

# Equity and Trends in Global CO<sub>2</sub> Emissions per Capita \*

\*For the fulfillment project proposal of AT82.01 Computer Programming for Data Science and Artificial Intelligence course by Dr. Chantri Polprasert

1<sup>st</sup> Thiri Shin Thant  
Department of Data Sciences and Artificial  
Intelligence  
Asian Institute of Technology  
Pathum Thani, Thailand  
st126018@ait.asia

2<sup>nd</sup> Aphisit Jaemyaem  
Department of Data Sciences and Artificial  
Intelligence  
Asian Institute of Technology  
Pathum Thani, Thailand  
st126130@ait.asia

**Abstract—** This project proposal aims to analyze equity and trends in global CO<sub>2</sub> emissions per capita by comparing small but developed countries with large developing countries. Using datasets from the Global Carbon Project (2000–2023) and world population data, we investigate which groups of countries contribute more CO<sub>2</sub> emissions per person and how these patterns relate to population size and development level. Exploratory data analysis, visualizations (bar charts, bubble plots, stacked sector analysis), and optional forecasting methods are applied to identify key insights. The expected outcome is to provide evidence of emission inequality, raise awareness, and support policy-making discussions on climate responsibility.

**Index Terms—** Carbon Emissions, CO<sub>2</sub> per capita, Global Equity, Climate Change, Data Science

## I. INTRODUCTION

### A. Introduction to CO<sub>2</sub> Emissions

Carbon dioxide (CO<sub>2</sub>) is the primary greenhouse gas contributing to global climate change. It originates from multiple sources including fossil fuel combustion, industrial activities, cement production, and flaring. Monitoring CO<sub>2</sub> emissions is critical to understanding environmental impact, guiding climate policy, and achieving global sustainability goals. Per capita CO<sub>2</sub> emissions, which measure emissions relative to population size, provide a fair indicator of how much each individual in a country contributes to global emissions.



Fig. 1: Carbon dioxide emissions to the atmosphere.

### B. Global Perspective on CO<sub>2</sub>

At a global scale, CO<sub>2</sub> emissions reveal disparities between developed and developing countries. Small but wealthy nations often produce disproportionately high per capita CO<sub>2</sub> due to energy-intensive industries and advanced infrastructure needs, while large developing countries generate high total emissions because of their population size, though their per capita values remain comparatively low. This global perspective is essential for climate negotiations,

as it highlights equity issues and identifies which countries bear greater responsibility in emission reductions.

### C. Local/Regional Perspective

In regional contexts such as Asia, the diversity in population size and economic development offers contrasting insights. Countries like Singapore and Qatar, despite small populations, demonstrate high per capita CO<sub>2</sub> emissions, while populous nations such as India, China, and Indonesia contribute massively to total global CO<sub>2</sub> but with lower per capita emissions. Understanding these dynamics helps regional policymakers and international organizations design fair strategies for emission reduction.

### D. Why Do We Do This Project?

The motivation of this project is to highlight inequities in carbon emissions across different types of countries and provide evidence for global awareness. By comparing per capita CO<sub>2</sub> emissions of small developed countries and large developing countries, this study aims to inform both policymakers and the public about responsibility sharing, climate justice, and the need for collaborative solutions in reducing carbon footprints.

## II. PROBLEM STATEMENT

We have been observing the problem of carbon emissions worldwide, and many initiatives such as electric vehicles, waste management, renewable energy, and other green projects have been introduced to reduce it. However, countries across the globe are still emitting different levels of carbon dioxide, and the amounts vary greatly between developing and developed countries.

This project aims to identify which countries are emitting large amounts of carbon and to examine whether carbon emissions are related to a country's development level. By analyzing this information, we will highlight which groups of countries bear greater responsibility for reducing carbon emissions and provide insights into global equity in climate change mitigation.

## III. RELATED WORKS

Several studies have explored carbon dioxide (CO<sub>2</sub>) emissions from different perspectives, including global monitoring, equity analysis, and forecasting future trends. The Global Carbon Project (GCP) regularly publishes comprehensive datasets on fossil CO<sub>2</sub> emissions, highlighting both total and per capita values across countries. These datasets are widely used in climate research to compare

national contributions and assess progress towards emission reduction targets.

