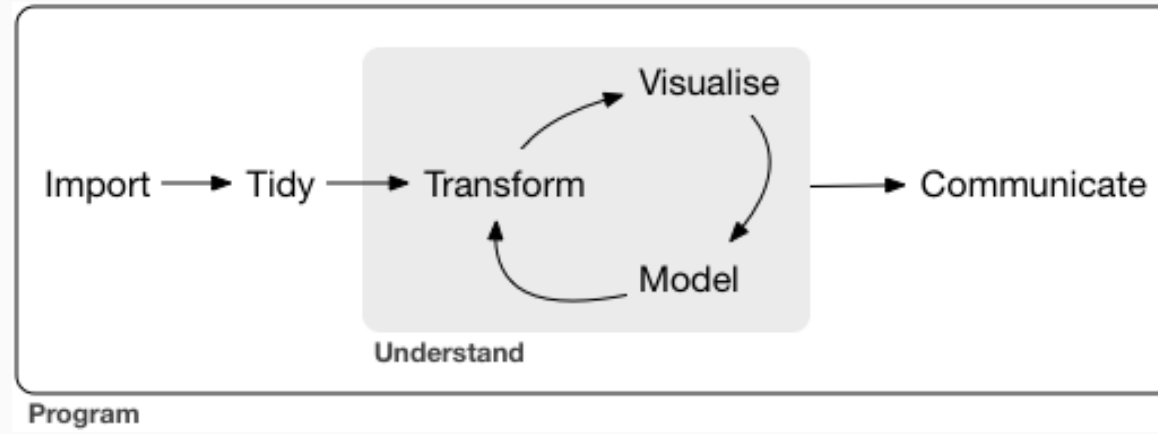
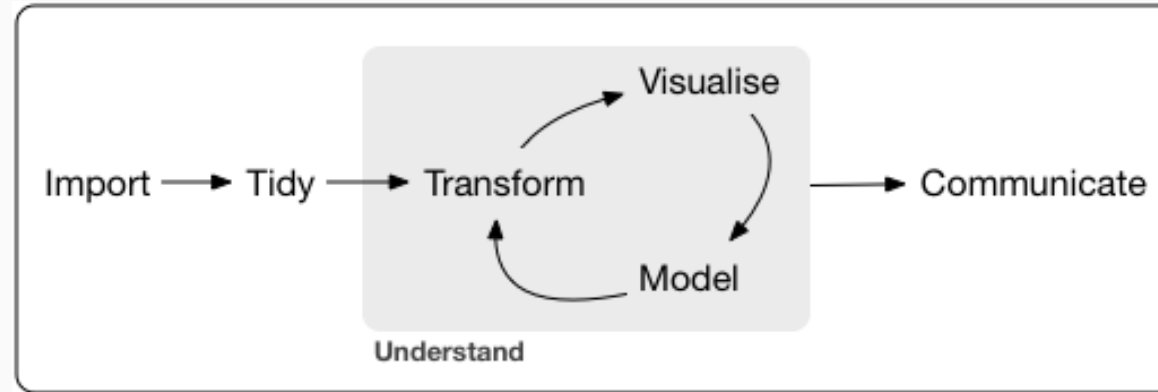


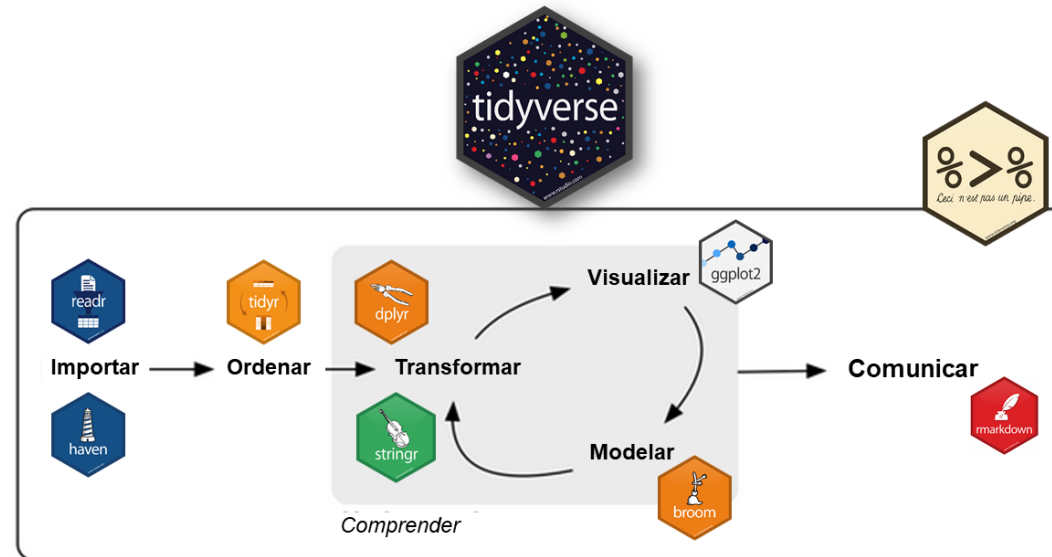
Procesando datos con el paquete tidyverse

Big Data e Inteligencia Territorial





Program



Programar

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`

✓ `dplyr`

✓ `select()` ✓ `filter()`

✓ `mutate()` ✓ `rename()`

✓ `arrange()` ✓ `summarise()`

✓ `group_by()`

✓ `magrittr`

✓ `%>%`

✓ `tidyr`

✓ `pivot_longer()` ✓ `pivot_wider()`

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)
```

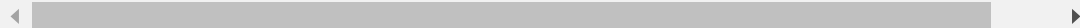
EL PIPE



Una forma de escribir

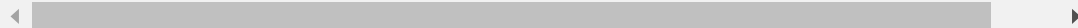
Sin EL PIPE:

```
### Calculo la media  
round(prop.table(table(base_inmuebles$pais)), digits
```



Sin EL PIPE:

```
### Calculo la media  
round(prop.table(table(base_inmuebles$pais)), digits
```



Con EL PIPE

```
### Cargo la librería  
library(tidyverse)  
  
### Calculo la media  
`base_inmuebles$pais` %>%  
  table() %>%  
  prop.table() %>%  
  round(digits = 2)
```

PRÁCTICA

Práctica

1. Calcular una frecuencia de la variable **situacion_inmueble** (función `table()`) en **r base**
2. Calcular una frecuencia de la variable **situacion_inmueble** (función `table()`) con el **pipe** (`%>%`)
3. En la siguiente sentencia se busca conocer la **cantidad** y **porcentaje** de inmuebles que hay por provincia (variable `provincia`). Identificar el error y corregir:

```
base_inmuebles %>%  
  table() %>%  
  prop.table()
```

1. Traduzca la siguiente sentencia al esquema *pipe*:

```
round(mean(base_inmuebles$superficie_aproximada_m2), digits = 1)
```


Funciones del paquete dplyr:

| Función | Acción |
|--------------------------|--|
| <code>select()</code> | <i>selecciona o descarta variables</i> |
| <code>filter()</code> | <i>selecciona filas</i> |
| <code>mutate()</code> | <i>crea / edita variables</i> |
| <code>rename()</code> | <i>renombra variables</i> |
| <code>group_by()</code> | <i>segmenta en funcion de una variable</i> |
| <code>summarize()</code> | <i>genera una tabla de resúmen</i> |

select()

Elije o descarta columnas de una base de datos

select()

La función tiene el siguiente esquema:

```
base_de_datos %>%  
  select(variable_1, variable_2, variable_n)
```

select()

Supongamos que quiero trabajar con las variables que permiten ubicar a cada uno de los inmuebles del Estado alrededor del país y también del mundo.

select() - por nombre de variable

selecciono las columnas que deseo de la base de datos:

```
base_inm_seleccion <- base_inmuebles %>%  
  select(codigo_del_inmueble, pais, provincia, departamento)
```

select() - por nombre de variable

selecciono las columnas que deseo de la base de datos:

```
base_inm_seleccion <- base_inmuebles %>%  
  select(codigo_del_inmueble, pais, provincia, departamento)
```

Chequeo la operación:

```
colnames(base_inm_seleccion)
```

```
## [1] "codigo_del_inmueble" "pais"          "provincia"  
## [4] "departamento"
```

select() - por posición de la columna

Supongamos que quiero las columnas **"codigo_del_inmueble"**, **"calle"** y **"numero"**

select() - por posición de la columna

Supongamos que quiero las columnas **"codigo_del_inmueble"**, **"calle"** y **"numero"**

1) Chequeo la posición de las columnas que deseo:

```
colnames(base_inmuebles)
```

```
## [1] "codigo_del_inmueble"      "organizacion"
## [3] "pais"                    "provincia"
## [5] "departamento"           "localidad"
## [7] "calle"                   "numero"
## [9] "situacion_inmueble"      "superficie_aproximada_m2"
## [11] "longitud"                "latitud"
## [13] "str_trim(calle)"
```

select() - por posición de la columna

2) Aplico la función `select()` en base a la posición de las columnas:

```
base_inm_seleccion2 <- base_inmuebles %>%  
  select(1, 7, 8)
```


select() - por posición de la columna

2) Aplico la función `select()` en base a la posición de las columnas:

