Investigating the Relationship Between Atmospheric CO_2 and Global Temperature Anomalies

Gideon Afriyie

Abstract

As concerns about climate change and global warming grow, I became curious about the direct connection between rising atmospheric carbon dioxide (CO_2) levels and global temperature anomalies. This project investigates that relationship using real-world data from NASA and NOAA, performing data wrangling, correlation analysis, and linear regression modeling in Python. The findings reveal a strong statistical association between CO_2 and temperature rise, supporting the broader scientific consensus on climate change.

1 Project Motivation

Climate change is one of the most urgent scientific and societal challenges of our time. While numerous scientific organizations publish findings linking atmospheric CO_2 levels to global temperature increases, I wanted to investigate the data myself. This project was designed to explore the question:

How closely is the rise in atmospheric CO_2 associated with changes in global temperature anomalies?

To answer this, I used open-access datasets from NOAA and NASA to build a merged dataset of monthly CO_2 readings and corresponding global temperature anomalies.

2 Data Sources

- **NOAA CO_2 Concentration Data** (co2_trend_gl.csv): Daily atmospheric CO_2 levels measured at Mauna Loa, using the "trend" column to exclude short-term noise. I filtered the dataset to the first day of each month.
- **NASA GISTEMP Global Temperature Anomaly Data** (GLB.Ts+dSST.csv): Monthly global temperature anomalies, expressed in Celsius relative to a baseline average.

3 Data Merging

To analyze temporal relationships, the datasets were merged based on the `Date` column. Monthly CO_2 readings were aligned with monthly temperature anomalies, enabling valid regression and correlation analysis.

4 Visual and Statistical Analysis

4.1 CO_2 Concentration Over Time

Plots of atmospheric CO_2 over time show a clear upward trend, particularly from the mid-20th century onward. The rise from approximately 320 ppm to over 420 ppm reflects industrial emissions and land-use changes.

4.2 Temperature Anomalies Over Time

Global temperature anomalies also show a rising trend, with increased variability after the 1970s. This pattern is consistent with the global warming hypothesis.

4.3 Correlation and Regression

- **Pearson Correlation:** $r = 0.73$, indicating a strong positive correlation between CO_2 concentration and temperature anomalies.
- **Linear Regression Model:**

$$\text{Temperature Anomaly} = 0.006 \cdot CO_2 - 2.281$$

- **R^2 Score:** 0.73, meaning 73% of the variance in temperature anomalies is explained by CO_2 levels.

4.4 Smoothed Trends

Plotting smoothed CO_2 values against smoothed temperature anomalies reveals a consistent relationship, confirming the observed correlation in raw data.

5 Interpretation and Implications

The results support the scientific consensus that rising CO_2 levels are a major factor contributing to global warming. While CO_2 is not the sole driver, the data shows a strong and significant relationship with temperature anomalies.

6 What I Learned

- Techniques for cleaning and aligning real-world datasets
- How to perform regression and correlation analysis in Python
- How to visualize and interpret climate trends using open-source tools

7 Tools and Libraries

- **Python:** pandas, NumPy, matplotlib, sklearn
- **Data Sources:** NASA GISTEMP, NOAA CO_2 Records
- **Analysis:** Regression modeling, correlation, time series visualization

8 GitHub Repository

All code, datasets, and visualizations for this project are available on GitHub:

`https://github.com/GideonAfriyie/Climate-Data-Analysis-CO2-Concentration-vs.-Global-Temperature-Anomalies`