Research on equity in emissions has emphasized the gap between developed and developing nations. Raupach et al. (2007) investigated emission responsibilities and highlighted that per capita emissions in developed countries remain disproportionately high compared to those in developing countries, despite overall growth in global emissions. Similarly, studies by Chakravarty et al. (2009) proposed frameworks for sharing emission reduction responsibilities based on per capita indicators.

Other works have focused on forecasting emissions trends. For example, Li et al. (2019) applied time-series methods to predict national CO<sub>2</sub> emissions and found that industrial growth and energy consumption were significant drivers. Machine learning techniques such as Random Forests and Support Vector Regression have also been applied to model and predict CO<sub>2</sub> emissions at both national and regional levels.

In addition, comparative studies between small developed countries (e.g., Singapore, Qatar) and large developing countries (e.g., China, India) have shown contrasting patterns: small wealthy nations exhibit very high per capita emissions, while populous developing countries contribute the largest total emissions but much lower per capita values. These works provide a foundation for our study, which focuses on visualizing and analyzing the equity gap between these two groups of countries.

#### IV. DATASET

##### A. Description

The dataset used in this project is sourced from Kaggle: Global Fossil CO<sub>2</sub> Emissions by Country (2000–2022), provided by thedevastator. It contains detailed annual fossil CO<sub>2</sub> emissions data for more than 200 countries, covering the years 2002 to 2022. The dataset is in CSV format and is suitable for both per capita and total CO<sub>2</sub> analysis.

For this study, we focus on two groups of countries:

- Large-population countries (e.g., China, India, United States, Indonesia, Nigeria, Pakistan, Brazil, Japan, Thailand, Russia)
- Small but developed countries (e.g., Singapore, Qatar, Norway, Maldives, Bhutan, Brunei Darussalam, Kuwait, United Arab Emirates, New Zealand, Switzerland)

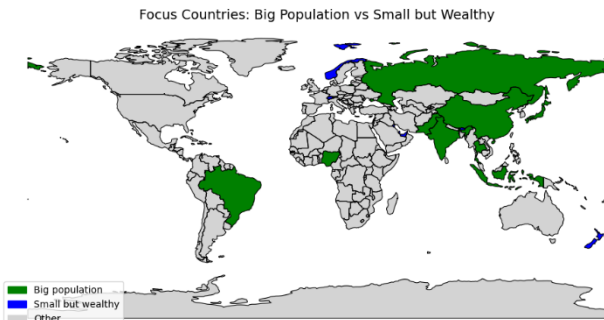


Fig. 2: The focus groups of countries.

This allows a comparative analysis of total vs. per capita CO<sub>2</sub> emissions and provides insights into the responsibility of different countries in contributing to global emissions.

##### B. Features

The dataset includes the following features:

- Country: The name of the country (string)
- ISO 3166-1 alpha-3: The three-letter country code (string)
- Year: The year of the data (integer, 2000–2022)
- Total: The total amount of CO<sub>2</sub> emissions for the country in the given year (float, MtCO<sub>2</sub>)
- Coal: The amount of CO<sub>2</sub> emissions from coal in the given year (float)
- Oil: The amount of CO<sub>2</sub> emissions from oil in the given year (float)
- Gas: The amount of CO<sub>2</sub> emissions from gas in the given year (float)
- Cement: The amount of CO<sub>2</sub> emissions from cement production in the given year (float)
- Flaring: The amount of CO<sub>2</sub> emissions from flaring operations in the given year (float)
- Other: The amount of CO<sub>2</sub> emissions from other sources in the given year (float)

We extend the dataset by combining with world population data to compute per capita CO<sub>2</sub> emissions (tCO<sub>2</sub>/person), which is critical for comparing countries with very different population sizes.

#### V. METHODOLOGY

The methodology of this study is structured into several stages, ensuring a systematic approach to analyzing CO<sub>2</sub> emissions across selected countries.

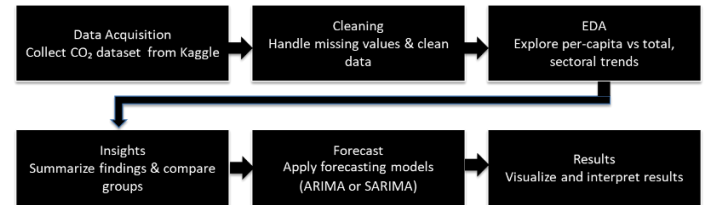


Fig. 3: Methodology workflow.