```
base_inm_seleccion2 <- base_inmuebles %>%  
  select(1, 7, 8)
```

chequeo seleccion:

```
colnames(base_inm_seleccion2)
```

```
## [1] "codigo_del_inmueble" "calle" "numero"
```

Otra forma de seleccionar

| | | |
|----------------|-------|---------------------|
| base_inmuebles | ## | codigo_del_inmueble |
| | ## 1 | 200002040 |
| | ## 2 | 200002059 |
| | ## 3 | 200002067 |
| | ## 4 | 200002075 |
| | ## 5 | 200002083 |
| | ## 6 | 200002091 |
| | ## 7 | 200002105 |
| | ## 8 | 200002113 |
| | ## 9 | 200002121 |
| | ## 10 | 200002131 |
| | ## 11 | 200002148 |
| | ## 12 | 200002156 |
| | ## 13 | 200002164 |
| | ## 14 | 200002172 |
| | ## 15 | 200002180 |
| | ## 16 | 200002199 |
| | ## 17 | 200002202 |
| | ## 18 | 200002210 |
| | ## 19 | 200002229 |
| | ## 20 | 200016521 |
| | ## 21 | 200043650 |
| | ## 22 | 200044010 |
| | ## 23 | 200044029 |
| | ## 24 | 200048261 |
| | ## 25 | 200048271 |
| | ## 26 | 200048288 |
| | ## 27 | 200048296 |

Otra forma de seleccionar

```
base_inmuebles %>%  
  select(8:10)
```

| ## | numero | situacion_inmueble |
|-------|-----------------|--------------------|
| ## 1 | 1355 | Propiedad del ENA |
| ## 2 | 1317 | Propiedad del ENA |
| ## 3 | 635 | Propiedad del ENA |
| ## 4 | 2221/51 | Propiedad del ENA |
| ## 5 | 675 | Propiedad del ENA |
| ## 6 | 53 | Propiedad del ENA |
| ## 7 | 1745 | Propiedad del ENA |
| ## 8 | 102-108-110 | Propiedad del ENA |
| ## 9 | 370 | Propiedad del ENA |
| ## 10 | 440 | Propiedad del ENA |
| ## 11 | 474/478/482/486 | Propiedad del ENA |
| ## 12 | 1378 | Propiedad del ENA |
| ## 13 | 841/45 | Propiedad del ENA |
| ## 14 | 7602/7618 | Propiedad del ENA |
| ## 15 | 350 | Propiedad del ENA |
| ## 16 | 463/465/467 | Propiedad del ENA |
| ## 17 | 1450/1452/1454 | Propiedad del ENA |
| ## 18 | 3102/38/40/42 | Propiedad del ENA |
| ## 19 | 162/8 | Propiedad del ENA |
| ## 20 | 1230 | Propiedad del ENA |
| ## 21 | 3102 | Propiedad del ENA |
| ## 22 | 131/135 | Alquilado |
| ## 23 | 1388 | Alquilado |
| ## 24 | 50 | Alquilado |
| ## 25 | 2499 | Alquilado |
| ## 26 | 2376 | Alquilado |
| ## 27 | 999 | Alquilado |

Una más!

Otra forma de seleccionar

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

Otra forma de seleccionar

```
base_inmuebles %>%  
  select(superficie_aproximada_m2:latitud)
```

| ## | superficie_aproximada_m2 | longitud | latitud |
|-------|--------------------------|-------------|------------|
| ## 1 | 946.18 | -58.3856773 | -34.608430 |
| ## 2 | 1361.25 | -58.3850685 | -34.608956 |
| ## 3 | 2740.15 | -58.3698314 | -34.615096 |
| ## 4 | 1461.43 | -58.3980500 | -34.610487 |
| ## 5 | 1137.77 | -58.3807414 | -34.600489 |
| ## 6 | 455.53 | -58.3804470 | -34.607741 |
| ## 7 | 841.70 | -58.3912102 | -34.608899 |
| ## 8 | 3295.14 | -58.3674397 | -34.621022 |
| ## 9 | 11165.97 | -58.3713990 | -34.609496 |
| ## 10 | 363.72 | -58.3740778 | -34.602866 |
| ## 11 | 1067.99 | -58.3800461 | -34.613787 |
| ## 12 | 761.85 | -58.3860127 | -34.606668 |
| ## 13 | 368.39 | -58.3785017 | -34.606920 |
| ## 14 | 473.79 | -58.4610415 | -34.543968 |
| ## 15 | 4399.31 | -58.3676263 | -34.611609 |
| ## 16 | 1151.19 | -58.3722211 | -34.613318 |
| ## 17 | 479.61 | -58.4310175 | -34.633905 |
| ## 18 | 19485.12 | -58.3903984 | -34.659623 |
| ## 19 | 1992.53 | -58.4146404 | -34.612290 |
| ## 20 | 594.02 | -58.3838762 | -34.610012 |
| ## 21 | 765.05 | -58.3900679 | -34.659937 |
| ## 22 | 442.18 | -58.3723035 | -34.609533 |
| ## 23 | 339.94 | -58.3862274 | -34.606544 |
| ## 24 | 725.51 | -58.4292277 | -34.618667 |
| ## 25 | 188.09 | -58.4016959 | -34.609678 |
| ## 26 | 500.44 | -58.4756772 | -34.606622 |
| ## 27 | 1027.84 | -58.4460231 | -34.568876 |

Una más!

Otra forma de seleccionar

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

Otra forma de seleccionar

```
base_inmuebles %>%  
  select(starts_with("s"))
```

| ## | situacion_inmueble | superficie_aproximada_m2 |
|-------|--------------------|--------------------------|
| ## 1 | Propiedad del ENA | 946.18 |
| ## 2 | Propiedad del ENA | 1361.25 |
| ## 3 | Propiedad del ENA | 2740.15 |
| ## 4 | Propiedad del ENA | 1461.43 |
| ## 5 | Propiedad del ENA | 1137.77 |
| ## 6 | Propiedad del ENA | 455.53 |
| ## 7 | Propiedad del ENA | 841.70 |
| ## 8 | Propiedad del ENA | 3295.14 |
| ## 9 | Propiedad del ENA | 11165.97 |
| ## 10 | Propiedad del ENA | 363.72 |
| ## 11 | Propiedad del ENA | 1067.99 |
| ## 12 | Propiedad del ENA | 761.85 |
| ## 13 | Propiedad del ENA | 368.39 |
| ## 14 | Propiedad del ENA | 473.79 |
| ## 15 | Propiedad del ENA | 4399.31 |
| ## 16 | Propiedad del ENA | 1151.19 |
| ## 17 | Propiedad del ENA | 479.61 |
| ## 18 | Propiedad del ENA | 19485.12 |
| ## 19 | Propiedad del ENA | 1992.53 |
| ## 20 | Propiedad del ENA | 594.02 |
| ## 21 | Propiedad del ENA | 765.05 |
| ## 22 | Alquilado | 442.18 |
| ## 23 | Alquilado | 339.94 |
| ## 24 | Alquilado | 725.51 |
| ## 25 | Alquilado | 188.09 |
| ## 26 | Alquilado | 500.44 |
| ## 27 | Alquilado | 1027.84 |

Una más!

Otra forma de seleccionar

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

Otra forma de seleccionar

```
base_inmuebles %>%  
  select(ends_with("o"))
```

| ## | departamento | numero |
|-------|--------------|-----------------|
| ## 1 | Comuna 1 | 1355 |
| ## 2 | Comuna 1 | 1317 |
| ## 3 | Comuna 1 | 635 |
| ## 4 | Comuna 3 | 2221/51 |
| ## 5 | Comuna 1 | 675 |
| ## 6 | Comuna 1 | 53 |
| ## 7 | Comuna 1 | 1745 |
| ## 8 | Comuna 1 | 102-108-110 |
| ## 9 | Comuna 1 | 370 |
| ## 10 | Comuna 1 | 440 |
| ## 11 | Comuna 1 | 474/478/482/486 |
| ## 12 | Comuna 1 | 1378 |
| ## 13 | Comuna 1 | 841/45 |
| ## 14 | Comuna 13 | 7602/7618 |
| ## 15 | Comuna 1 | 350 |
| ## 16 | Comuna 1 | 463/465/467 |
| ## 17 | Comuna 7 | 1450/1452/1454 |
| ## 18 | Comuna 4 | 3102/38/40/42 |
| ## 19 | Comuna 5 | 162/8 |
| ## 20 | Comuna 1 | 1230 |
| ## 21 | Comuna 4 | 3102 |
| ## 22 | Comuna 1 | 131/135 |
| ## 23 | Comuna 1 | 1388 |
| ## 24 | Comuna 6 | 50 |
| ## 25 | Comuna 15 | 2499 |
| ## 26 | Comuna 11 | 2376 |
| ## 27 | Comuna 14 | 999 |

Una más!

Otra forma de seleccionar

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

Otra forma de seleccionar

