

Laboratorio 28-MD

Leislle R. Manjarrez O.

2023-02-23

Laboratorio 28- Tidy data- Valores faltantes

Objetivo: Introducción práctica a los datos ordenados (o tidy data) y a las herramientas que provee el paquete tidy

En este ejercicio vamos a: 1. Cargar datos (tibbles) 2. Llenar los valores faltantes

Un valor faltante explícito es la presencia de una ausencia Un valor faltante implícito es la ausencia de una presencia

Pre-requisitos: instalar paquetería tidyverse `install.packages("tidyverse")`

Llamar a la librería de tidyverse

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.0      ✓ readr      2.1.4
## ✓ forcats    1.0.0      ✓ stringr   1.5.0
## ✓ ggplot2    3.4.1      ✓ tibble     3.1.8
## ✓ lubridate  1.9.2      ✓ tidyr      1.3.0
## ✓ purrr      1.0.1
## — Conflicts — tidyverse_conflicts() —
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()     masks stats::lag()
## i Use the [8];http://conflicted.r-lib.org/conflicted package[8]; to force all conflicts to become errors
```

Cargar datos

```
acciones <- tibble(
  anio = c(2015, 2015, 2015, 2015, 2016, 2016, 2016),
  trimestre = c(1, 2, 3, 4, 2, 3, 4),
  retorno = c(1.88, 0.59, 0.35, NA, 0.92, 0.17, 2.66))
```

Dataframe1 (DF1) Convertir tabla a dataframe y ver resultado

```
DF1 = acciones
```

```
head(DF1)
```

```
## # A tibble: 6 × 3
##   anio trimestre retorno
##   <dbl>     <dbl>   <dbl>
## 1  2015         1     1.88
## 2  2015         2     0.59
## 3  2015         3     0.35
## 4  2015         4      NA
## 5  2016         2     0.92
## 6  2016         3     0.17
```

Dos valores faltantes, un trimestre (explícito) y el NA (implícito)

Dataframe2 (DF2) Llenar los datos faltantes y ver resultado

```
DF2 = DF1 %>%
  spread(anio, retorno)
```

```
head(DF2)
```

```
## # A tibble: 4 × 3
##   trimestre `2015` `2016`
##   <dbl>   <dbl>   <dbl>
## 1         1   1.88    NA
## 2         2   0.59   0.92
## 3         3   0.35   0.17
## 4         4    NA    2.66
```

Dataframe3 (DF3) Eliminar los datos faltantes y ver resultado

```
DF3 = acciones %>%
  pivot_wider(names_from = anio, values_from = retorno) %>%
  pivot_longer(
    cols = c("2015", "2016"),
    names_to = "anio",
    values_to = "retorno",
    values_drop_na = TRUE)

head(DF3)
```

```
## # A tibble: 6 × 3
##   trimestre anio  retorno
##   <dbl> <chr>    <dbl>
## 1       1 2015      1.88
## 2       2 2015      0.59
## 3       2 2016      0.92
## 4       3 2015      0.35
## 5       3 2016      0.17
## 6       4 2016      2.66
```

Dataframe4 (DF4) Otra herramienta para hacer explicitos los valores faltante (conservar la estructura original) y ver resultado

```
DF4 = DF1 %>%
  complete(anio, trimestre)

head(DF4)
```

```
## # A tibble: 6 × 3
##   anio trimestre retorno
##   <dbl>    <dbl>    <dbl>
## 1 2015         1      1.88
## 2 2015         2      0.59
## 3 2015         3      0.35
## 4 2015         4      NA
## 5 2016         1      NA
## 6 2016         2      0.92
```