##### A. Data Acquisition

- The dataset is obtained from the Global Fossil CO<sub>2</sub> Emissions by Country, available on Kaggle [1].
- The dataset contains annual CO<sub>2</sub> emissions by source sectors (coal, oil, gas, cement, flaring) as well as total emissions per country.
- Additional population data is integrated to calculate per-capita CO<sub>2</sub> emissions for fair comparison.

##### B. Data Preprocessing & Cleaning

- Handle missing values by imputation or replacement with zero (where emissions are not reported).
- Standardize country names and codes (ISO Alpha-3) to ensure consistency across datasets.
- Select the study period (2000–2022) to capture both long-term and recent trends.

### C. Exploratory Data Analysis (EDA)

The exploratory data analysis in this project was designed to investigate patterns and trends of CO<sub>2</sub> emissions between 2000 and 2022, focusing on two distinct groups of countries: big population countries (e.g., China, India, Indonesia, United States, Brazil, Nigeria, Pakistan) and small but wealthy countries (e.g., Singapore, Qatar, Norway, Kuwait, UAE, Switzerland, New Zealand).

The analysis first compared total versus per-capita CO<sub>2</sub> emissions. Large population countries generally exhibited high total emissions due to their industrial scale and energy demand, yet their per-capita emissions varied widely. In contrast, small but wealthy countries contributed relatively low total emissions but often displayed disproportionately high per-capita emissions, reflecting their more energy-intensive lifestyles and higher carbon footprints at the individual level.

Sectoral examination further revealed differences in energy dependency. Coal and oil remained dominant sources of emissions in many large population countries, while small but wealthy nations showed higher reliance on oil and gas. These differences indicate distinct structural drivers of emissions across groups.

By combining per-capita, total, and sectoral perspectives, the EDA provides evidence of inequities in global CO<sub>2</sub> emissions. The findings highlight that responsibility for emission reduction cannot be assessed solely by absolute values; instead, both large developing nations and small affluent countries must be considered to ensure fairness and effectiveness in climate policy.

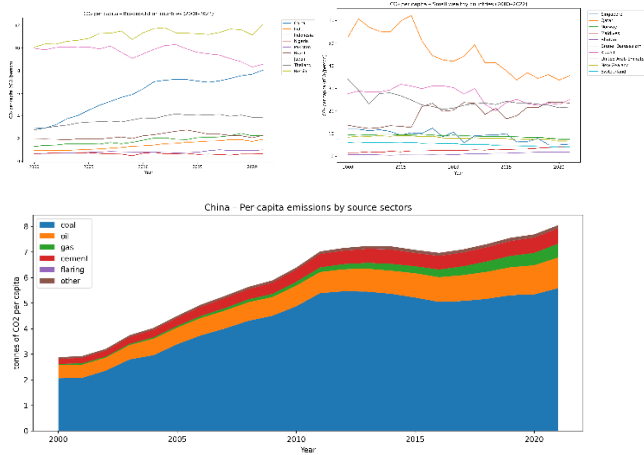


Fig. 4: Exploratory Data Analysis of CO<sub>2</sub> Emissions — Trends, Groups, and Sectoral Breakdown.

### D. Model

To forecast CO<sub>2</sub> emissions beyond 2022, this study applies both a simple baseline model and advanced statistical forecasting models.

The baseline model chosen is the Naïve Forecast, which assumes that the next year's emission will be equal to the most recent observed value. Although very simple, this model provides a standard benchmark; any effective forecasting approach should outperform it [2],[3].

The first advanced method is the ARIMA (Autoregressive Integrated Moving Average) model, which is well-suited for analyzing time series with trends and autocorrelation. It uses past values and past errors to generate forecasts, making it a reliable baseline for long-term emission data.

The second method is the SARIMA (Seasonal ARIMA) model, which extends ARIMA by including seasonal components. This is particularly important when data show repeating patterns or cycles, such as annual fluctuations or structural shifts related to economic activity.

Model performance will be evaluated using common error metrics (e.g., RMSE, MAE) and information criteria (AIC, BIC). By comparing ARIMA and SARIMA with the naïve baseline, the study ensures that improvements in forecasting accuracy are both measurable and meaningful.

