

Data reshaping

The `unite` function



Structure of the `unite` function

`unite` combines many columns into one column.

```
name_of_dataset %>%  
  unite(  
    column_united,  
    column_one,  
    column_two  
    sep = "_" # The symbol placed between united columns  
  )
```

- The first input is the name to give the united column
- The subsequent inputs (`column_one` and `column_two`) are the columns you wish to unite into a single column
- The `sep =` input specifies the symbol that will be placed between united columns

Untidy data example

Problem: Untidy data frame stored in `table5`

country	century	year	rate
Afghanistan	19	99	745/19987071
Afghanistan	20	00	2666/20595360
Brazil	19	99	37737/172006362
Brazil	20	00	80488/174504898
China	19	99	212258/1272915272
China	20	00	213766/1280428583

Untidy data example

Goal: Use `unite` to transform `table5` back to this:

country	year	rate
Afghanistan	1999	745/19987071
Afghanistan	2000	2666/20595360
Brazil	1999	37737/172006362
Brazil	2000	80488/174504898
China	1999	212258/1272915272
China	2000	213766/1280428583

Afterwards, you can follow the instructions in `separate` lecture to restore `table1`.

unite schematic

`unite` combines many columns into one column.

country	year	rate
Afghanistan	1999	745 / 19987071
Afghanistan	2000	2666 / 20595360
Brazil	1999	37737 / 172006362
Brazil	2000	80488 / 174504898
China	1999	212258 / 1272915272
China	2000	213766 / 1280428583

country	century	year	rate
Afghanistan	19	99	745 / 19987071
Afghanistan	20	0	2666 / 20595360
Brazil	19	99	37737 / 172006362
Brazil	20	0	80488 / 174504898
China	19	99	212258 / 1272915272
China	20	0	213766 / 1280428583

table6

Source: Figure 12.5 in *R for Data Science* by Garrett Golemund and Hadley Wickham.

unite example

```
table5 %>%  
  unite(  
    new,  
    century,  
    year  
  )
```

country	new	rate
Afghanistan	19_99	745/19987071
Afghanistan	20_00	2666/20595360
Brazil	19_99	37737/172006362
Brazil	20_00	80488/174504898
China	19_99	212258/1272915272
China	20_00	213766/1280428583

unite example

```
table5 %>%  
  unite(  
    new,  
    century,  
    year,  
    sep = ""  
  )
```

country	new	rate
Afghanistan	1999	745/19987071
Afghanistan	2000	2666/20595360
Brazil	1999	37737/172006362
Brazil	2000	80488/174504898
China	1999	212258/1272915272
China	2000	213766/1280428583

Example: tidyr + dplyr functions

```
table5 %>%  
  unite(new, century, year, sep = "") %>%  
  mutate(new = as.integer(new)) %>% # Change data type to integer  
  rename(year = new) # Rename column to year
```

country	year	rate
Afghanistan	1999	745/19987071
Afghanistan	2000	2666/20595360
Brazil	1999	37737/172006362
Brazil	2000	80488/174504898
China	1999	212258/1272915272
China	2000	213766/1280428583

Credits

License

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International

Acknowledgments

Content adapted from *R for Data Science* by Garrett Golemund and Hadley Wickham, [chapter 12](#), made available under the [CC BY-NC-ND 3.0 license](#).