

SPEC Lab R Resources: Data Management II- Homework

Ben Graham and Alix Ziff

Summer 2020

Manipulating Data using the Tidyverse

Complete the following assignment and then save your R script and merged dataset into the associated google drive folder. The R script should be titled HW_DM2_[YOUR INITIALS]

Exercise 1:

Prep the relative political capacity data:

A Download the relative political capacity data from <https://transresearchconsortium.com/data>. You don't need the components, just the file with the main measures is fine.

B Append gleditsch ward numbers. Check for countries missing gleditsch ward numbers. Thoughtfully address each case of duplicates.

Exercise 2:

Download the following variables for all countries and all years from the World Bank Ed Stats database: Barro and Lee's measures of the number of 15-19 year-olds that have completed secondary school for both the whole population and for females only. <https://databank.worldbank.org/reports.aspx?source=Education%20Statistics>

A Reshape the data to country year, with concise, informative variable names

B Append gleditsch ward numbers. Check for countries missing gleditsch ward numbers and for duplicate observations by country and year. Thoughtfully address each case of duplicates.

Exercise 3:

Merge these two datasets together.