

Exercise: Reshaping

Day 2, Part B

1. Load the Nigeria health metrics data set ('data/nigeria_healthmap.csv').

- Drop the `location_id`, `geography`, `indicator_type`, and `units` columns and keep only rows with data for 2000 and 2010.
- Reshape this data long such that the data in the `estimate`, `ci_lb`, and `ci_ub` columns is combined into one column, e.g.:

	location_name	year	indicator	variable	value
1	Nigeria	2000	Antenatal care, 1 visit (ANC1) (%)	estimate	73
2	Nigeria	2010	Antenatal care, 1 visit (ANC1) (%)	estimate	68
3	Abia	2000	Antenatal care, 1 visit (ANC1) (%)	estimate	94
4	Abia	2010	Antenatal care, 1 visit (ANC1) (%)	estimate	93
5	Adamawa	2000	Antenatal care, 1 visit (ANC1) (%)	estimate	74
6	Adamawa	2010	Antenatal care, 1 visit (ANC1) (%)	estimate	75

- Reshape this data wide so that there are separate columns for the estimate, lower bound, and upper bound in each year, e.g.:

	location_name					
1	Abia					
2	Abia					
3	Abia					
4	Abia					
5	Abia					
6	Abia					
					indicator	
1					Antenatal care, 1 visit (ANC1) (%)	
2					Antenatal care, 4 visits (ANC4) (%)	
3					BCG immunization coverage (%)	
4					Diphtheria-pertussis-tetanus, three doses (DPT3) immunization coverage (%)	
5					Exclusive breastfeeding (EBF) (%)	
6					Household ownership of at least one insecticide-treated net (ITN) (%)	
		estimate_2000	estimate_2010	ci_lb_2000	ci_lb_2010	ci_ub_2000 ci_ub_2010
1		94	93	86	89	97 95
2		92	88	83	82	97 91
3		84	93	64	86	93 97
4		66	80	42	66	84 90
5		7	15	2	8	18 28
6		0	8	0	3	0 22

- Create a new column called `increase` which is `TRUE` if the estimate for 2010 is higher than for 2000 and `FALSE` otherwise.
- Remove all objects in your work space.

Bonus:

2. Still using the original data set, `nigeria_healthmap.csv`:

- Find and read the help docs for `reshape` from the stats library.
- Use `reshape()` rather than `dcast()` to create the output from 1c above, starting with the original nigeria health metrics data.