

Tarea

Pregunta 1

Describe cómo están organizadas las variables y observaciones en cada una de las cinco familias de tablas `table1`,... `table5` de `tidyverse`.

Pregunta 2

Calcula la columna de `rate` para `table2` y para la combinación de `table4a` y `table4b` sin usar las funciones `gather` o `spread`.

Pregunta 3

Recrea el plot que hemos hecho en la clase para mostrar los casos de infección a lo largo de los años usando la `table2` en lugar de la `table1`. ¿En qué difiere el proceso?

Pregunta 4

Las funciones `spread` y `gather` no son absolutamente simétricas. Toma el siguiente ejemplo para explicarlo correctamente:

```
roi <- tibble(  
  year = c(rep(2016,4), rep(2017,4), 2018),  
  quarter = c(rep(c(1,2,3,4),2),1),  
  return = rnorm(9, mean = 0.5, sd = 1)  
)  
  
roi %>%  
  spread(year, return) %>%  
  gather("year", "return", `2016`:`2018`)
```

Pregunta 5

Las funciones de `spread` y `gather` comparten un argumento `convert`. Investiga su uso.

Pregunta 6

Sin ejecutar, investiga por qué falla el siguiente código

```
table4a %>% gather(1999,2000, key = "year", value = "cases")
```

Pregunta 7

Explica por qué falla la función `spread` aplicada a la siguiente tribble:

```
people <- tribble(
  ~name,      ~key,    ~value,
  #-----/-----/-----
  "Juan Gabriel", "age",    30,
  "Juan Gabriel", "weight", 71,
  "Juan Gabriel", "age",    18,
  "Ricardo",      "age",    55,
  "Ricardo",      "age",    75,
)
```

¿Crees que añadiendo alguna nueva columna se soluciona el problema? Di cual.

Pregunta 8

Limpia la siguiente tibble con la función de `spread` o `gather` que creas más útil.

```
pregnancy <- tribble(
  ~pregnant, ~male, ~female,
  #-----/-----/-----
  "yes",      NA,    32,
  "no",       85,    43
)
```

Indica cuales son las nuevas variables después de aplicar la función.

Pregunta 9

Investiga los parámetros `extra` y `fill` de la función `separate`. Experimenta con varias opciones de las mismas con las dos tibbles siguientes:

```
tibble(x = c("a,b,c", "d,e,f,g","h,i,j")) %>%
  separate(x, c("x", "y", "z"))

tibble(x = c("a,b,c", "d,e","f,g,h")) %>%
  separate(x, c("x", "y", "z"))
```

Pregunta 10

`unite` y `separate` tienen un argumento llamado `remove`. ¿Cómo funciona? ¿Se te ocurre cuando lo pondrías a `false`?

Pregunta 11

Compara las funciones `separate` y `unite`. ¿Por qué existen tres variantes de separación (basándonos en posición, separador o por grupos) pero solamente una para unir?

Pregunta 12

Compara el argumento `fill` de `spread` y el de `complete`. Investiga también el argumento `direction` de la función `fill`.