

Laboratorio 2 - Niveles, Bases de datos, .qmd

Métodos Cuantitativos

Carlos Eduardo Molina Berumen

El Colegio de México

24 de octubre de 2025



Repaso

Variables: - texto - numéricas - lógicas

¿Qué hace el siguiente código?

```
library(fivethirtyeight)  
  
trump <- trumpworld_issues
```

¿Cómo saber qué tipo de variable es una de las variables de mi bdd?

```
class(trump$approve)  
  
[1] "integer"
```

Creemos una nueva variable.

Si en un país la aprobación sobre un issue de trump es mayor a 0, hagamos una variable lógica.

```
trump$liked <- (trump$net_approval > 0)
```

Hágamos más cosas con las bases de datos. ¿Qué nos dice esta tabla?

```
library(dplyr)  
count(trump, liked)
```

```
# A tibble: 2 x 2  
  liked     n  
  <lgl> <int>  
1 FALSE    181  
2 TRUE      4
```

dplyr: gramática nueva para R

Imaginemos que queremos saber cual es la aprobación neta media del issue 2 según mi base de datos. Usaríamos algo así:

```
wall_approval <- mean(trump$net_approval[trump$issue == "1"])
```

```
[1] -51.27027
```

Podríamos hacerlo más sencillo utilizando pipas:

- %>% (command + shift + m)

Una pipa (%>%) sirve para encadenar pasos sin repetir el nombre de tu base de datos.

Es como decirle a R:

“Toma esto, y luego hazle esto otro, y luego esto otro...”

Así lees el código como una receta en orden, en vez de escribir la base una y otra vez.

Creemos un nuevo objeto del issue 1

```
approval_issue <- trump %>%
  group_by(issue) %>%
  summarise(media_aprobacion = mean(net_approval, na.rm = TRUE))
```

Recordemos

```
x <- c(1, 2, 2, 3, 3, 3, 4)  
mean()
```

```
median()
```

```
library(modeest)  
mlv(x, method = "mfv")
```

```
var()
```

Niveles

```
aprobacion <- c("Mucho", "Poco", "Nada")
```

```
aprobacion <- as.factor(aprobacion)  
aprobacion
```

```
[1] Mucho Poco  Nada  
Levels: Mucho Nada Poco
```

```
aprobacion1 <- factor(aprobacion,  
                      levels = c("Mucho", "Poco", "Nada"))
```