Visualizing Global Temperature and Environmental Trends

Aheli Das

B.Tech in Data Science

Swami Vivekananda University

Period of Internship: 25th August 2025 - 19th September 2025

Report submitted to: IDEAS – Institute of Data Engineering, Analytics and Science Foundation, ISI Kolkata

1. Abstract

This project presents a comprehensive visual analysis of global environmental trends, with a primary focus on rising temperatures from the year 2000 to 2024. Using a curated dataset containing key environmental indicators for various countries, this study performs an in-depth exploratory data analysis. The core objective is to uncover and illustrate the patterns of climate change through intuitive visualizations. The analysis investigates the overall trend of global temperatures, compares these trends across different nations, and examines the correlation between temperature, carbon dioxide emissions, and sea-level rise. Key findings reveal a consistent and unmistakable upward trajectory in global average temperatures. The visualizations effectively communicate the urgency of this trend and highlight the interconnected nature of critical environmental factors, providing a data-driven narrative on the progression of climate change.

2. Introduction

The issue of global warming and climate change is one of the most pressing challenges of our time. Understanding the patterns and correlations within historical climate data is crucial for raising awareness and informing future policies. This project aims to contribute to this understanding by leveraging the power of data visualization. The relevance of this work lies in its ability to transform complex numerical data into clear, accessible, and impactful visual stories that underscore the reality of global temperature changes.

The project was developed using the Python programming language within a Google Colab environment. Key technologies included the Pandas library for robust data manipulation and analysis, along with Matplotlib and Seaborn for creating a wide array of static and informative plots. The procedure involved sourcing a relevant dataset, performing necessary data aggregation, and systematically generating visualizations to explore different facets of the data.

During the initial phase of the internship, training was received on several foundational topics that were directly applied in this project, including:

- Data Loading and Cleaning with Pandas
- Fundamentals of Time-Series Analysis
- Exploratory Data Analysis (EDA) Techniques
- Advanced Data Visualization using Matplotlib and Seaborn
- Statistical Grouping and Aggregation Methods

3. Project Objective

The primary goal of this project is to analyze and visualize multi-country environmental data to draw meaningful inferences about climate trends. The specific objectives are as follows:

- To analyze and visualize the overall trend of the average global temperature from 2000 to 2024.
- To conduct a comparative analysis of average temperature trends across a diverse set of countries to identify regional patterns and disparities.
- To investigate the potential relationships between rising average temperatures, CO2 emissions, and sea-level rise using correlation plots.
- To generate a series of clear and compelling visualizations—including line charts, bar plots, heatmaps, and pair plots—to effectively communicate the analytical findings.
- To synthesize the results to draw data-driven conclusions about the progression and key indicators of global climate change.

4. Methodology

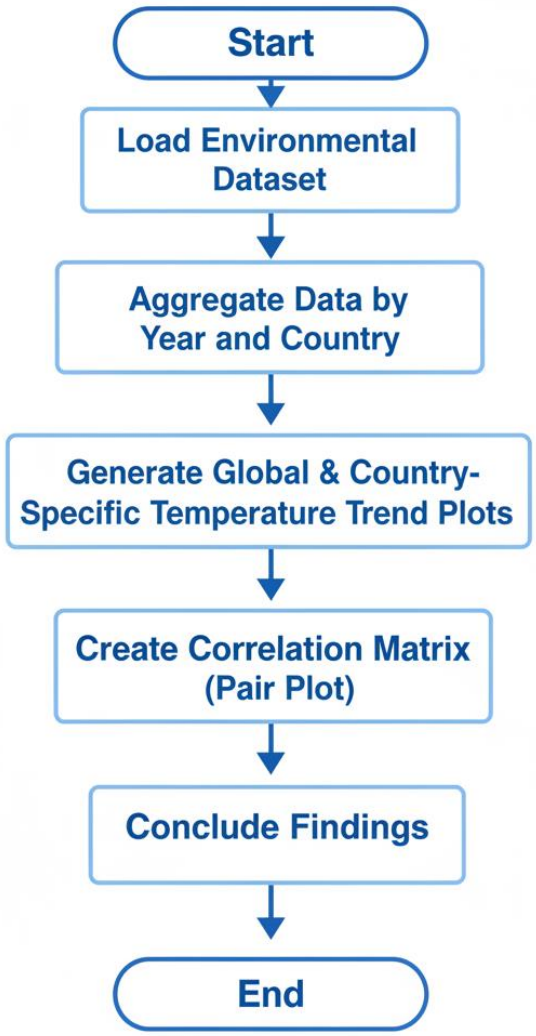
The project followed a structured methodology to ensure a thorough and effective analysis of the environmental data.