```
base_inmuebles %>%  
  select(contains("it"))
```

| ## | situacion_inmueble | longitud | latitud |
|-------|--------------------|-------------|------------|
| ## 1 | Propiedad del ENA | -58.3856773 | -34.608430 |
| ## 2 | Propiedad del ENA | -58.3850685 | -34.608956 |
| ## 3 | Propiedad del ENA | -58.3698314 | -34.615096 |
| ## 4 | Propiedad del ENA | -58.3980500 | -34.610487 |
| ## 5 | Propiedad del ENA | -58.3807414 | -34.600489 |
| ## 6 | Propiedad del ENA | -58.3804470 | -34.607741 |
| ## 7 | Propiedad del ENA | -58.3912102 | -34.608899 |
| ## 8 | Propiedad del ENA | -58.3674397 | -34.621022 |
| ## 9 | Propiedad del ENA | -58.3713990 | -34.609496 |
| ## 10 | Propiedad del ENA | -58.3740778 | -34.602866 |
| ## 11 | Propiedad del ENA | -58.3800461 | -34.613787 |
| ## 12 | Propiedad del ENA | -58.3860127 | -34.606668 |
| ## 13 | Propiedad del ENA | -58.3785017 | -34.606920 |
| ## 14 | Propiedad del ENA | -58.4610415 | -34.543968 |
| ## 15 | Propiedad del ENA | -58.3676263 | -34.611609 |
| ## 16 | Propiedad del ENA | -58.3722211 | -34.613318 |
| ## 17 | Propiedad del ENA | -58.4310175 | -34.633905 |
| ## 18 | Propiedad del ENA | -58.3903984 | -34.659623 |
| ## 19 | Propiedad del ENA | -58.4146404 | -34.612290 |
| ## 20 | Propiedad del ENA | -58.3838762 | -34.610012 |
| ## 21 | Propiedad del ENA | -58.3900679 | -34.659937 |
| ## 22 | Alquilado | -58.3723035 | -34.609533 |
| ## 23 | Alquilado | -58.3862274 | -34.606544 |
| ## 24 | Alquilado | -58.4292277 | -34.618667 |
| ## 25 | Alquilado | -58.4016959 | -34.609678 |
| ## 26 | Alquilado | -58.4756772 | -34.606622 |
| ## 27 | Alquilado | -58.4460231 | -34.568876 |

PRÁCTICA

Práctica

- 1) Crear un objeto en donde importamos la base de inmuebles del Estado (recordar tener en cuenta la extensión del archivo)
- 2) Crear otro objeto en donde selecciono 3 columnas de interés según sus nombres
- 3) Crear otro objeto en donde selecciono 3 columnas de interés según su posición

filter()

Define los casos (filas) en base a una condición

filter()

La función tiene el siguiente esquema:

```
base_de_datos %>%  
  filter(variable_x ≥ categoria)
```

filter()

La función tiene el siguiente esquema:

```
base_de_datos %>%  
  filter(variable_x ≥ categoria)
```

Caso: Supongamos que deseo quedarme sólo con los inmuebles ubicados en Estados Unidos:

filter()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

filter()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, pais)
```

| ## | codigo_del_inmueble | pais |
|-------|---------------------|-----------|
| ## 1 | 200002040 | Argentina |
| ## 2 | 200002059 | Argentina |
| ## 3 | 200002067 | Argentina |
| ## 4 | 200002075 | Argentina |
| ## 5 | 200002083 | Argentina |
| ## 6 | 200002091 | Argentina |
| ## 7 | 200002105 | Argentina |
| ## 8 | 200002113 | Argentina |
| ## 9 | 200002121 | Argentina |
| ## 10 | 200002131 | Argentina |
| ## 11 | 200002148 | Argentina |
| ## 12 | 200002156 | Argentina |
| ## 13 | 200002164 | Argentina |
| ## 14 | 200002172 | Argentina |
| ## 15 | 200002180 | Argentina |
| ## 16 | 200002199 | Argentina |
| ## 17 | 200002202 | Argentina |
| ## 18 | 200002210 | Argentina |
| ## 19 | 200002229 | Argentina |
| ## 20 | 200016521 | Argentina |
| ## 21 | 200043650 | Argentina |
| ## 22 | 200044010 | Argentina |
| ## 23 | 200044029 | Argentina |
| ## 24 | 200048261 | Argentina |
| ## 25 | 200048271 | Argentina |
| ## 26 | 200048288 | Argentina |
| ## 27 | 200048296 | Argentina |

filter()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, pais) %>%  
  filter(pais == "Estados Unidos")
```

| ## | codigo_del_inmueble | pais |
|-------|---------------------|----------------|
| ## 1 | 876 | Estados Unidos |
| ## 2 | 884 | Estados Unidos |
| ## 3 | 892 | Estados Unidos |
| ## 4 | 787 | Estados Unidos |
| ## 5 | 795 | Estados Unidos |
| ## 6 | 809 | Estados Unidos |
| ## 7 | 9900000275 | Estados Unidos |
| ## 8 | 9900000283 | Estados Unidos |
| ## 9 | 9900000291 | Estados Unidos |
| ## 10 | 9900000305 | Estados Unidos |
| ## 11 | 9900000313 | Estados Unidos |

filter()

Operadores para filtrar:

| Condición | Acción |
|-----------|--------------------------|
| = | <i>igual</i> |
| %in% | <i>incluye</i> |
| ≠ | <i>distinto</i> |
| > | <i>mayor que</i> |
| < | <i>menor que</i> |
| ≥ | <i>mayor o igual que</i> |
| ≤ | <i>menor o igual que</i> |

| Operador | Descripción |
|----------|---|
| & | y - Cuando se cumplen ambas condiciones |
| | o - Cuando se cumple una u otra condición |

filter()

Caso: Quiero quedarme tanto con los inmuebles de la provincia de *Cordoba* como los de **(o)** la provincia *Buenos aires*, **y** que sean *propios* (no alquilados):

- Provincias: Córdoba o Buenos Aires (`provincia = "Cordoba" | provincia = "Buenos Aires"`);

--> Alternativa:

- Provincias: Córdoba o Buenos Aires (`provincia %in% c("Cordoba", "Buenos Aires")`);

filter()

Caso: Quiero quedarme tanto con los inmuebles de la provincia de *Cordoba* como los de **(o)** la provincia *Buenos aires*, **y** que sean *propios* (no alquilados):

- Provincias: Córdoba o Buenos Aires (`provincia = "Cordoba" | provincia = "Buenos Aires"`);

--> Alternativa:

- Provincias: Córdoba o Buenos Aires (`provincia %in% c("Cordoba", "Buenos Aires")`);
- Inmuebles propios (`situacion_inmueble = "Propiedad del ENA"`);

filter

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

```
base_inmuebles %>%
  select(pais,
         provincia,
         situacion_inmueble)
```

| ## | pais | | | | provincia | situacion_inmueble |
|-------|-----------|--------|----------|-----------------|-------------------|--------------------|
| ## 1 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 2 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 3 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 4 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 5 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 6 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 7 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 8 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 9 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 10 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 11 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 12 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 13 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 14 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 15 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 16 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 17 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 18 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 19 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 20 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 21 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 22 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 23 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 24 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 25 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 26 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 27 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |

filter

```
base_inmuebles %>%  
  select(pais,  
         provincia,  
         situacion_inmueble) %>%  
  filter(provincia %in% c("Cordoba", "Buenos Aires"))
```

| | | pais | provincia | situacion_inmueble |
|-------|-----------|--------------|-------------------|--------------------|
| ## 1 | Argentina | Buenos Aires | Alquilado | |
| ## 2 | Argentina | Cordoba | Propiedad del ENA | |
| ## 3 | Argentina | Buenos Aires | Propiedad del ENA | |
| ## 4 | Argentina | Buenos Aires | Propiedad del ENA | |
| ## 5 | Argentina | Buenos Aires | Propiedad del ENA | |
| ## 6 | Argentina | Buenos Aires | Propiedad del ENA | |
| ## 7 | Argentina | Cordoba | Propiedad del ENA | |
| ## 8 | Argentina | Buenos Aires | Propiedad del ENA | |
| ## 9 | Argentina | Buenos Aires | Propiedad del ENA | |
| ## 10 | Argentina | Cordoba | Propiedad del ENA | |

filter

```
base_inmuebles %>%  
  select(pais,  
         provincia,  
         situacion_inmueble) %>%  
  filter(provincia %in% c("Cordoba", "Buenos Aires")) %>%  
  filter(situacion_inmueble == "Propiedad del ENA")
```

| | pais | provincia | situacion_inmueble |
|------|-----------|--------------|--------------------|
| ## 1 | Argentina | Cordoba | Propiedad del ENA |
| ## 2 | Argentina | Buenos Aires | Propiedad del ENA |
| ## 3 | Argentina | Buenos Aires | Propiedad del ENA |
| ## 4 | Argentina | Buenos Aires | Propiedad del ENA |
| ## 5 | Argentina | Buenos Aires | Propiedad del ENA |
| ## 6 | Argentina | Cordoba | Propiedad del ENA |
| ## 7 | Argentina | Buenos Aires | Propiedad del ENA |
| ## 8 | Argentina | Buenos Aires | Propiedad del ENA |
| ## 9 | Argentina | Cordoba | Propiedad del ENA |

PRÁCTICA

Práctica

- Crear un objeto que contenga las variables **codigo_del_inmueble**, **pais**, **superficie**. Filtrar los inmuebles que se encuentren en argentina y cuya superficie sea menor a 100 mts².

mutate()

Creoa / edita variables (columnas)

mutate()

La función tiene el siguiente esquema:

