



Analysis of Land-Ocean Global Temperature Anomalies (1880-2024)

Scenario:

The dataset you have is an important historical record of global temperature anomalies. These anomalies are essential in understanding long-term climate patterns. By analyzing temperature changes over the years, you can gain valuable insights into trends such as global warming, seasonal changes, and regional variations. Your goal is to perform an **Exploratory Data Analysis (EDA)** to explore these patterns and present insights visually.

Dataset Description:

This dataset provides **monthly temperature anomalies** from 1880 to the present, with each row representing a different year. The temperature anomalies are compared to a long-term baseline (typically the 20th-century average) and are measured in **degrees Celsius**.

Columns in the Dataset:

1. **Year:** The year of observation.
2. **Month Columns (Jan-Dec):** The temperature anomaly for each month (from January to December).
3. **J-D:** The **annual mean temperature anomaly** (based on all 12 months of the year).
4. **D-N:** The **December to November** temperature anomaly.
5. **Seasonal Columns (DJF, MAM, JJA, SON):**
 - **DJF:** Winter temperature anomaly (December, January, February).
 - **MAM:** Spring temperature anomaly (March, April, May).
 - **JJA:** Summer temperature anomaly (June, July, August).
 - **SON:** Fall temperature anomaly (September, October, November).

Task Phases for EDA using Python:

1. Data Preprocessing:

- **Load the Data:**
 - Load the dataset into a Pandas DataFrame.
 - Inspect the first few rows to get a sense of the structure and format of the data.
- **Clean the Data:**
 - Check for missing values and handle them accordingly (e.g., filling or removing missing data).
 - Ensure that all columns (especially the year and temperature columns) have the correct data types (numeric for temperature anomalies, integers for the year).
- **Data Transformation:**
 - Convert the **Year** column to integer format if it is not already.
 - Extract the year as a separate column if necessary for easier manipulation.



- Ensure that the month and seasonal temperature columns are in a usable numeric format.

2. Exploratory Data Analysis (EDA):

- **Descriptive Statistics:**

- Use `.describe()` to summarize the data for each column (mean, median, min, max, etc.).
- Identify any unusually high or low temperature anomalies (outliers).

- **Visualize the Data:**

- **Time Series Plot (Overall Global Trend):**

- Plot the annual temperature anomalies (J-D) across the years to visualize the overall trend (is the world warming?).
- Use **line plots** to show temperature changes over time.

- **Monthly Anomalies Visualization:**

- Create a **line plot** for each month (Jan-Dec) showing how the global temperature anomalies for that month have changed over the years.
- Compare the temperature anomalies for different months over time.

- **Seasonal Anomalies Visualization:**

- Plot the seasonal temperature anomalies (DJF, MAM, JJA, SON) across the years to observe how different seasons have been affected by climate change.
- Use **line plots** or **stacked bar charts** to visualize the trends in seasonal anomalies.

- **Annual Temperature Distribution:**

- Visualize the **distribution of temperature anomalies** across all years using a **histogram** or **boxplot**.
- This will help you understand if the temperature anomalies tend to stay within a specific range or if there are more extreme values (outliers).

- **Seasonal Comparison:**

- Create **side-by-side bar plots** to compare the temperature anomalies for each season over the years.

- **Moving Averages:**

- Calculate and plot **moving averages** (e.g., 5-year or 10-year) for the temperature anomalies to smooth the fluctuations and highlight long-term trends.
- This will help identify periods of consistent warming or cooling over time.

3. Visualizing Correlations:

- **Correlation Analysis:**

- Explore the correlation between the temperature anomalies in different months and seasons using a **heatmap**.
- This will show how temperature anomalies for, for example, January relate to those in July, or how winter temperatures correlate with summer temperatures.

- **Month vs. Year Correlation:**

- Use a **heatmap** or **pairplot** to visualize how temperature anomalies in each month correlate with each other across the years.



4. Detecting Anomalies:

- **Identify Extreme Anomalies:**
 - Use **scatter plots** or **highlighted points** to identify years that have unusually high or low temperature anomalies.
 - These points may represent extreme weather events, volcanic eruptions, or other phenomena that have influenced the climate.

5. Insights and Findings:

- **Identify Key Trends:**
 - Discuss any visible patterns you see, such as:
 - Is there a clear upward or downward trend in the global temperature anomalies over time?
 - Which months or seasons show the greatest warming or cooling?
 - Are there any sudden changes in temperature anomalies during specific periods?
- **Compare Seasonal and Monthly Trends:**
 - Which seasons have seen the most significant temperature changes?
 - Are certain months (e.g., January, July) experiencing more dramatic shifts in temperature?
- **Report Visual Insights:**
 - Summarize your visual findings, focusing on the most notable trends, seasonal comparisons, and any anomalies in the dataset.

Deliverables:

1. **Cleaned Dataset:** A properly formatted dataset with missing values handled and necessary transformations done.
 2. **Python Script:** A detailed syntax with proper heading and inferences.
 3. **Presentation:** A written summary of the insights gained from the analysis, highlighting key trends, seasonal variations, and any detected anomalies.
-

****NOTE: Late Submission will have effect on their scores, no extension will be given****