

Curso 6

Data Frames Wide and Tall

Factores

- *factor()* asigna nombres a códigos numéricos.
- *factor(c(0, 0, 0, 1, 2, 2, 1) , labels = c('Total', 'Hombre', 'Mujer'))*

Cambiar el data.frame usando tidyr

- Otra librería.
- Cambia la forma de un Data Frame.
- Dos formas: Wide y Tall.
 - *spread()* <- extender un valor clave en varias columnas
 - *gather()* <- reúne varias columnas en pares agrupados por un valor.

(Notebook)

Operaciones Agrupadas

- `group_by()` agrupa filas con el mismo valor. No cambia la tabla.
- `summarize()` crea nuevas columnas con el resumen.
 - `n()`
 - `n_distinct()`

(Notebook)

Agrupar usando varias columnas

- `group_by` también agrupa usando varias columnas.
- En este caso agrupa elementos que comparten la misma combinación de características.

(Notebook)

Unir Data Frames

- *inner_join(d1, d2)* contiene solo las filas de d1 comunes a d2.
- *left_join(d1, d2)* contiene todas las filas de d1 y NA en las filas de d2 que no están en d1.
- *semi_join(d1, d2)* no añade columnas a d1. Contiene las filas de d1 comunes a d2.

d1

Nombre	Instrumento
Pedro	Guitarra
Máximo	Pianola
Widinson	Voz

d2

Nombre	Grupo
Pedro	Si
Alejandro	No
Widinson	Si
Esteban	No

inner_join(d1, d2)

Nombre	Instrumento	Grupo
Pedro	Guitarra	Si
Widinson	Voz	Si

d1

Nombre	Instrumento
Pedro	Guitarra
Máximo	Pianola
Widinson	Voz

d2

Nombre	Grupo
Pedro	Si
Alejandro	No
Widinson	Si
Esteban	No

left_join(d1, d2)

Nombre	Instrumento	Grupo
Pedro	Guitarra	Si
Máximo	Pianola	NA
Widinson	Voz	Si

d1

Nombre	Instrumento
Pedro	Guitarra
Máximo	Pianola
Widinson	Voz

d2

Nombre	Grupo
Pedro	Si
Alejandro	No
Widinson	Si
Esteban	No

semi_join(d1, d2)

Nombre	Instrumento
Pedro	Guitarra
Widinson	Voz