```
base_de_datos %>%  
  mutate(var_nueva = var_1 + var_2)
```

Caso: Supongamos que quiero crear la variable `direccion`, uniendo la información de la `calle` y el `numero`:

mutate()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

mutate()

```
base_inmuebles %>%
  select(pais, calle, numero)
```

| ## | pais | calle |
|-------|-----------|-------------------------|
| ## 1 | Argentina | Avenida Rivadavia |
| ## 2 | Argentina | Avenida de Mayo |
| ## 3 | Argentina | Avenida Paseo Colon |
| ## 4 | Argentina | Hipolito Yrigoyen |
| ## 5 | Argentina | Carlos Pellegrini |
| ## 6 | Argentina | Carlos Pellegrini |
| ## 7 | Argentina | Avenida Rivadavia |
| ## 8 | Argentina | Humberto Primo |
| ## 9 | Argentina | Hipolito Yrigoyen |
| ## 10 | Argentina | San Martin |
| ## 11 | Argentina | Bernardo de Irigoyen |
| ## 12 | Argentina | Tte. Gral Juan D. Peron |
| ## 13 | Argentina | Bartolome Mitre |
| ## 14 | Argentina | Avda. del Libertador |
| ## 15 | Argentina | Azopardo |
| ## 16 | Argentina | Defensa |
| ## 17 | Argentina | Beauchef |
| ## 18 | Argentina | Zepita |
| ## 19 | Argentina | Virrey Liniers |
| ## 20 | Argentina | Hipólito Yrigoyen |
| ## 21 | Argentina | Zepita |
| ## 22 | Argentina | Defensa |
| ## 23 | Argentina | Perón |
| ## 24 | Argentina | Rosario |
| ## 25 | Argentina | Av. Rivadavia |
| ## 26 | Argentina | Av. Alvarez Jonte |
| ## 27 | Argentina | Av. Cabildo |

mutate()

```
base_inmuebles %>%  
  select(pais, calle, numero) %>%  
  filter(pais == "Argentina")
```

| ## | pais | calle |
|-------|-----------|-------------------------|
| ## 1 | Argentina | Avenida Rivadavia |
| ## 2 | Argentina | Avenida de Mayo |
| ## 3 | Argentina | Avenida Paseo Colon |
| ## 4 | Argentina | Hipolito Yrigoyen |
| ## 5 | Argentina | Carlos Pellegrini |
| ## 6 | Argentina | Carlos Pellegrini |
| ## 7 | Argentina | Avenida Rivadavia |
| ## 8 | Argentina | Humberto Primo |
| ## 9 | Argentina | Hipolito Yrigoyen |
| ## 10 | Argentina | San Martin |
| ## 11 | Argentina | Bernardo de Irigoyen |
| ## 12 | Argentina | Tte. Gral Juan D. Peron |
| ## 13 | Argentina | Bartolome Mitre |
| ## 14 | Argentina | Avda. del Libertador |
| ## 15 | Argentina | Azopardo |
| ## 16 | Argentina | Defensa |
| ## 17 | Argentina | Beauchef |
| ## 18 | Argentina | Zepita |
| ## 19 | Argentina | Virrey Liniers |
| ## 20 | Argentina | Hipólito Yrigoyen |
| ## 21 | Argentina | Zepita |
| ## 22 | Argentina | Defensa |
| ## 23 | Argentina | Perón |
| ## 24 | Argentina | Rosario |
| ## 25 | Argentina | Av. Rivadavia |
| ## 26 | Argentina | Av. Alvarez Jonte |
| ## 27 | Argentina | Av. Cabildo |

mutate()

```
base_inmuebles %>%  
  select(pais, calle, numero) %>%  
  filter(pais == "Argentina") %>%  
  mutate(direccion = paste0(calle, " ", numero))
```

| ## | pais | calle |
|-------|-----------|-------------------------|
| ## 1 | Argentina | Avenida Rivadavia |
| ## 2 | Argentina | Avenida de Mayo |
| ## 3 | Argentina | Avenida Paseo Colon |
| ## 4 | Argentina | Hipolito Yrigoyen |
| ## 5 | Argentina | Carlos Pellegrini |
| ## 6 | Argentina | Carlos Pellegrini |
| ## 7 | Argentina | Avenida Rivadavia |
| ## 8 | Argentina | Humberto Primo |
| ## 9 | Argentina | Hipolito Yrigoyen |
| ## 10 | Argentina | San Martin |
| ## 11 | Argentina | Bernardo de Irigoyen |
| ## 12 | Argentina | Tte. Gral Juan D. Peron |
| ## 13 | Argentina | Bartolome Mitre |
| ## 14 | Argentina | Avda. del Libertador |
| ## 15 | Argentina | Azopardo |
| ## 16 | Argentina | Defensa |
| ## 17 | Argentina | Beauchef |
| ## 18 | Argentina | Zepita |
| ## 19 | Argentina | Virrey Liniers |
| ## 20 | Argentina | Hipólito Yrigoyen |
| ## 21 | Argentina | Zepita |
| ## 22 | Argentina | Defensa |
| ## 23 | Argentina | Perón |
| ## 24 | Argentina | Rosario |
| ## 25 | Argentina | Av. Rivadavia |
| ## 26 | Argentina | Av. Alvarez Jonte |
| ## 27 | Argentina | Av. Cabildo |

mutate() - case_when()

Función complementaria: `case_when()`, utilizada para recodificación de variables:

```
base_de_trabajo <- base_de_trabajo %>%  
  mutate(var_nueva = case_when(  
    -----  
    var_vieja == categoría_vieja_1 ~ "nueva categoria 1",  
    var_vieja == categoría_vieja_2 ~ "nueva categoria 2",  
    var_vieja == categoría_vieja_3 ~ "nueva categoria 3"))
```

Operador de
asignación

Condición a
verificar

Nuevo
valor

Recodificando con mutate() y case_when()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

Recodificando con mutate() y case_when()

```
base_inmuebles %>%  
  select(pais, situacion_inmueble)
```

| | ## | pais | situacion_inmueble |
|-------|-----------|-------------------|--------------------|
| ## 1 | Argentina | Propiedad del ENA | |
| ## 2 | Argentina | Propiedad del ENA | |
| ## 3 | Argentina | Propiedad del ENA | |
| ## 4 | Argentina | Propiedad del ENA | |
| ## 5 | Argentina | Propiedad del ENA | |
| ## 6 | Argentina | Propiedad del ENA | |
| ## 7 | Argentina | Propiedad del ENA | |
| ## 8 | Argentina | Propiedad del ENA | |
| ## 9 | Argentina | Propiedad del ENA | |
| ## 10 | Argentina | Propiedad del ENA | |
| ## 11 | Argentina | Propiedad del ENA | |
| ## 12 | Argentina | Propiedad del ENA | |
| ## 13 | Argentina | Propiedad del ENA | |
| ## 14 | Argentina | Propiedad del ENA | |
| ## 15 | Argentina | Propiedad del ENA | |
| ## 16 | Argentina | Propiedad del ENA | |
| ## 17 | Argentina | Propiedad del ENA | |
| ## 18 | Argentina | Propiedad del ENA | |
| ## 19 | Argentina | Propiedad del ENA | |
| ## 20 | Argentina | Propiedad del ENA | |
| ## 21 | Argentina | Propiedad del ENA | |
| ## 22 | Argentina | Alquilado | |
| ## 23 | Argentina | Alquilado | |
| ## 24 | Argentina | Alquilado | |
| ## 25 | Argentina | Alquilado | |
| ## 26 | Argentina | Alquilado | |
| ## 27 | Argentina | Alquilado | |

Recodificando con mutate() y case_when()

```
base_inmuebles %>%
  select(pais, situacion_inmueble) %>%
  mutate(sit_inmb_2 = case_when(situacion_inmueble == "Propiedad del ENA" ~
                                situacion_inmueble == "Alquilado" ~
```



| ## | | pais | situacion_inmueble | sit_inmb_2 |
|-------|--|-----------|--------------------|------------|
| ## 1 | | Argentina | Propiedad del ENA | Propio |
| ## 2 | | Argentina | Propiedad del ENA | Propio |
| ## 3 | | Argentina | Propiedad del ENA | Propio |
| ## 4 | | Argentina | Propiedad del ENA | Propio |
| ## 5 | | Argentina | Propiedad del ENA | Propio |
| ## 6 | | Argentina | Propiedad del ENA | Propio |
| ## 7 | | Argentina | Propiedad del ENA | Propio |
| ## 8 | | Argentina | Propiedad del ENA | Propio |
| ## 9 | | Argentina | Propiedad del ENA | Propio |
| ## 10 | | Argentina | Propiedad del ENA | Propio |
| ## 11 | | Argentina | Propiedad del ENA | Propio |
| ## 12 | | Argentina | Propiedad del ENA | Propio |
| ## 13 | | Argentina | Propiedad del ENA | Propio |
| ## 14 | | Argentina | Propiedad del ENA | Propio |
| ## 15 | | Argentina | Propiedad del ENA | Propio |
| ## 16 | | Argentina | Propiedad del ENA | Propio |
| ## 17 | | Argentina | Propiedad del ENA | Propio |
| ## 18 | | Argentina | Propiedad del ENA | Propio |
| ## 19 | | Argentina | Propiedad del ENA | Propio |
| ## 20 | | Argentina | Propiedad del ENA | Propio |
| ## 21 | | Argentina | Propiedad del ENA | Propio |
| ## 22 | | Argentina | Alquilado | Alquilado |
| ## 23 | | Argentina | Alquilado | Alquilado |
| ## 24 | | Argentina | Alquilado | Alquilado |
| ## 25 | | Argentina | Alquilado | Alquilado |
| ## 26 | | Argentina | Alquilado | Alquilado |
| ## 27 | | Argentina | Alquilado | Alquilado |

mutate()

Caso: Quiero crear una variable superficie en rangos de mts2:

mutate()

Caso: *Quiero crear una variable superficie en rangos de mts2:*

Breve chequeo de la distribución de la variable numérica:

