3: Tidy Data with `tidyr`

Suppose we have the following toy data set, named `df`, on tennis players. The first column contains the player's max serve speed and handedness, the second contains their rank in the year 1980, and the third contains their rank in the year 1981.

xvar	Rank1980	Rank1981
100-RH	2	6
110-LH	30	19
99-RH	31	30

Split One Column into Two with `separate()`

df %>% separate(xvar, into = c("Speed", ~~"Handedness"~~

"Handedness"),

✓ name of variable   names of new variables

sep = " - ")

↳ name of separator

=>

Rank1980	Rank1981	Speed	Handedness
2	6	100	RH
⋮	⋮	⋮	⋮

`pivot_longer()` to Gather Multiple Columns

df %>% pivot_longer(c(Rank1980, Rank1981),

columns to pivot

names_to = "Year", ← name of new column

values_to = "Rank") ← name of new column

=>

xvar	Year	Rank
100-RH	Rank1980	2
100-RH	Rank1981	6
110-LH	Rank1980	30
110-LH	Rank1981	19
99-RH	Rank1980	31
99-RH	Rank1981	30

data becomes longer ↓

`unite()` and `pivot_wider()`

→ "opposite" of `separate()`   "opposite" of `pivot_longer()`