

## Detailed Description of the Data Used

This report examines emerging perspectives on the relationship between climate change and various influencing factors. To explore these connections, I selected five datasets and conducted exploratory data analysis (EDA). The chosen datasets aim to provide insights into the present and potential future causes and effects of climate change, as well as the capacity and direction of adaptation strategies.

To analyze factors affecting climate change across different countries and regions, I utilized the following datasets:

### 1. World Regions Dataset (world\_regions)

This dataset consists of a single categorical variable listing 248 countries across different global regions. It specifies each country's continent (e.g., Asia, Europe) and whether it belongs to the Global North or Global South.

From this dataset, I extracted the country names into a separate dataframe to facilitate merging with other datasets. To enhance clarity and usability, I separated the dataset into three distinct variables:

- Country Name
- Region
- Global Classification (Global North/South)

This restructuring allows for a more effective comparison of climate-related variables across different regions.

### 2. Natural Gas Dataset (natural\_gas)

This dataset provides insights into the distribution and trends of various types of natural gas across different countries and regions worldwide. It contains 2,079 observations across 75 variables, including:

- Region (API): A categorical variable representing different regions and countries, with the first observation representing global totals.
- Natural Gas Type: A categorical variable specifying different types of natural gas, including gross natural gas, dry natural gas, vented and flared gas, reinjected natural gas, and carbon dioxide emissions.
- Yearly Data (1949–2021): Numeric variables (double) representing the quantity of each natural gas type per year.

The motivation behind selecting this dataset is to examine how the rising production of natural gas contributes to climate change. By analyzing the variations in natural gas production over time, we can investigate correlations between increasing fossil fuel use and escalating climate risks.

### 3. Primary Energy Dataset (primary\_energy)

This dataset provides a comprehensive overview of primary energy production and consumption across multiple countries over time. It includes 3,465 observations and 42 variables, such as:

- Region (API): A categorical variable listing different countries, with "World" as the top entry.
- Primary Energy Type: A categorical variable categorizing energy production into total production, total consumption, and five specific energy sources: coal, natural gas, petroleum and other liquids, nuclear, and renewables.
- Yearly Data (1980–2019): Numeric (double) variables representing energy production and consumption trends.

To enhance data clarity, I created separate datasets for:

- Total primary energy production (global and country-level) over time
- Total primary energy consumption (global and country-level) over time
- Production trends of each primary energy type over time

These datasets, joined with the world\_regions dataset, help examine the relationship between climate change and global energy production and consumption patterns. Specifically, this analysis aims to explore whether increasing energy use is necessary for climate adaptation or if dependence on fossil fuels exacerbates climate risks.

### 4. Emissions Dataset (emissions)

To deepen the analysis of how climate change correlates with natural factors, I extracted relevant data from the emissions dataset, which contains 231 observations across 73 variables, including:

- Region: A categorical variable listing different countries.
- Emission Type: A categorical variable capturing various emissions, from which I specifically extracted CO<sub>2</sub> emissions.
- Yearly Data (1949–2019): Numeric (double) variables indicating emissions per year.

This dataset is instrumental in analyzing the role of greenhouse gas emissions in climate change and how emissions trends vary by region and time.

## 5. Social Capital County Dataset (social\_capital\_county)

While the previous datasets focus on natural factors, this dataset explores the social and economic dimensions of climate change adaptation. It includes 3,089 observations across 26 variables, such as:

- State Name & County Name: Categorical variables listing different U.S. states and their corresponding counties.
- County Code: An integer variable uniquely identifying counties.
- Population: An integer variable representing county populations.
- Economic Connectedness (EC): A numeric variable measuring interactions between high- and low-income individuals.
- Socioeconomic Connectedness (SC): A numeric variable indicating the level of socioeconomic integration.
- Children's Economic and Socioeconomic Connectedness: Numeric variables capturing economic and social mobility for children in different income groups.
- Economic and Socioeconomic Connectedness of Communities: Additional numeric variables assessing social cohesion at a community level.
- Income Group Distribution: Numeric variables indicating the proportion of high- and low-income populations.

As climate change increasingly threatens human societies, understanding adaptation strategies and collective action is critical. Adaptation decisions are often made at individual, community, and societal levels, and their effectiveness depends on social cohesion and economic integration.

By analyzing economic connectedness, this dataset helps assess how social networks influence climate adaptation, shedding light on how different communities respond to climate-related challenges.