```
summary(base_inmuebles$superficie_aproximada_m2)
```

| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|----|------|---------|--------|-------|---------|---------|
| ## | 30 | 428 | 702 | 15628 | 1329 | 4189756 |

Recodificando con mutate() y case_when()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

Recodificando con mutate() y case_when()

```
base_inmuebles %>%
```

```
  select(pais, superficie_aproximada_m2)
```

| ## | pais | superficie_aproximada_m2 |
|-------|-----------|--------------------------|
| ## 1 | Argentina | 946.18 |
| ## 2 | Argentina | 1361.25 |
| ## 3 | Argentina | 2740.15 |
| ## 4 | Argentina | 1461.43 |
| ## 5 | Argentina | 1137.77 |
| ## 6 | Argentina | 455.53 |
| ## 7 | Argentina | 841.70 |
| ## 8 | Argentina | 3295.14 |
| ## 9 | Argentina | 11165.97 |
| ## 10 | Argentina | 363.72 |
| ## 11 | Argentina | 1067.99 |
| ## 12 | Argentina | 761.85 |
| ## 13 | Argentina | 368.39 |
| ## 14 | Argentina | 473.79 |
| ## 15 | Argentina | 4399.31 |
| ## 16 | Argentina | 1151.19 |
| ## 17 | Argentina | 479.61 |
| ## 18 | Argentina | 19485.12 |
| ## 19 | Argentina | 1992.53 |
| ## 20 | Argentina | 594.02 |
| ## 21 | Argentina | 765.05 |
| ## 22 | Argentina | 442.18 |
| ## 23 | Argentina | 339.94 |
| ## 24 | Argentina | 725.51 |
| ## 25 | Argentina | 188.09 |
| ## 26 | Argentina | 500.44 |
| ## 27 | Argentina | 1027.84 |

Recodificando con mutate() y case_when()

```
base_inmuebles %>%  
  select(pais, superficie_aproximada_m2) %>%  
  mutate(sup_rango = case_when(superficie_aproximada_m2 %in% c(0:1000) ~ "Menor a 1000",  
                                superficie_aproximada_m2 %in% c(2000:3000) ~ "Mayor a 1400",  
                                superficie_aproximada_m2 %in% c(4000:5000) ~ "Mayor a 1400",  
                                superficie_aproximada_m2 %in% c(6000:7000) ~ "Mayor a 1400",  
                                superficie_aproximada_m2 > 8000 ~ "Mayor a 1400"))
```

| | pais | superficie_aproximada_m2 | sup_rango |
|-------|-----------|--------------------------|--------------|
| ## | | | |
| ## 1 | Argentina | 946.18 | Mayor a 1400 |
| ## 2 | Argentina | 1361.25 | Mayor a 1400 |
| ## 3 | Argentina | 2740.15 | Mayor a 1400 |
| ## 4 | Argentina | 1461.43 | Mayor a 1400 |
| ## 5 | Argentina | 1137.77 | Mayor a 1400 |
| ## 6 | Argentina | 455.53 | <NA> |
| ## 7 | Argentina | 841.70 | Mayor a 1400 |
| ## 8 | Argentina | 3295.14 | Mayor a 1400 |
| ## 9 | Argentina | 11165.97 | Mayor a 1400 |
| ## 10 | Argentina | 363.72 | <NA> |
| ## 11 | Argentina | 1067.99 | Mayor a 1400 |
| ## 12 | Argentina | 761.85 | <NA> |
| ## 13 | Argentina | 368.39 | <NA> |
| ## 14 | Argentina | 473.79 | <NA> |
| ## 15 | Argentina | 4399.31 | Mayor a 1400 |
| ## 16 | Argentina | 1151.19 | Mayor a 1400 |
| ## 17 | Argentina | 479.61 | <NA> |
| ## 18 | Argentina | 19485.12 | Mayor a 1400 |
| ## 19 | Argentina | 1992.53 | Mayor a 1400 |
| ## 20 | Argentina | 594.02 | <NA> |
| ## 21 | Argentina | 765.05 | <NA> |
| ## 22 | Argentina | 442.18 | <NA> |
| ## 23 | Argentina | 339.94 | <NA> |
| ## 24 | Argentina | 725.51 | <NA> |
| ## 25 | Argentina | 188.09 | <NA> |
| ## 26 | Argentina | 500.44 | <NA> |
| ## 27 | Argentina | 1027.84 | Mayor a 1400 |

Recodificando con mutate() y case_when()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

Recodificando con mutate() y case_when()

```
base_inmuebles %>%
```

```
  select(pais, superficie_aproximada_m2)
```

| ## | pais | superficie_aproximada_m2 |
|-------|-----------|--------------------------|
| ## 1 | Argentina | 946.18 |
| ## 2 | Argentina | 1361.25 |
| ## 3 | Argentina | 2740.15 |
| ## 4 | Argentina | 1461.43 |
| ## 5 | Argentina | 1137.77 |
| ## 6 | Argentina | 455.53 |
| ## 7 | Argentina | 841.70 |
| ## 8 | Argentina | 3295.14 |
| ## 9 | Argentina | 11165.97 |
| ## 10 | Argentina | 363.72 |
| ## 11 | Argentina | 1067.99 |
| ## 12 | Argentina | 761.85 |
| ## 13 | Argentina | 368.39 |
| ## 14 | Argentina | 473.79 |
| ## 15 | Argentina | 4399.31 |
| ## 16 | Argentina | 1151.19 |
| ## 17 | Argentina | 479.61 |
| ## 18 | Argentina | 19485.12 |
| ## 19 | Argentina | 1992.53 |
| ## 20 | Argentina | 594.02 |
| ## 21 | Argentina | 765.05 |
| ## 22 | Argentina | 442.18 |
| ## 23 | Argentina | 339.94 |
| ## 24 | Argentina | 725.51 |
| ## 25 | Argentina | 188.09 |
| ## 26 | Argentina | 500.44 |
| ## 27 | Argentina | 1027.84 |

Recodificando con mutate() y case_when()

```
base_inmuebles %>%  
  select(pais, superficie_aproximada_m2) %>%  
  mutate(superficie_aproximada_m2 = round(superficie_aproximada_m2,
```

| | ## | pais | superficie_aproximada_m2 |
|--|-------|-----------|--------------------------|
| | ## 1 | Argentina | 946 |
| | ## 2 | Argentina | 1361 |
| | ## 3 | Argentina | 2740 |
| | ## 4 | Argentina | 1461 |
| | ## 5 | Argentina | 1138 |
| | ## 6 | Argentina | 456 |
| | ## 7 | Argentina | 842 |
| | ## 8 | Argentina | 3295 |
| | ## 9 | Argentina | 11166 |
| | ## 10 | Argentina | 364 |
| | ## 11 | Argentina | 1068 |
| | ## 12 | Argentina | 762 |
| | ## 13 | Argentina | 368 |
| | ## 14 | Argentina | 474 |
| | ## 15 | Argentina | 4399 |
| | ## 16 | Argentina | 1151 |
| | ## 17 | Argentina | 480 |
| | ## 18 | Argentina | 19485 |
| | ## 19 | Argentina | 1993 |
| | ## 20 | Argentina | 594 |
| | ## 21 | Argentina | 765 |
| | ## 22 | Argentina | 442 |
| | ## 23 | Argentina | 340 |
| | ## 24 | Argentina | 726 |
| | ## 25 | Argentina | 188 |
| | ## 26 | Argentina | 500 |
| | ## 27 | Argentina | 1028 |

Recodificando con mutate() y case_when()