### E. Results and Visualization

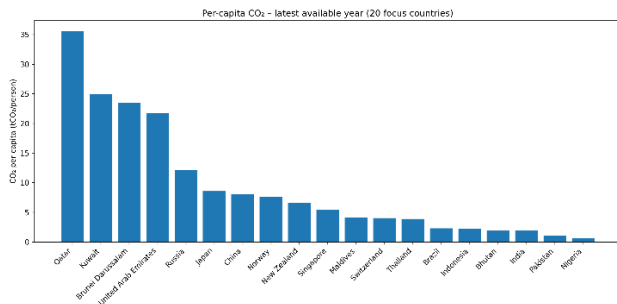
The results of this study will provide a comparative understanding of CO<sub>2</sub> emissions between big population countries and small but wealthy countries, both in terms of total and per-capita values. The analysis will highlight inequities in emissions distribution and identify dominant contributing sectors. The expected outcomes include:

- **Emission Trends:** Time-series plots will demonstrate changes in emissions between 2000 and 2022, showing both total and per-capita perspectives.
- **Group Comparisons:** Bar charts and line charts will compare the two focus groups, illustrating differences in carbon intensity.
- **Sectoral Contributions:** Stacked bar charts will visualize the share of coal, oil, gas, cement, and flaring for selected countries.
- **Geographical Patterns:** Choropleth world maps will highlight spatial differences, marking focus countries in contrasting colors (e.g., green for big population, blue for small wealthy).
- **Forecasting Results:** Forecast plots (ARIMA and SARIMA) will project CO<sub>2</sub> emissions beyond 2022 and compare them against the naïve baseline to evaluate predictive improvements.

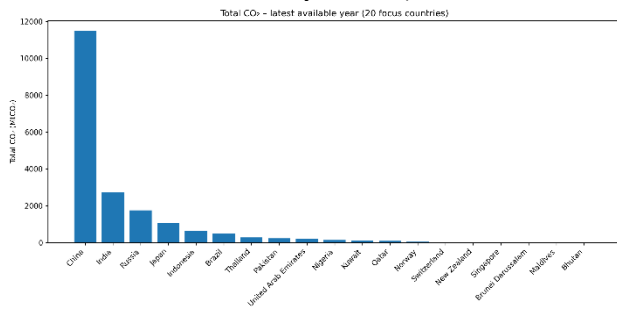
Through these visualizations, the study will provide clear evidence of the imbalance between total and per-capita CO<sub>2</sub> emissions, allowing for a more nuanced discussion of fairness and responsibility in global climate change mitigation [3].

### REFERENCES

- [1] The Devastator, "Global Fossil CO<sub>2</sub> Emissions by Country" Kaggle Dataset, 2023. [Online]. Available: <https://www.kaggle.com/datasets/thedevastator/global-fossil-co2-emissions-by-country>.
- [2] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ, USA: Wiley, 2016.
- [3] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, Australia: OTexts, 2021. [Online]. Available: <https://otexts.com/fpp3/>.
- [4] Intergovernmental Panel on Climate Change (IPCC), Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK: Cambridge Univ. Press, 2022.



(a) Per-capita CO<sub>2</sub> emissions for 20 focus countries (latest available year, 2022).



(b) Total CO<sub>2</sub> emissions for 20 focus countries (latest available year, 2022).

Fig. 5: Comparison of CO<sub>2</sub> emissions.

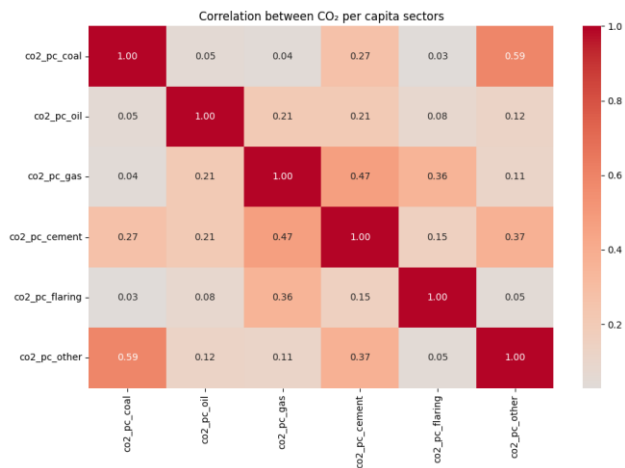


Fig. 6: Correlation between per-capita CO<sub>2</sub> emissions from coal, oil, gas, cement, flaring, and other sources.