

Global Carbon Emissions and Impact Analysis Report

Abstract

This report details the analysis of global carbon emissions to understand trends, identify primary contributing sectors, and compare environmental impact across nations on a per-capita basis. The project successfully merged disparate global datasets, cleaned critical geographic and temporal inconsistencies using Python, and modelled the final data in Power BI for an interactive dashboard. The primary challenge of resolving country name mismatches and handling sparse population data was overcome to produce reliable metrics for effective data visualization and actionable insight generation.

Tools Used

Python (Pandas, NumPy)- Data ingestion, cleaning, transformation (unpivoting population data), country name standardization via mapping, and calculation of derived metrics.

Excel- data check and format data.

Power BI Desktop -Final data modelling (unpivoting sectoral data, filtering nulls, setting geographic categories), interactive dashboard creation, and visualization.

Steps Involved in Building the Project

The project required a multi-stage process to integrate and cleanse the complex time-series data from two independent sources: CO_Emissions by Sector, World Population.

1. Data Integration and Cleaning (Python)

- **Initial Merging:** The carbon Emissions data was merged with the Population data based on the Entity (Country Name) and Year columns.
- **Data Standardization:** A custom Python mapping dictionary was applied to standardize country names (e.g., merging "Korea, Rep." with "South Korea") to maximize the number of successful joins between the datasets.
- **Metric Calculation: Two key metrics were calculated:**
 - **Total carbon Emissions**-(sum of all sectoral emissions).
 - **Carbon Emissions per Capita**-(Total Emissions/ Population).
- **Data Filtration:** Rows where Population was Not announced (over 85\% of the original data due to missing years and aggregate entities like "World" or "Europe") were filtered out to ensure the per-capita metric was accurate.

2. Final Data Modelling (Power BI)

- **Geographic Tagging:** The Code (3-letter ISO) and Entity columns were explicitly categorized as Country/Region to enable geographic plotting on the map visual.
- **Dimensional Reshaping (Unpivoting):** The eight specific emissions columns (e.g., Carbon dioxide emissions from buildings) were transformed from a "wide" format to a "tall" format using the Unpivot Other Columns function. This created the necessary Sector and Emissions Value columns for the bar charts.

3. Dashboard Creation

The finalized, clean data model was used to create an interactive dashboard featuring the following elements:

- **Map Visual:** Displays the geographical distribution and magnitude of Emissions per Capita (using the reliable Code column for location).
- **Line Chart:** Shows the trend of Total Emissions over time.
- **Bar Chart:** Highlights the sectoral breakdown of carbon contributions (Electricity & Heat, Transport, Industry, etc.).
- **Slicers and KPIs:** Enabled interactive filtering by Country and Year, along with cards showing global summaries.

Conclusion

The completed dashboard provides a clear, actionable view of global carbon emissions. Key insights derived from the analysis include:

1. **Dominant Sector:** The Electricity and Heat sector remains the single largest source of carbon emissions globally, highlighting it as the most critical area for targeted decarbonization efforts.
2. **Per Capita Disparity:** When normalized by population, Emissions per Capita metrics expose significant differences in carbon intensity, revealing that smaller, industrialized, or resource-rich nations often have a disproportionately high environmental impact compared to large developing nations.
3. **Temporal Trends:** The line charts demonstrate distinct trends across regions, with many developed nations showing a slight decline in Total Emissions since 2010, while other rapidly industrializing nations show sharp increases.

The project successfully transformed complex, messy global data into a high-utility business intelligence asset ready for strategic decision-making.