```
base_inmuebles %>%  
  select(pais, superficie_aproximada_m2) %>%  
  mutate(superficie_aproximada_m2 = round(superficie_aproximada_m2))  
  mutate(sup_rango = case_when(  
    superficie_aproximada_m2 %in% c(0:1000) ~ "Menor a 1000",  
    superficie_aproximada_m2 %in% c(2000:3000) ~ "Entre 2000 y 3000",  
    superficie_aproximada_m2 %in% c(4000:6000) ~ "Entre 4000 y 6000",  
    superficie_aproximada_m2 %in% c(6000:10000) ~ "Entre 6000 y 10000",  
    superficie_aproximada_m2 > 8000 ~ "Mayor a 8000")
```

| | pais | superficie_aproximada_m2 | sup_rango |
|-------|-----------|--------------------------|--------------|
| ## | | | |
| ## 1 | Argentina | 946 | Mayor a 1400 |
| ## 2 | Argentina | 1361 | Mayor a 1400 |
| ## 3 | Argentina | 2740 | Mayor a 1400 |
| ## 4 | Argentina | 1461 | Mayor a 1400 |
| ## 5 | Argentina | 1138 | Mayor a 1400 |
| ## 6 | Argentina | 456 | 801 a 1200 |
| ## 7 | Argentina | 842 | Mayor a 1400 |
| ## 8 | Argentina | 3295 | Mayor a 1400 |
| ## 9 | Argentina | 11166 | Mayor a 1400 |
| ## 10 | Argentina | 364 | 401 a 800 |
| ## 11 | Argentina | 1068 | Mayor a 1400 |
| ## 12 | Argentina | 762 | 1201 a 1400 |
| ## 13 | Argentina | 368 | 401 a 800 |
| ## 14 | Argentina | 474 | 801 a 1200 |
| ## 15 | Argentina | 4399 | Mayor a 1400 |
| ## 16 | Argentina | 1151 | Mayor a 1400 |
| ## 17 | Argentina | 480 | 801 a 1200 |
| ## 18 | Argentina | 19485 | Mayor a 1400 |
| ## 19 | Argentina | 1993 | Mayor a 1400 |
| ## 20 | Argentina | 594 | 801 a 1200 |
| ## 21 | Argentina | 765 | 1201 a 1400 |
| ## 22 | Argentina | 442 | 801 a 1200 |
| ## 23 | Argentina | 340 | 401 a 800 |
| ## 24 | Argentina | 726 | 1201 a 1400 |
| ## 25 | Argentina | 188 | 0 a 400 |
| ## 26 | Argentina | 500 | 801 a 1200 |
| ## 27 | Argentina | 1028 | Mayor a 1400 |

PRÁCTICA

Práctica

- Recodificar la variable **provincia**, para que los valores *Ciudad Autónoma de Buenos Aires* pasen a ser *CABA*.
- Crear una nueva variable con la **superficie en mts2** redondeada a 1 dígito

summarise()

Resume la información en una nueva tabla

summarise()

```
base_de_datos %>%  
  summarise(var1_resumen = sum(var1),  
            var2_media   = mean(var2),  
            var2_desvio  = sd(var2),  
            var2_cv      = var2_desvio / var2_media * 100)
```

summarise()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

summarise()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, superficie_aproximada_m2)
```

| ## | codigo_del_inmueble | superficie_aproximada_m2 |
|-------|---------------------|--------------------------|
| ## 1 | 200002040 | 946.18 |
| ## 2 | 200002059 | 1361.25 |
| ## 3 | 200002067 | 2740.15 |
| ## 4 | 200002075 | 1461.43 |
| ## 5 | 200002083 | 1137.77 |
| ## 6 | 200002091 | 455.53 |
| ## 7 | 200002105 | 841.70 |
| ## 8 | 200002113 | 3295.14 |
| ## 9 | 200002121 | 11165.97 |
| ## 10 | 200002131 | 363.72 |
| ## 11 | 200002148 | 1067.99 |
| ## 12 | 200002156 | 761.85 |
| ## 13 | 200002164 | 368.39 |
| ## 14 | 200002172 | 473.79 |
| ## 15 | 200002180 | 4399.31 |
| ## 16 | 200002199 | 1151.19 |
| ## 17 | 200002202 | 479.61 |
| ## 18 | 200002210 | 19485.12 |
| ## 19 | 200002229 | 1992.53 |
| ## 20 | 200016521 | 594.02 |
| ## 21 | 200043650 | 765.05 |
| ## 22 | 200044010 | 442.18 |
| ## 23 | 200044029 | 339.94 |
| ## 24 | 200048261 | 725.51 |
| ## 25 | 200048271 | 188.09 |
| ## 26 | 200048288 | 500.44 |
| ## 27 | 200048296 | 1027.84 |

summarise()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, superficie_aproximada_m2) %>%  
  summarise(min = min(superficie_aproximada_m2, na.rm = TRUE),  
            max = max(superficie_aproximada_m2, na.rm = TRUE),  
            media = mean(superficie_aproximada_m2, na.rm = TRUE),  
            mediana = median(superficie_aproximada_m2, na.rm = TRUE),  
            desvio = sd(superficie_aproximada_m2, na.rm = TRUE),  
            cv = desvio / media * 100)
```

```
##      min      max      media mediana      desvio      cv  
## 1 29.76 4189756 15628.43   702.22 172603.8 1104.422
```

group_by()

Aplica una operación sobre la población de forma segmentada

group_by()

```
base_de_datos %>%  
  group_by(variable_de_corte) #<<
```

group_by()

```
base_inmuebles
```

```
##      codigo_del_inmueble
## 1      200002040
## 2      200002059
## 3      200002067
## 4      200002075
## 5      200002083
## 6      200002091
## 7      200002105
## 8      200002113
## 9      200002121
## 10     200002131
## 11     200002148
## 12     200002156
## 13     200002164
## 14     200002172
## 15     200002180
## 16     200002199
## 17     200002202
## 18     200002210
## 19     200002229
## 20     200016521
## 21     200043650
## 22     200044010
## 23     200044029
## 24     200048261
## 25     200048271
## 26     200048288
## 27     200048296
```

group_by()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, situacion_inmueble, superficie_
```



| ## | codigo_del_inmueble | situacion_inmueble | superficie_aproximada_m2 |
|-------|---------------------|--------------------|--------------------------|
| ## 1 | 200002040 | Propiedad del ENA | 946.18 |
| ## 2 | 200002059 | Propiedad del ENA | 1361.25 |
| ## 3 | 200002067 | Propiedad del ENA | 2740.15 |
| ## 4 | 200002075 | Propiedad del ENA | 1461.43 |
| ## 5 | 200002083 | Propiedad del ENA | 1137.77 |
| ## 6 | 200002091 | Propiedad del ENA | 455.53 |
| ## 7 | 200002105 | Propiedad del ENA | 841.70 |
| ## 8 | 200002113 | Propiedad del ENA | 3295.14 |
| ## 9 | 200002121 | Propiedad del ENA | 11165.97 |
| ## 10 | 200002131 | Propiedad del ENA | 363.72 |
| ## 11 | 200002148 | Propiedad del ENA | 1067.99 |
| ## 12 | 200002156 | Propiedad del ENA | 761.85 |
| ## 13 | 200002164 | Propiedad del ENA | 368.39 |
| ## 14 | 200002172 | Propiedad del ENA | 473.79 |
| ## 15 | 200002180 | Propiedad del ENA | 4399.31 |
| ## 16 | 200002199 | Propiedad del ENA | 1151.19 |
| ## 17 | 200002202 | Propiedad del ENA | 479.61 |
| ## 18 | 200002210 | Propiedad del ENA | 19485.12 |
| ## 19 | 200002229 | Propiedad del ENA | 1992.53 |
| ## 20 | 200016521 | Propiedad del ENA | 594.02 |
| ## 21 | 200043650 | Propiedad del ENA | 765.05 |
| ## 22 | 200044010 | Alquilado | 442.18 |
| ## 23 | 200044029 | Alquilado | 339.94 |
| ## 24 | 200048261 | Alquilado | 725.51 |
| ## 25 | 200048271 | Alquilado | 188.09 |
| ## 26 | 200048288 | Alquilado | 500.44 |
| ## 27 | 200048296 | Alquilado | 1027.84 |

group_by()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, situacion_inmueble, superficie_  
    group_by(situacion_inmueble)
```

```
## # A tibble: 677 x 3  
## # Groups:   situacion_inmueble [2]  
##   codigo_del_inmueble situacion_inmueble superficie_aproximada_m2  
##           <dbl> <chr>           <dbl>  
## 1      200002040 Propiedad del ENA           946.  
## 2      200002059 Propiedad del ENA          1361.  
## 3      200002067 Propiedad del ENA          2740.  
## 4      200002075 Propiedad del ENA          1461.  
## 5      200002083 Propiedad del ENA          1138.  
## 6      200002091 Propiedad del ENA           456.  
## 7      200002105 Propiedad del ENA           842.  
## 8      200002113 Propiedad del ENA          3295.  
## 9      200002121 Propiedad del ENA         11166.  
## 10     200002131 Propiedad del ENA           364.  
## # ... with 667 more rows
```

group_by()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, situacion_inmueble, superficie_  
group_by(situacion_inmueble) %>%  
  summarise(min = min(superficie_aproximada_m2, na.rm = TRUE),  
            max = max(superficie_aproximada_m2, na.rm = TRUE),  
            media = mean(superficie_aproximada_m2, na.rm = TRUE),  
            mediana = median(superficie_aproximada_m2, na.rm = TRUE),  
            desvio = sd(superficie_aproximada_m2, na.rm = TRUE),  
            cv = desvio / media * 100)
```

```
## # A tibble: 2 x 7  
##   situacion_inmueble    min      max  media mediana  desvio    cv  
##   <chr>          <dbl>    <dbl>  <dbl>   <dbl>   <dbl>  <dbl>  
## 1 Alquilado        74.0  185904.  5250.    388.  27555.  525.  
## 2 Propiedad del ENA  29.8 4189756. 16864.    725. 182318. 1081.
```


summarise()