Data Collection:
The analysis is based on the "Global Environmental Trends" dataset, which was sourced from a public repository. This dataset contains annual records from 2000 to 2024 for 19 countries, with columns including Avg_Temperature_degC, CO2_Emissions_tons_per_capita, Sea_Level_Rise_mm, Population, and Renewable_Energy_pct.

Data Processing and Analysis:
The entire analysis was performed using Python in a Google Colab notebook. The process was as follows:

1. Data Loading: The dataset was loaded from a CSV file into a Pandas DataFrame.
2. Data Aggregation: The groupby() function in Pandas was extensively used to aggregate data by year and/or country to calculate mean values for temperature and other indicators. This was a critical step for preparing the data for time-series and comparative visualizations.
3. Visualization Tools: Matplotlib and Seaborn were the primary libraries used for plotting.
    - plt.plot() was used for creating time-series line graphs showing temperature trends.
    - sns.barplot() was utilized for comparing average temperatures across different countries.
    - sns.FacetGrid() was employed to create a matrix of line plots, allowing for simultaneous visualization of temperature trends for each individual country.
    - sns.pairplot() was generated to create a scatter matrix, which helped in visually assessing the pairwise relationships between temperature, CO2 emissions, sea-level rise, and the year.

This workflow can be visualized as a simple flow chart:

