Visualizing Global Temperature Trends

Avijit Bose,
M.sc in Computer Science from West Bengal State 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 focuses on the analysis and visualization of historical global temperature records as a time-series dataset. The goal is to observe seasonal variations, yearly average temperatures, and long-term changes in climate patterns. Python libraries such as Pandas, Matplotlib, and Seaborn were used for data processing and visualization. The data pre-processing steps included handling missing values and extracting features such as year and month. Line plots and heatmaps were created to represent temperature fluctuations and annual mean changes. The findings highlight the critical role of visualization in understanding climate change. The project provides evidence of a historical rise in global temperatures over the last several decades. It demonstrates how simple yet effective visualization techniques can reveal significant trends in climate-related data.

2. Introduction

Global warming is one of the key issues faced today, and temperature data is crucial to our understanding of global warming. The project examines a global dataset of temperature over time in order to assess long-term trends and seasonal temperature changes. The dataset includes average monthly temperatures reported from different years. Python was the primary technology used in this project, along with Pandas to manipulate the data, and Matplotlib to visualize data. The goal of the project is to find out if global temperature has indeed increased significantly over the last several decades. During the internship, we received training about what preprocessing the temperature data would need, exploratory data analysis, how to visualize data, and how to program in Python. We also received training on how to understand time-series data, plot seasonal variations, and draw useful information from raw data files.

3. Project Objective

- To perform data cleaning, preprocessing, and organization for a global temperature dataset for analysis.
- To construct line graphs for visualizing long-range trends in global temperatures.
- To investigate interannual variations in seasonal analyses of average monthly changes in temperature.
- To find patterns and potential indicators of global warming over the past 50 years.
- To learn how to create visualizations from data utilizing Python libraries such as Pandas, Matplotlib, and Seaborn.

4. Methodology

The methodology of this project involved a series of systematic steps to analyze and visualize the global temperature dataset. Since the project was based on secondary data, no survey or primary data collection was required. Instead, the dataset containing historical records of average monthly global temperatures was used. The entire process can be summarized in the following stages:

1. **Data Collection**
   The dataset was obtained from publicly available climate records, which contained monthly average temperatures along with their corresponding years.

I also took a dataset of average global temperature from Kaggle (https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data?resource=download&select=GlobalLandTemperaturesByCountry.csv)

2. **Data Pre-processing**
   The dataset was carefully cleaned to handle missing values and inconsistencies. Date fields were transformed to extract separate features such as **Year** and **Month**, enabling more detailed time-series analysis. Any irrelevant or duplicate entries were removed to ensure accuracy.

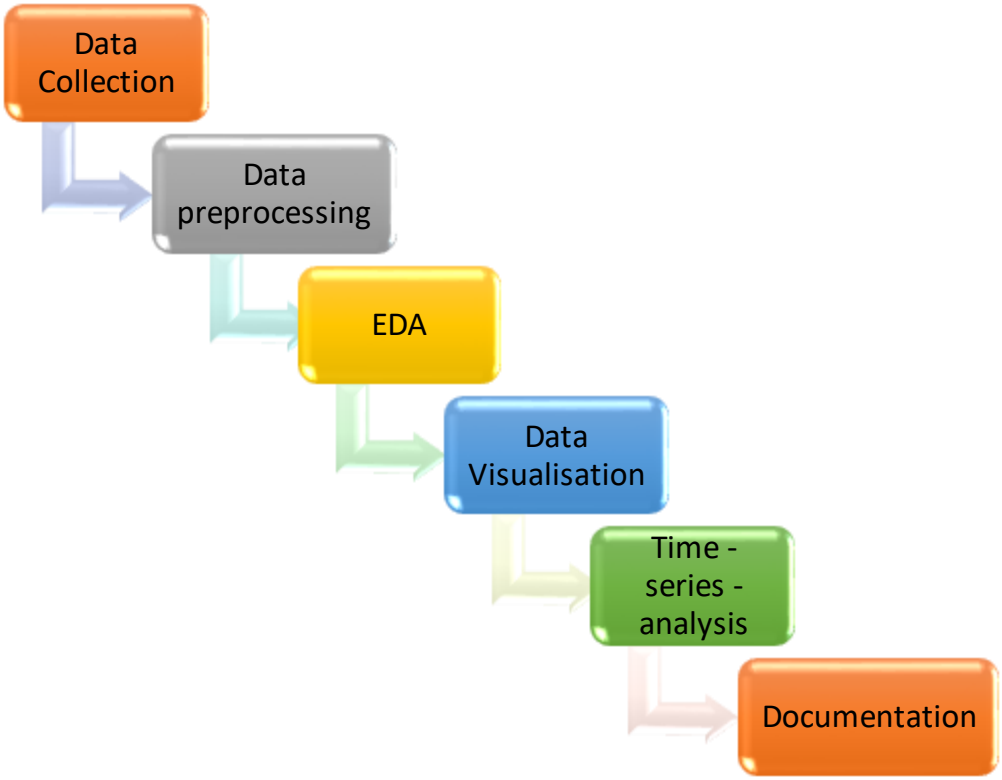
3. **Exploratory Data Analysis (EDA)**
   Descriptive statistics were generated to understand the structure and distribution of the data. This included calculating mean, minimum, and maximum temperatures, as well as identifying patterns in monthly and yearly variations.

