Data Module

28 March, 2021

# Abstract

Climate change data are vast and multi-faceted. In particular, data on emissions have been used in all sorts of ways to paint a certain narrative in the global conversation on climate. In this module we will help students understand how emissions and their measured impacts in terms of climate-related disasters could impact the trajectory of international cooperation.

In the first part of the module students will learn how to access and retrieve raw data on emissions from two different datasets, and prepare it for analysis using data wrangling techniques and best practices. Through this specific process, students will learn about the timeliness and reliability of data by exploring missing data and how to report on it.

The next section will ask to create several visualizations that look at how emissions can be aggregated differently depending on who you are as a member state and what makes your case look good in climate negotiations. They will explore the difference between CO2 only versus complete GHG emissions inventories, and explore how trade changes the story about emissions in order to exploit the net contributions to emissions at a global scale.

Finally, students may have the chance to explore the differential impacts of climate-related events by looking at disaster mortality data. They will first try to assess which climate hazards are the most deadly and which regions are affected the most. Students will also explore how the frequency of these events have changed over time, and which regions deserve the most urgent attention.

# Problem Set

You are a climate policy analyst working for a non-governmental organization participating in the next round of negotiations at the Conference of Parties. Your manager asks you to visualize the most recent data on GHG emissions and to create a one-page memo that includes the following visualizations:

1. A map of CO2 emissions by country.
2. A map of GHG emissions (minus CO2) by country.
3. A map of net emissions by country.

The one-page memo should synthesize the data you’ve explored to answer some of the following analytical questions:

1. Who is the historically biggest emitter at this conference? And by what standard of measurement?
2. What are some of the data deprivations that might affect your final analysis and assessment?

# Steps to complete the assignment

1. Clone the following repository into your local computer **[They will clone the raw data folder from our GitHub repository]**
2. Open the pre-formatted R markdown file to start your one-page memo.
3. Import your datasets (emissions data from UNFCCC[[1]](#footnote-22) and net contributions data from Global Carbon Project[[2]](#footnote-24)) using readxl package
4. Clean your datasets
   1. Rename the column names in a consistent manner
   2. Recode the data types that’s appropriate for analysis. Example, change year from charactor to factor, same with country names
   3. Convert data from wide to long to prepare for analysis
5. Merge the two datasets based on country names and years using tidyverse package
6. Create maps using ggplot2

# References

Friedlingstein, Pierre, Michael O’Sullivan, Matthew W. Jones, Robbie M. Andrew, Judith Hauck, Are Olsen, Glen P. Peters, et al. “Global Carbon Budget 2020.” *Earth System Science Data* 12, no. 4 (December 11, 2020): 3269–3340. <https://doi.org/ghn75s>.

“Greenhouse Gas Inventory Data - Flexible Queries Non Annex I Countries.” Accessed March 27, 2021. <https://di.unfccc.int/flex_non_annex1>.

1. Pierre Friedlingstein et al., “Global Carbon Budget 2020,” *Earth System Science Data* 12, no. 4 (December 11, 2020): 3269–3340, <https://doi.org/ghn75s>. [↑](#footnote-ref-22)
2. “Greenhouse Gas Inventory Data - Flexible Queries Non Annex I Countries,” accessed March 27, 2021, <https://di.unfccc.int/flex_non_annex1>. [↑](#footnote-ref-24)