Caso: Queremos conocer la cantidad de inmuebles por provincia y la superficie media, según si el inmueble es propio o alquilado. Al mismo tiempo, queremos eliminar casos outliers, que puedan estar distorcionando los resultados de los promedios.

group_by()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

group_by()

```
base_inmuebles %>%
  select(pais, provincia, situacion_inmueble, superficie_aproximada)
```

| ## | pais | | | | provincia | situacion_inmueble |
|-------|-----------|--------|----------|----|--------------|--------------------|
| ## 1 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 2 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 3 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 4 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 5 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 6 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 7 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 8 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 9 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 10 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 11 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 12 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 13 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 14 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 15 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 16 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 17 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 18 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 19 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 20 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 21 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Propiedad del ENA |
| ## 22 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Alquilado |
| ## 23 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Alquilado |
| ## 24 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Alquilado |
| ## 25 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Alquilado |
| ## 26 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Alquilado |
| ## 27 | Argentina | Ciudad | Autonoma | de | Buenos Aires | Alquilado |

group_by()

```
base_inmuebles %>%
  select(pais, provincia, situacion_inmueble, superficie_aproximada)
  filter(pais == "Argentina")
```

| ## | pais | | | | provincia | situacion_inmueble |
|-------|-----------|--------|----------|-----------------|-------------------|--------------------|
| ## 1 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 2 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 3 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 4 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 5 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 6 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 7 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 8 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 9 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 10 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 11 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 12 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 13 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 14 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 15 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 16 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 17 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 18 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 19 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 20 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 21 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 22 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 23 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 24 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 25 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 26 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 27 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |

group_by()

```
base_inmuebles %>%  
  select(pais, provincia, situacion_inmueble, superficie_aproximada_m2)  
  filter(pais == "Argentina") %>%  
  filter(superficie_aproximada_m2 < 10000)
```

| ## | pais | | | | provincia | situacion_inmueble |
|-------|-----------|--------|----------|-----------------|-------------------|--------------------|
| ## 1 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 2 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 3 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 4 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 5 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 6 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 7 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 8 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 9 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 10 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 11 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 12 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 13 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 14 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 15 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 16 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 17 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 18 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 19 | Argentina | Ciudad | Autonoma | de Buenos Aires | Propiedad del ENA | |
| ## 20 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 21 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 22 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 23 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 24 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 25 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 26 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |
| ## 27 | Argentina | Ciudad | Autonoma | de Buenos Aires | Alquilado | |

group_by()

```
base_inmuebles %>%
  select(pais, provincia, situacion_inmueble, superficie_aproximada_m2)
  filter(pais == "Argentina") %>%
  filter(superficie_aproximada_m2 < 10000) %>%
  mutate(provincia = case_when(provincia == "Ciudad Autonoma de Buenos Aires" ~ "CABA",
                                TRUE ~ provincia),
          situacion_inmueble = case_when(situacion_inmueble == "Propio" ~ "Propio",
                                          situacion_inmueble == "Alquilado" ~ "Alquilado"))
```

| | pais | provincia | situacion_inmueble | superficie_aproximada_m2 |
|-------|-----------|-----------|--------------------|--------------------------|
| ## 1 | Argentina | CABA | Propio | 946.18 |
| ## 2 | Argentina | CABA | Propio | 1361.25 |
| ## 3 | Argentina | CABA | Propio | 2740.15 |
| ## 4 | Argentina | CABA | Propio | 1461.43 |
| ## 5 | Argentina | CABA | Propio | 1137.77 |
| ## 6 | Argentina | CABA | Propio | 455.53 |
| ## 7 | Argentina | CABA | Propio | 841.70 |
| ## 8 | Argentina | CABA | Propio | 3295.14 |
| ## 9 | Argentina | CABA | Propio | 363.72 |
| ## 10 | Argentina | CABA | Propio | 1067.99 |
| ## 11 | Argentina | CABA | Propio | 761.85 |
| ## 12 | Argentina | CABA | Propio | 368.39 |
| ## 13 | Argentina | CABA | Propio | 473.79 |
| ## 14 | Argentina | CABA | Propio | 4399.31 |
| ## 15 | Argentina | CABA | Propio | 1151.19 |
| ## 16 | Argentina | CABA | Propio | 479.61 |
| ## 17 | Argentina | CABA | Propio | 1992.53 |
| ## 18 | Argentina | CABA | Propio | 594.02 |
| ## 19 | Argentina | CABA | Propio | 765.05 |
| ## 20 | Argentina | CABA | Alquilado | 442.18 |
| ## 21 | Argentina | CABA | Alquilado | 339.94 |
| ## 22 | Argentina | CABA | Alquilado | 725.51 |
| ## 23 | Argentina | CABA | Alquilado | 188.09 |
| ## 24 | Argentina | CABA | Alquilado | 500.44 |
| ## 25 | Argentina | CABA | Alquilado | 1027.84 |
| ## 26 | Argentina | CABA | Alquilado | 1475.81 |
| ## 27 | Argentina | CABA | Alquilado | 1035.40 |

group_by()

```
base_inmuebles %>%
  select(pais, provincia, situacion_inmueble, superficie_aproximada_m2)
  filter(pais == "Argentina") %>%
  filter(superficie_aproximada_m2 < 10000) %>%
  mutate(provincia = case_when(provincia == "Ciudad Autonoma de Buenos Aires" ~ "CABA",
                                TRUE ~ provincia),
          situacion_inmueble = case_when(situacion_inmueble == "Propio" ~ "Propio",
                                          situacion_inmueble == "Alquilado" ~ "Alquilado"))
  group_by(provincia, situacion_inmueble)
```



```
## # A tibble: 569 x 4
## # Groups:   provincia, situacion_inmueble [7]
##   pais      provincia situacion_inmueble superficie_aproximada_m2
##   <chr>      <chr>      <chr>                  <dbl>
## 1 Argentina CABA      Propio                  946.
## 2 Argentina CABA      Propio                 1361.
## 3 Argentina CABA      Propio                 2740.
## 4 Argentina CABA      Propio                 1461.
## 5 Argentina CABA      Propio                 1138.
## 6 Argentina CABA      Propio                  456.
## 7 Argentina CABA      Propio                  842.
## 8 Argentina CABA      Propio                 3295.
## 9 Argentina CABA      Propio                  364.
## 10 Argentina CABA      Propio                 1068.
## # ... with 559 more rows
```

group_by()

```
base_inmuebles %>%
  select(pais, provincia, situacion_inmueble, superficie_aproximada_m2)
  filter(pais == "Argentina") %>%
  filter(superficie_aproximada_m2 < 10000) %>%
  mutate(provincia = case_when(provincia == "Ciudad Autonoma de Buenos Aires" ~ "CABA",
                                TRUE ~ provincia),
          situacion_inmueble = case_when(situacion_inmueble == "Propio" ~ "Propio",
                                          situacion_inmueble == "Alquilado" ~ "Alquilado"),
          cantidad = 1)
  group_by(provincia, situacion_inmueble) %>%
  summarise(cantidad = n(),
            media_sup_mts2 = mean(superficie_aproximada_m2, na.rm = TRUE))
```

```
## # A tibble: 7 x 4
## # Groups:   provincia [6]
##   provincia    situacion_inmueble cantidad media_sup_mts2
##   <chr>         <chr>          <int>      <dbl>
## 1 Buenos Aires Propio             2        870.
## 2 CABA          Alquilado         68        670.
## 3 CABA          Propio          494       1128.
## 4 Catamarca    Alquilado         1        386.
## 5 Chaco         Propio            1         33
## 6 Cordoba       Propio            2      4767.
## 7 Rio Negro    Alquilado         1        314.
```


group_by()

```
base_inmuebles %>%
  select(pais, provincia, situacion_inmueble, superficie_aproximada_m2)
  filter(pais == "Argentina") %>%
  filter(superficie_aproximada_m2 < 10000) %>%
  mutate(provincia = case_when(provincia == "Ciudad Autonoma de Buenos Aires" ~ "CABA",
                                TRUE ~ provincia),
          situacion_inmueble = case_when(situacion_inmueble == "Propio" ~ "Propio",
                                          situacion_inmueble == "Alquilado" ~ "Alquilado"),
          cantidad = 1)
  group_by(provincia, situacion_inmueble) %>%
  summarise(cantidad = n(),
            media_sup_mts2 = mean(superficie_aproximada_m2, na.rm=TRUE))
  arrange(-cantidad)
```

```
## # A tibble: 7 x 4
## # Groups:   provincia [6]
##   provincia    situacion_inmueble cantidad media_sup_mts2
##   <chr>         <chr>          <int>      <dbl>
## 1 CABA          Propio             494      1128.
## 2 CABA          Alquilado          68       670.
## 3 Buenos Aires Propio              2       870.
## 4 Cordoba       Propio              2     4767.
## 5 Catamarca     Alquilado           1       386.
## 6 Chaco          Propio              1         33
## 7 Rio Negro     Alquilado           1       314.
```