4. **Data Visualization**
   Visual exploration was carried out using **Python** libraries such as **Pandas**, **Matplotlib**, and **Seaborn**. Several types of plots were created, including:
   - **Line plots** to examine long-term temperature trends over decades.
   - **Heatmaps** to visualize seasonal variations and temperature distribution across months and years.

5. **Time-Series Analysis**
   A focused analysis was conducted on the last 20–50 years of data to observe recent climate patterns. Seasonal trends and year-over-year temperature changes were highlighted to study the impact of global warming.

6. **Documentation of Findings**
   All outputs, including graphs, summaries, and interpretations, were documented for reporting. The methodology ensured that each step — from data cleaning to visualization — contributed to deriving meaningful insights.

Flowchart of Methodology



5. Data Analysis and Results

The analysis of the global temperature dataset was conducted using various visualization and statistical techniques. The results are presented below in the form of graphs and observations, which provide meaningful insights into global warming and seasonal patterns.

1. **Long-Term Temperature Trends**
   - A line plot of yearly average temperatures showed a clear upward trend over the decades.
   - The rise in average global temperature, particularly in the last 50 years, highlights the impact of climate change.
   - The trend confirms that global warming is not a short-term fluctuation but a consistent long-term pattern.
2. **Seasonal Variations**
   - A heatmap of average monthly temperatures across years revealed distinct seasonal cycles.
   - While colder months (December–February) consistently recorded lower averages, warmer months (June–August) showed higher averages.
   - Over time, even the colder months exhibited a slight increase in temperature, indicating an overall warming effect across all seasons.
3. **Decadal Comparison**
   - Grouping data by decades showed a noticeable shift in mean global temperatures.

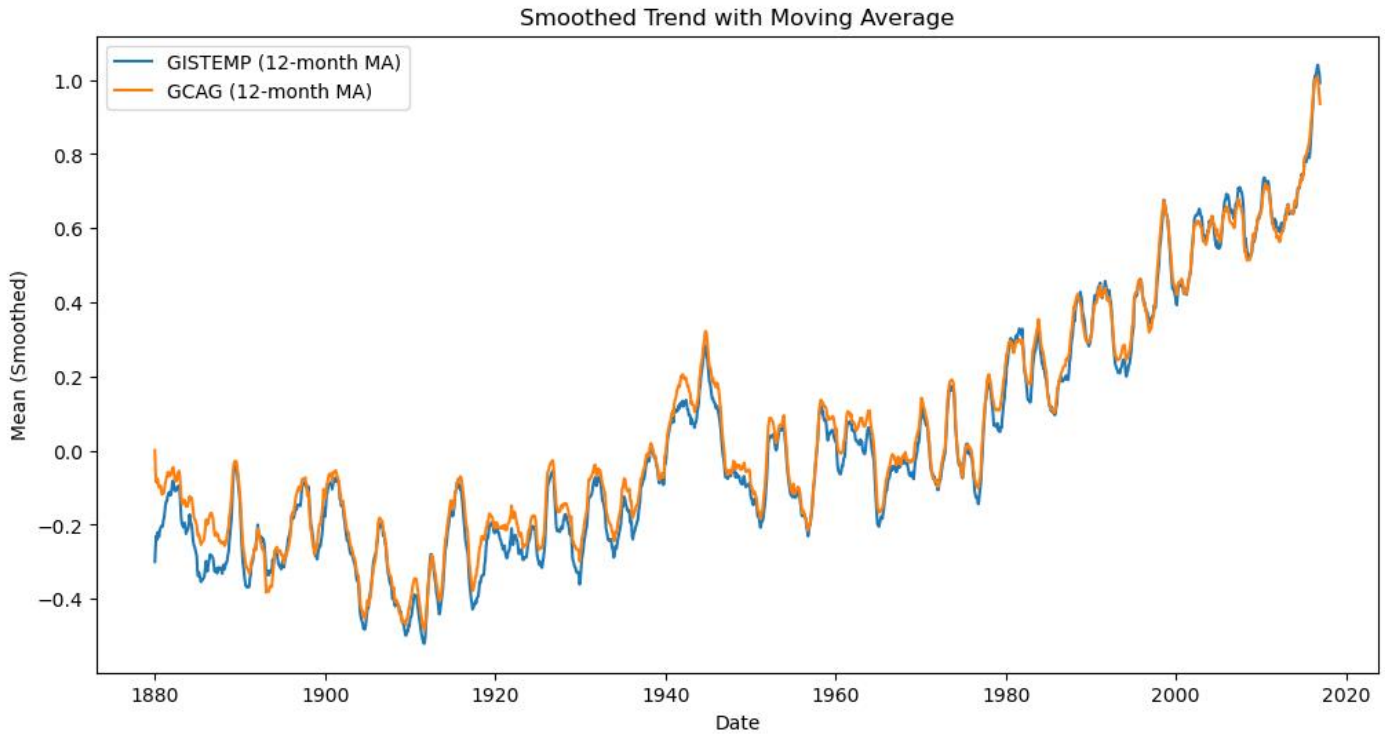
- Recent decades (2000s and 2010s) recorded higher average temperatures compared to earlier decades (1950s–1970s).