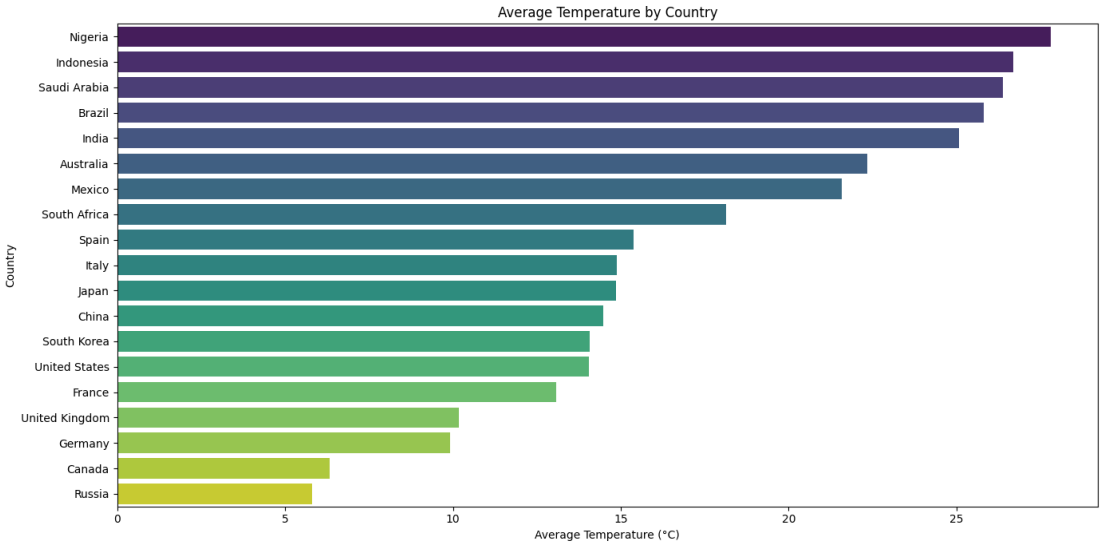


5. Data Analysis and Results

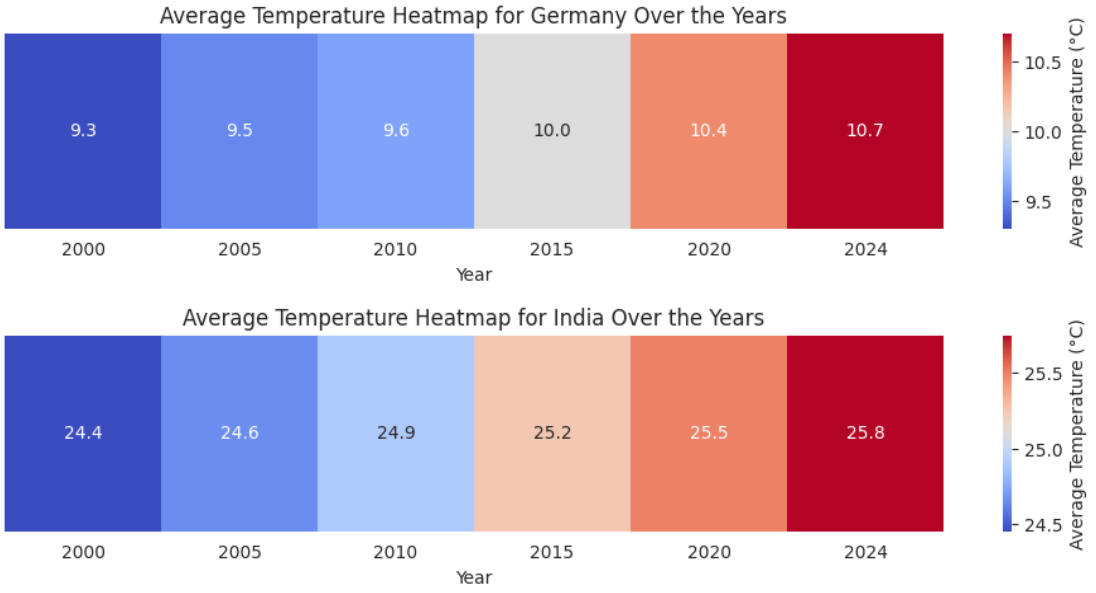
The analysis yielded several key insights into global environmental trends, presented through the following visualizations.

Descriptive Analysis:

