SPEC Lab R Resources:

Merging & Reshaping Data

Group Practice

*Alix Ziff & Benjamin A.T. Graham*

*Version: April 2021*

In this assignment, you will continue to work with the IDC powersharing data. In our last assignment, we used data that was already in country-year format. The goal for this assignment is to walk through some of the steps we used to get from individual coder sheets which each covered a single country, to a merged country-year dataset.

You will need the following files:

1. LATVIA\_NJ\_fortraining.csv
2. FINLAND\_JC\_fortraining.csv

**Steps 1-5:** Write a detailed header for your new R script, clear old objects out of your environment, set your working directory, load the appropriate packages (e.g. tidyverse), and then load each spreadsheet.

**Exercise 1**: Pivot the data in each spreadsheet to make it tidy – i.e. variables as columns and years as rows.

**Exercise 2**: Use as.numeric to make the values that should be numeric (like years) actually numeric.

**Exercise 3**: Merge these two spreadsheets together.

**Exercise 4:** Identify a subset of six variables you find most interesting and rename those variables to simple, descriptive names. Remember to keep everything() else! We don’t want to subset the data in this exercise. The IDC powersharing codebook is in the Training Data folder.

**Exercise 5:** Save your merged dataset as a .rds file in your homework submission folder. Also save a copy of your R script in there. Don’t forget to annotate the heck out of your R script for future you!