

Data reshaping

The `spread` function



Structure of the `spread` function

`spread` takes rows and turns them into columns.

```
name_of_dataset %>%  
  spread(  
    key = column_one,  
    value = column_three  
  )
```

- The `key =` input is the name of a column of categorical data. The categories will become the names of new columns after running the `spread` function.
- The `value =` input is the name of a column containing values that you want to spread into the new columns you created with the `key =` input.

Untidy data example

Untidy data frame stored in `table2`

Afghanistan	1999	cases	745
Afghanistan	1999	population	19987071
Afghanistan	2000	cases	2666
Afghanistan	2000	population	20595360
Brazil	1999	cases	37737
Brazil	1999	population	172006362
Brazil	2000	cases	80488
Brazil	2000	population	174504898
China	1999	cases	212258
China	1999	population	1272915272
China	2000	cases	213766
China	2000	population	1280428583

Untidy data example

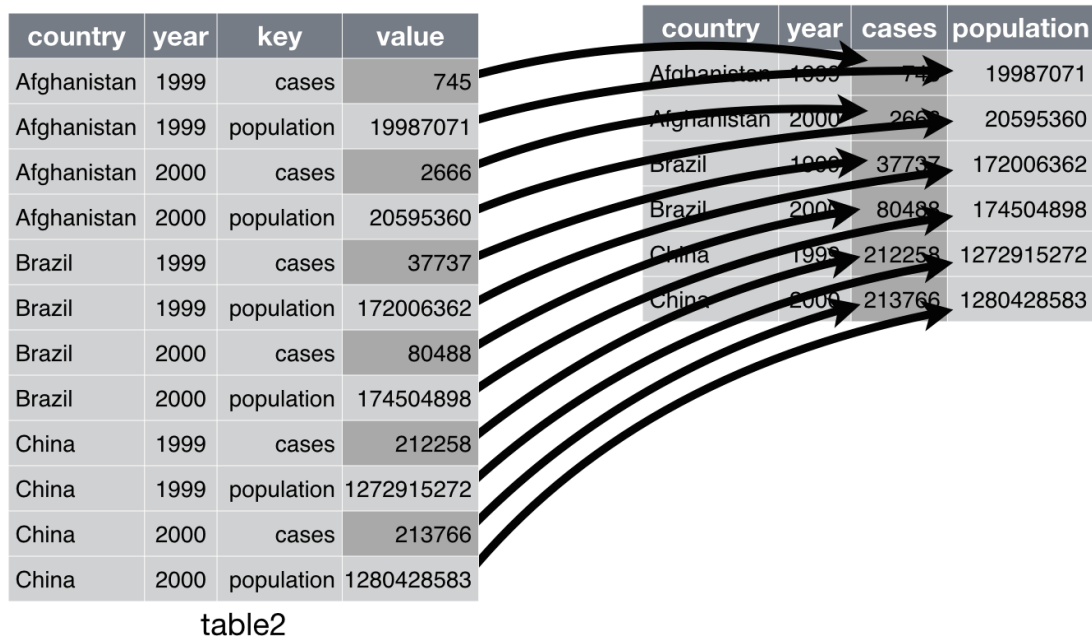
spread

table2

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

spread schematic

spread takes rows and turns them into columns.



Source: Figure 12.3 in

by Garrett Golemund and Hadley Wickham.

spread example

```
table2 %>%  
  spread(  
    key = type,  
    value = count  
  )
```

spread example

```
table2 %>%  
  spread(  
    key = type,  
    value = count  
  )
```

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 Grolmund and Hadley Wickham, [chapter 12](#), made available under the [CC BY-NC-ND 3.0 license](#).