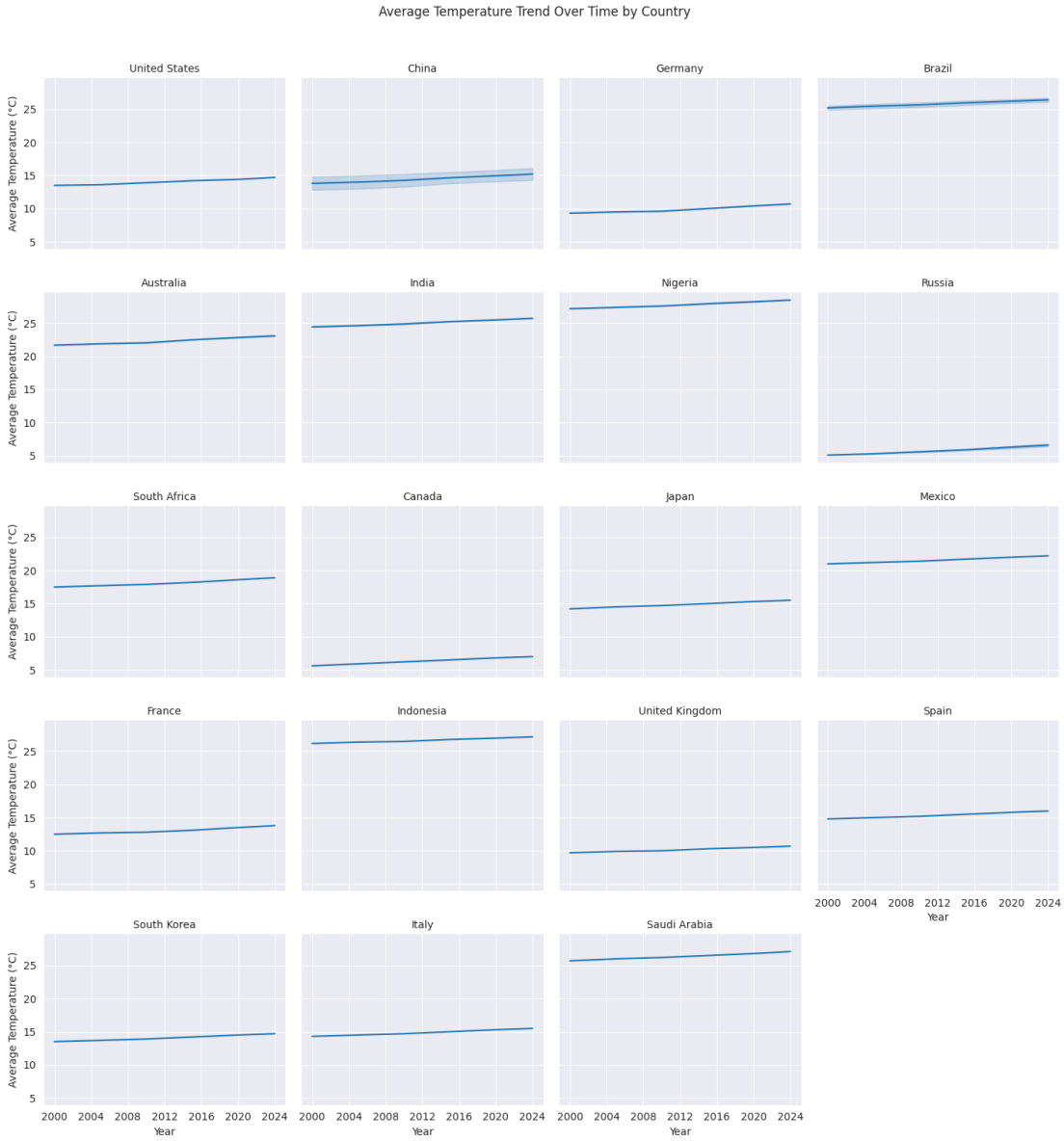
- **Global Temperature Trend**: A line plot of the average global temperature across all countries from 2000 to 2024 shows a clear and steady upward trend. This visualization confirms that, on average, the planet has been getting progressively warmer over the past two and a half decades.



- **Average Temperature Heatmap for India and Germany**: The heatmap for India provides a granular look at how temperatures have evolved. The color gradient, shifting from cooler blues to warmer reds, visually represents a consistent increase in average temperature over the years, with 2024 appearing as the warmest period.