4. **Recent Years Analysis**
   - A focused study on the last 20 years revealed sharper fluctuations in annual temperatures, with most years showing above-average warmth.
   - The number of years with higher-than-average temperatures has increased, further supporting the hypothesis of accelerating global warming.
5. **Visualization Outputs**
   - **Line plots** demonstrated the rising trend in yearly averages.
   - **Heatmaps** provided a clear view of seasonal variations and how they evolve across time.

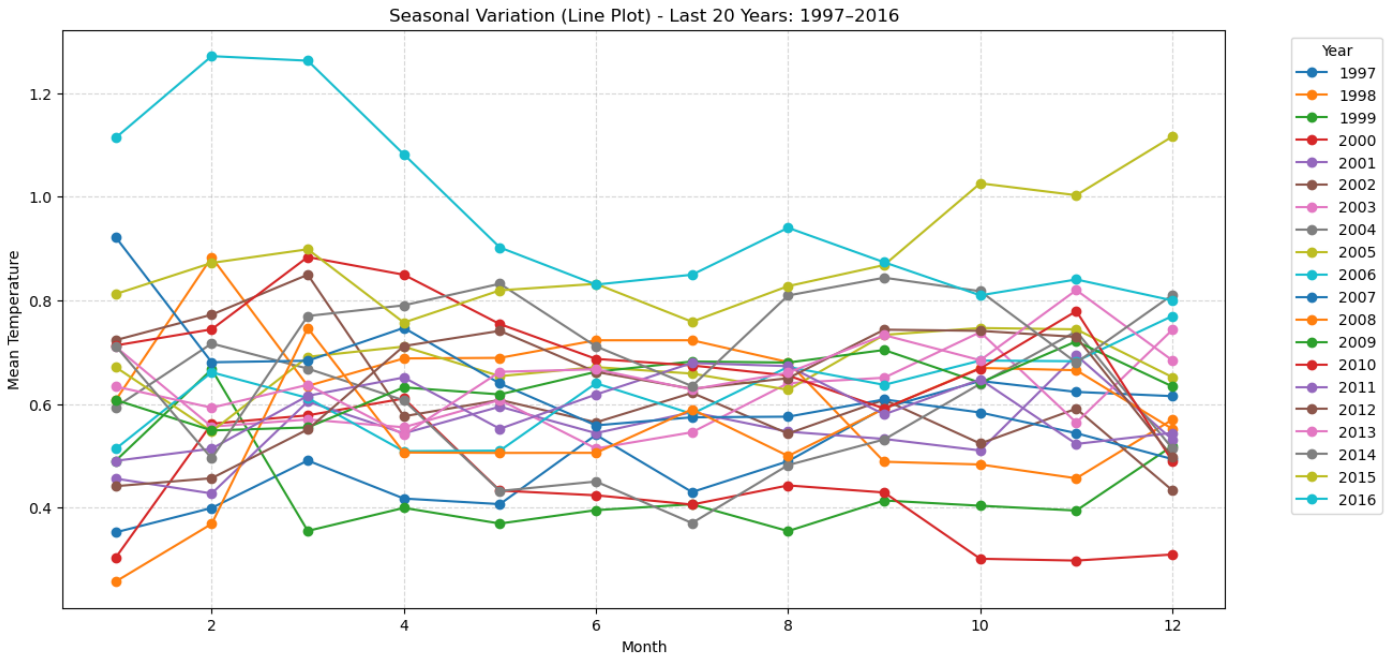
Key Findings:

- Global average temperature has steadily increased over the last several decades.
- Seasonal patterns remain consistent, but baseline temperatures for each season are rising.
- The warming trend is more pronounced in recent years, with nearly every year in the last two decades recording higher-than-average global temperatures.