Funciones del paquete tidyr:

| Función | Acción |
|-----------------------------|--|
| <code>pivot_longer()</code> | <i>Transforma en filas varias columnas</i> |
| <code>pivot_wider()</code> | <i>transforma en columnas varias filas</i> |

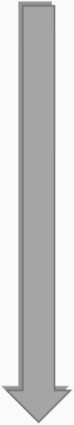
A lo Ancho

| id | provincia | edad_media | edad_desvio | edad_cv |
|----|-----------|------------|-------------|---------|
| 1 | BsAs | 38 | 18 | 47 |
| 2 | CABA | 40 | 18 | 46 |



A lo Largo

| id | provincia | Variable | valor |
|----|-----------|-------------|-------|
| 1 | BsAs | edad_media | 38 |
| 2 | CABA | edad_media | 40 |
| 1 | BsAs | edad_desvio | 18 |
| 2 | CABA | edad_desvio | 18 |
| 1 | BsAs | edad_cv | 47 |
| 2 | CABA | edad_cv | 46 |



pivot_longer()

Reestructura la base, apilando varias columnas en una. De ancho a largo

pivot_longer()

base_inmuebles

| ## | codigo_del_inmueble |
|-------|---------------------|
| ## 1 | 200002040 |
| ## 2 | 200002059 |
| ## 3 | 200002067 |
| ## 4 | 200002075 |
| ## 5 | 200002083 |
| ## 6 | 200002091 |
| ## 7 | 200002105 |
| ## 8 | 200002113 |
| ## 9 | 200002121 |
| ## 10 | 200002131 |
| ## 11 | 200002148 |
| ## 12 | 200002156 |
| ## 13 | 200002164 |
| ## 14 | 200002172 |
| ## 15 | 200002180 |
| ## 16 | 200002199 |
| ## 17 | 200002202 |
| ## 18 | 200002210 |
| ## 19 | 200002229 |
| ## 20 | 200016521 |
| ## 21 | 200043650 |
| ## 22 | 200044010 |
| ## 23 | 200044029 |
| ## 24 | 200048261 |
| ## 25 | 200048271 |
| ## 26 | 200048288 |
| ## 27 | 200048296 |

pivot_longer()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, situacion_inmueble, superficie_
```



| ## | codigo_del_inmueble | situacion_inmueble | superficie_aproximada_m2 |
|-------|---------------------|--------------------|--------------------------|
| ## 1 | 200002040 | Propiedad del ENA | 946.18 |
| ## 2 | 200002059 | Propiedad del ENA | 1361.25 |
| ## 3 | 200002067 | Propiedad del ENA | 2740.15 |
| ## 4 | 200002075 | Propiedad del ENA | 1461.43 |
| ## 5 | 200002083 | Propiedad del ENA | 1137.77 |
| ## 6 | 200002091 | Propiedad del ENA | 455.53 |
| ## 7 | 200002105 | Propiedad del ENA | 841.70 |
| ## 8 | 200002113 | Propiedad del ENA | 3295.14 |
| ## 9 | 200002121 | Propiedad del ENA | 11165.97 |
| ## 10 | 200002131 | Propiedad del ENA | 363.72 |
| ## 11 | 200002148 | Propiedad del ENA | 1067.99 |
| ## 12 | 200002156 | Propiedad del ENA | 761.85 |
| ## 13 | 200002164 | Propiedad del ENA | 368.39 |
| ## 14 | 200002172 | Propiedad del ENA | 473.79 |
| ## 15 | 200002180 | Propiedad del ENA | 4399.31 |
| ## 16 | 200002199 | Propiedad del ENA | 1151.19 |
| ## 17 | 200002202 | Propiedad del ENA | 479.61 |
| ## 18 | 200002210 | Propiedad del ENA | 19485.12 |
| ## 19 | 200002229 | Propiedad del ENA | 1992.53 |
| ## 20 | 200016521 | Propiedad del ENA | 594.02 |
| ## 21 | 200043650 | Propiedad del ENA | 765.05 |
| ## 22 | 200044010 | Alquilado | 442.18 |
| ## 23 | 200044029 | Alquilado | 339.94 |
| ## 24 | 200048261 | Alquilado | 725.51 |
| ## 25 | 200048271 | Alquilado | 188.09 |
| ## 26 | 200048288 | Alquilado | 500.44 |
| ## 27 | 200048296 | Alquilado | 1027.84 |

pivot_longer()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, situacion_inmueble, superficie_  
    group_by(situacion_inmueble)
```

```
## # A tibble: 677 x 3  
## # Groups:   situacion_inmueble [2]  
##   codigo_del_inmueble situacion_inmueble superficie_aproximada_m2  
##           <dbl> <chr>           <dbl>  
## 1      200002040 Propiedad del ENA           946.  
## 2      200002059 Propiedad del ENA          1361.  
## 3      200002067 Propiedad del ENA          2740.  
## 4      200002075 Propiedad del ENA          1461.  
## 5      200002083 Propiedad del ENA          1138.  
## 6      200002091 Propiedad del ENA           456.  
## 7      200002105 Propiedad del ENA           842.  
## 8      200002113 Propiedad del ENA          3295.  
## 9      200002121 Propiedad del ENA         11166.  
## 10     200002131 Propiedad del ENA           364.  
## # ... with 667 more rows
```


pivot_longer()

```
base_inmuebles %>%  
  select(codigo_del_inmueble, situacion_inmueble, superficie_  
group_by(situacion_inmueble) %>%  
  summarise(min = min(superficie_aproximada_m2, na.rm = TRUE),  
            max = max(superficie_aproximada_m2, na.rm = TRUE),  
            media = mean(superficie_aproximada_m2, na.rm = TRUE),  
            mediana = median(superficie_aproximada_m2, na.rm = TRUE),  
            desvio = sd(superficie_aproximada_m2, na.rm = TRUE),  
            cv = desvio / media * 100)
```

```
## # A tibble: 2 x 7  
##   situacion_inmueble   min     max  media mediana  desvio    cv  
##   <chr>             <dbl>   <dbl> <dbl>   <dbl>   <dbl> <dbl>  
## 1 Alquilado         74.0 185904. 5250.    388.  27555.  525.  
## 2 Propiedad del ENA  29.8 4189756. 16864.    725. 182318. 1081.
```

pivot_longer()

```
base_inmuebles %>%
  select(codigo_del_inmueble, situacion_inmueble, superficie_
  group_by(situacion_inmueble) %>%
  summarise(min = min(superficie_aproximada_m2, na.rm = TRUE),
    max = max(superficie_aproximada_m2, na.rm = TRUE),
    media = mean(superficie_aproximada_m2, na.rm = TRUE),
    mediana = median(superficie_aproximada_m2, na.rm = TRUE),
    desvio = sd(superficie_aproximada_m2, na.rm = TRUE),
    cv = desvio / media * 100) %>%
  select(situacion_inmueble, media, mediana, desvio)
```

```
## # A tibble: 2 x 4
##   situacion_inmueble  media mediana  desvio
##   <chr>              <dbl>   <dbl>   <dbl>
## 1 Alquilado          5250.    388.  27555.
## 2 Propiedad del ENA  16864.    725. 182318.
```

pivot_longer()

```
base_inmuebles %>%
  select(codigo_del_inmueble, situacion_inmueble, superficie_
  group_by(situacion_inmueble) %>%
  summarise(min = min(superficie_aproximada_m2, na.rm = TRUE),
    max = max(superficie_aproximada_m2, na.rm = TRUE),
    media = mean(superficie_aproximada_m2, na.rm = TRUE),
    mediana = median(superficie_aproximada_m2, na.rm = TRUE),
    desvio = sd(superficie_aproximada_m2, na.rm = TRUE),
    cv = desvio / media * 100) %>%
  select(situacion_inmueble, media, mediana, desvio) %>%
  pivot_longer(cols = c(media, mediana, desvio), #<<
    names_to = "variable",
    values_to = "valor")
```

```
## # A tibble: 6 x 3
##   situacion_inmueble variable   valor
##   <chr>              <chr>     <dbl>
## 1 Alquilado          media      5250.
## 2 Alquilado          mediana     388.
## 3 Alquilado          desvio    27555.
## 4 Propiedad del ENA  media     16864.
## 5 Propiedad del ENA  mediana     725.
## 6 Propiedad del ENA  desvio   182318.
```

pivot_wider()

Reestructura la base, encolumnando varias filas de una variable. De largo a ancho

pivot_wider()

```
base_largo
```

```
## # A tibble: 6 x 3
##   situacion_inmueble variable  valor
##   <chr>              <chr>    <dbl>
## 1 Alquilado          media    5250.
## 2 Alquilado          mediana    388.
## 3 Alquilado          desvio   27555.
## 4 Propiedad del ENA  media    16864.
## 5 Propiedad del ENA  mediana    725.
## 6 Propiedad del ENA  desvio  182318.
```

pivot_wider()

```
base_largo %>%  
  pivot_wider(names_from = "variable", #<<  
              values_from = "valor")
```

```
## # A tibble: 2 x 4  
##   situacion_inmueble  media mediana  desvio  
##   <chr>              <dbl>   <dbl>   <dbl>  
## 1 Alquilado          5250.    388.  27555.  
## 2 Propiedad del ENA  16864.    725. 182318.
```