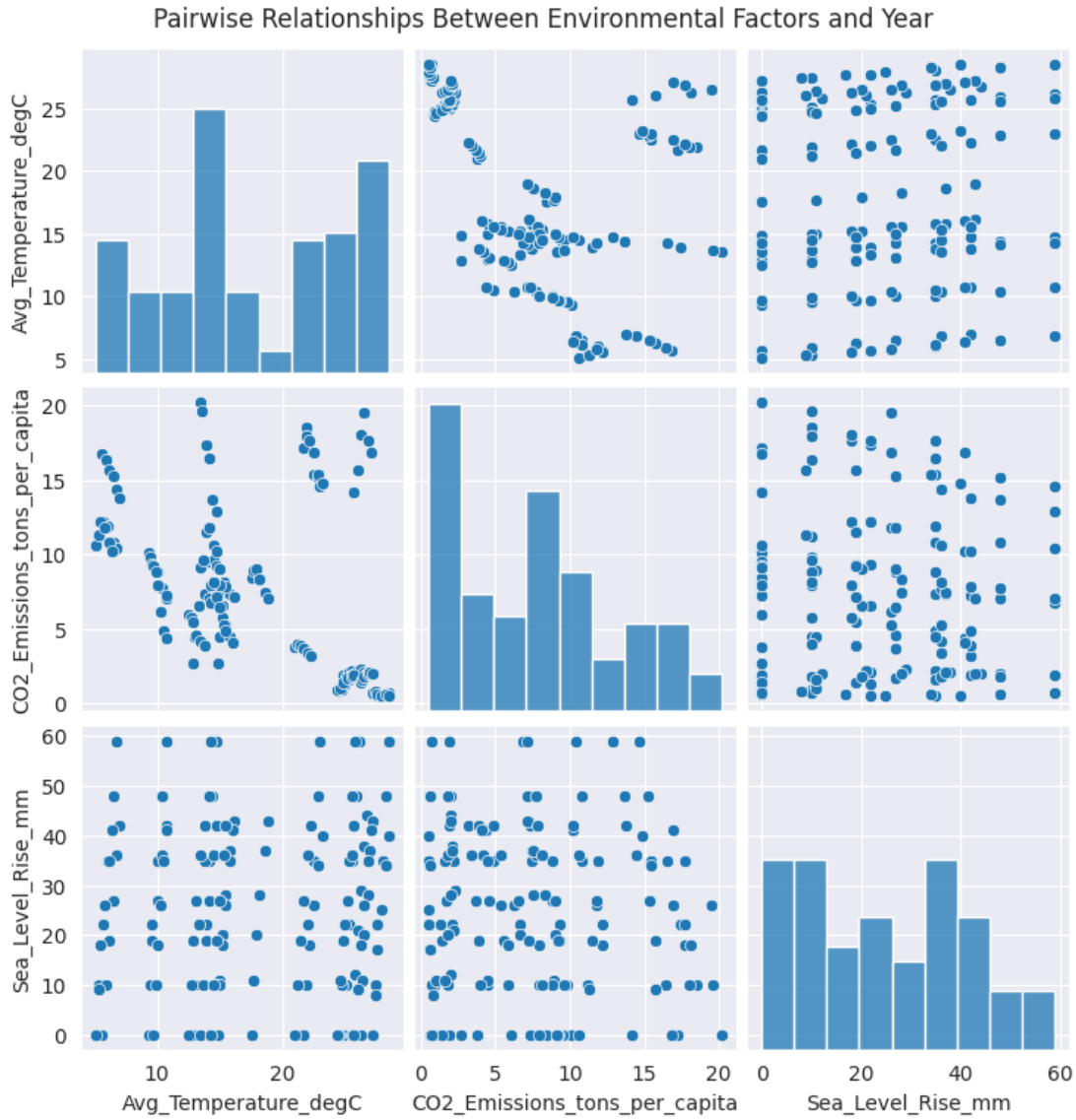


- Country-Specific Temperature Trends: To understand regional differences, a facet grid of line plots was created. This powerful visualization shows the temperature trend for each country individually. While nearly all countries exhibit a warming trend, the slope and variability of this trend differ, highlighting that climate change does not impact all regions uniformly.



Inferential Analysis:

- Correlations Between Environmental Factors: A pair plot was generated to explore the relationships between Year, Average Temperature, CO2 Emissions, and Sea Level Rise. The scatter plots revealed strong positive correlations:



- As the Year increases, the Average Temperature tends to rise.
- There is a visible positive relationship between increasing Average Temperature and rising CO2 Emissions and Sea Level Rise.
- This suggests that these critical environmental metrics are interconnected and have been increasing in tandem over the studied period.

Conclusion

The analytical work conducted in this project leads to several data-supported conclusions. The visualizations unequivocally demonstrate a significant and persistent global warming trend from 2000 to 2024. This is not an isolated phenomenon but a widespread pattern observable across numerous countries, albeit with regional variations.

Furthermore, the strong positive correlations found between rising temperatures, increasing CO2 emissions, and sea-level rise reinforce the multifaced nature of climate change. The findings are consistent with the

broader scientific consensus and highlight the value of data visualization in making these complex trends understandable to a wider audience.
For future work, this analysis could be expanded by incorporating a longer time series to observe more historical patterns. Additionally, predictive time-series models like ARIMA or Prophet could be applied to forecast future temperature trends based on the historical data. Finally, a deeper analysis could be performed to quantify the impact of renewable energy adoption on CO2 emissions within the dataset.

7. APPENDICES

- The "Global Environmental Trends" dataset was sourced from https://www.kaggle.com/datasets/adilshamim8/temperature.
- Github link- https://github.com/dasaheli14/IDEAS
- https://colab.research.google.com/drive/1sXBVnYkYUC6_nQOYeFSfdd_vNG3p97eH