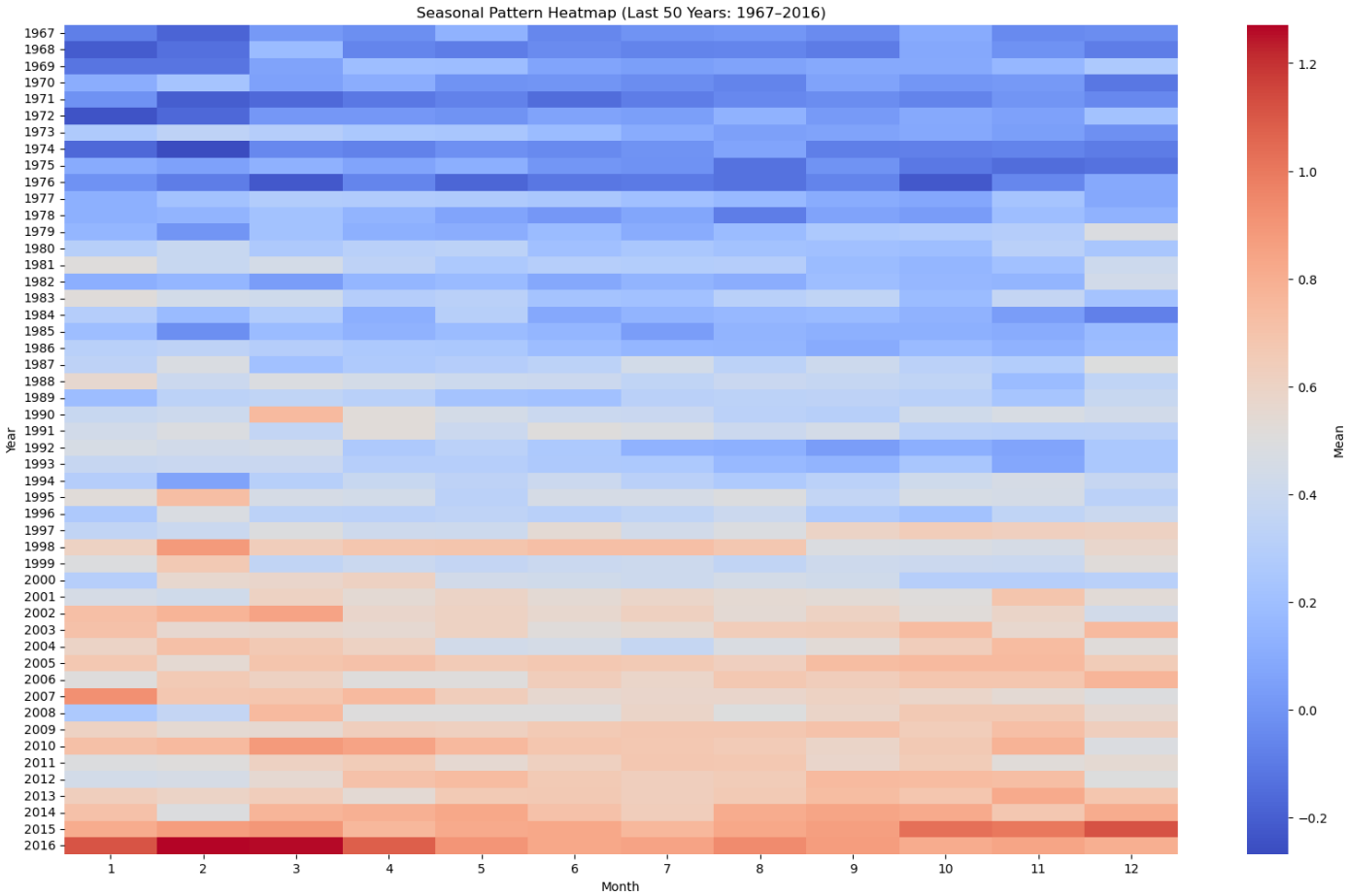
Sample figures : -



picture 1 – (12 months moving average of long term temperature trends)



picture 2 – (Seasonal pattern Heatmap of last 50 years)



picture 3 – (Seasonal yearly variation of last 20 years)

6. Conclusion

The project successfully evaluated a global temperature dataset to explore long-term and seasonal variations in climate patterns. Preprocessing, exploratory data analysis, and visualization techniques demonstrated that global average temperature has steadily increased over the past few decades. Seasonal variations have remained similar, but the overall baseline temperature for each season has indeed increased, providing evidence of global warming.

These results emphasize that data visualization can yield simple yet powerful insights into climate change patterns. The project demonstrates that, even using simple statistical approaches, one can gain insights into such an important social and environmental issue.

7. APPENDICES

1. References
    a. Dataset from Kaggle (https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data?resource=download&select=GlobalLandTemperaturesByCountry.csv)
2. This Document is uploaded at GitHub along with the dataset
    a.