

Title: Assessing the Role of Individual Consumption Behavior in Climate Change

Author: Zahra Zamanoghli

Date: June 6, 2023

Question

How do the consumption patterns of individuals contribute to carbon emissions across diverse categories of goods and services?

Data Sources

Chosen Data Sources

Datasource1: Eurostat - Consumption footprint

- Metadata URL: https://ec.europa.eu/eurostat/cache/metadata/en/cei_gsr010_esmsip2.htm
- Data URL: https://ec.europa.eu/eurostat/api/dissemination/sdmx/2.1/data/cei_gsr010?format=SDMX-CSV&compressed=false
- Data Type: CSV
- License: Eurostat Copyright/Licence Policy (Can be found in the 8. Comment section of the metadata. Based on the license terms, it is acceptable for individuals to use this dataset for personal or educational endeavors, provided we acknowledge the source appropriately.)
- **Reason for Use:** This dataset meets the need for detailed insights into individual consumption patterns, including household expenditures, consumer behavior, and lifestyle choices related to various goods and services.

The indicator consumption footprint estimates the environmental impacts of EU and Member States consumption by combining data on consumption intensity and environmental impacts of representative products. The indicator covers five areas of consumption: food, mobility, housing, appliances, and household goods. Consumption intensities are calculated based on consumption statistics.

Sample Data Entry:

```
DATAFLOW, LAST UPDATE, freq, cons_fp, unit, geo, TIME_PERIOD, OBS_VALUE, OBS_FLAG  
ESTAT:CEI_GSR010(1.0), 25/03/24 23:00:00, A, AC, I10, AT, 2010, 100,
```

Datasource2: Eurostat - EU CO2 emissions from the production and consumption (footprint) perspectives (FIGARO application)

- Metadata URL: https://ec.europa.eu/eurostat/cache/metadata/en/env_ac_co2fp_esms.htm
- Data URL: https://ec.europa.eu/eurostat/api/dissemination/sdmx/2.1/data/egi_co2_1?format=SDMX-CSV&compressed=false
- Data Type: CSV
- License: CC BY-NC-ND 4.0 (Can be found in the 8.3. Release policy - user access section of the metadata. Based on the license terms, individuals are permitted to use this dataset for personal or educational purposes, provided that proper attribution is given to the source.)
- **Reason for Use:** This dataset meets the need for quantifying the environmental impact and carbon footprint of various consumer activities, including greenhouse gas emissions associated with different products and services.

The dataset presents modelling estimates of carbon dioxide (CO2) 'embodied' in products (goods and services) for final demand - also referred to as 'footprints'. The estimates are the result of environmental input-output modelling and cover the entire world economy.

Sample Data Entry:

```
DATAFLOW, LAST_UPDATE, freq, indic_env, unit, geo, TIME_PERIOD, OBS_VALUE, OBS_FLAG  
ESTAT:EGI_CO2_1(1.0), 05/04/24 23:00:00, A, CO2_CONS_FP, THS_T, EU27_2020, 2010, 4209578.474,
```

Data Pipeline

Overview:

- Download Data: The pipeline begins by downloading CSV files from specified URLs.
- Transform Data: The downloaded data is then read into a pandas DataFrame, where certain columns are dropped to clean and transform the data.
- Save Data: Finally, the transformed data is saved into an SQLite database.

Technologies Used:

- Python: Main programming language used.
- pandas: Library for data manipulation and analysis.
- requests: Library to handle HTTP requests for downloading data.
- SQLite: Database to store the transformed data.
- OS and URL libraries: For handling file paths and URLs.

Transformation Steps:

Dropping Columns: The columns "DATAFLOW", "LAST_UPDATE", "freq" and "OBS_FLAG" are dropped from the DataFrame.

Reason for Transformation:

The rationale behind dropping the first three columns is that they contain identical values across all data entries, rendering them redundant for analysis. The last column was excluded because it was entirely devoid of data, offering no meaningful contribution to the dataset.

Result and Limitations:

Output Data:

The data is stored in an SQLite database, with each dataset being a table in the database. The reason is that SQLite databases are general-purpose formats and are easy to work with.

Limitations

Incompatibility for Merging: Despite initial intentions to merge the datasets, their structures posed challenges for seamless integration.

Data Structure and Quality

- **Accuracy:** Both the Consumption Footprint and Carbon Dioxide Emission Footprints datasets from Eurostat are accurate, offering reliable estimates for understanding environmental impacts and carbon footprints at both the EU and global levels.
- **Completeness:** Both datasets contain all necessary information to understand and analyze the respective environmental indicators.
- **Consistency:** Both datasets are consistent in their format.
- **Timeliness:** The datasets are updated regularly and disseminated annually, ensuring that the age of the data remains appropriate.
- **Relevancy:** The datasets are updated regularly, providing current information on consumption footprints and carbon dioxide emissions