STATISTICS AND TRENDS REPORT

Student Name: Muhammad Areeb Gul **Student ID:** 24177904
GitHub: Areebgul/ADS1-stats-trends-assignment: First assignment **Dataset:** Kaggle (Global Temperature Data)

INTRODUCTION

In this report, I have explored a dataset about **Global Temperature**. The purpose of this report is to show my understanding of data preparation, visualisation, and statistical moments in a clear and simple way.

Goal:
The goal is to study how the temperature varies over the time and learn the overall tendencies of global warming along with the help of simple data analysis and visualisation.

Dataset Details:
The dataset holds monthly and yearly temperature anomalies of the globe for years **1860 to 2020**.

DATA CLEANING AND PREPARATION

1. converted all numeric columns into float type
2. eliminated empty columns
3. replaced missing values using forward fill

STATISTICAL MOMENTS

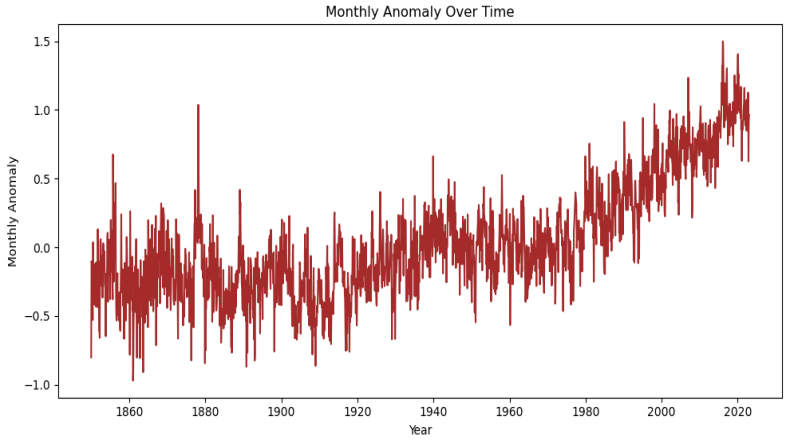
To get better understanding of data, I focused on column **"Monthly Anomaly".**
The findings indicate that the mean temperature anomaly is slightly higher than zero meaning that global temperatures have increased generally over the last few years. The skewness and kurtosis help to understand that most data points are near the mean, but there are some higher peaks in some years.

The following are the four statistical moments that I estimated:

- **Mean:** shows the average temperature anomaly.
- **Variance:** shows the level of distribution of the values around the mean.
- **Skewness:** indicates whether the data is more on the left or on the right.
- **Kurtosis:** shows how sharp or flat the data distribution is.

DATA ANALYSIS

1. ***Plot 1: Line Plot (Relational Plot)***

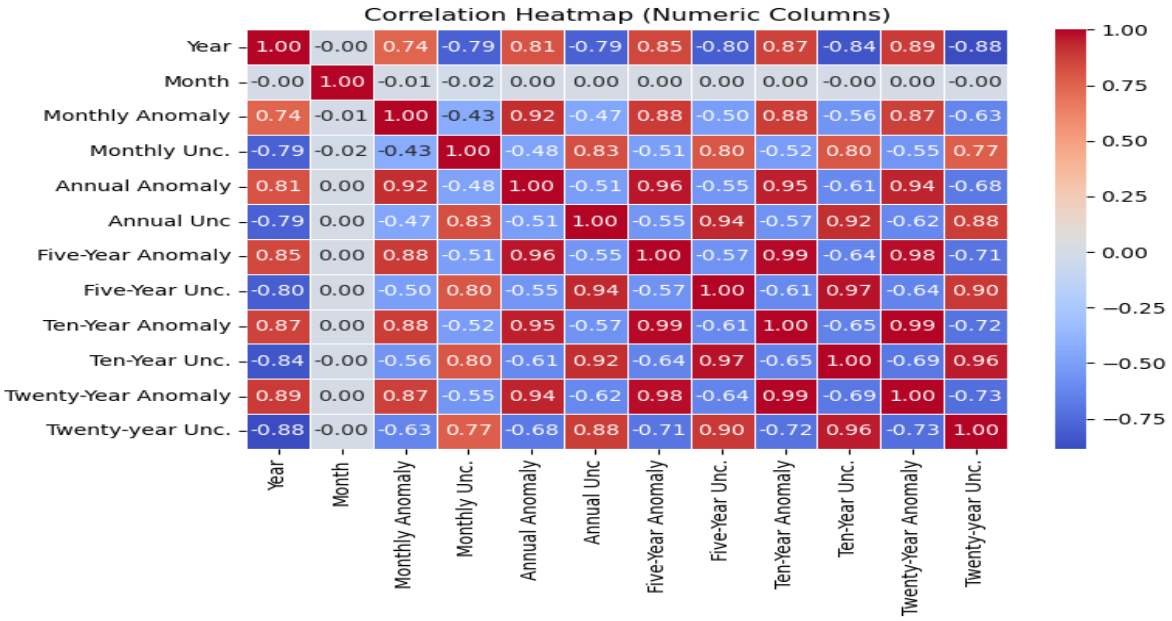


Description: As it occurs in the line plot, we can distinctly notice that the temperature is
rising gradually over time.

Insights:

- This shows a rising global temperature trend
- one of the key indicators of climate change. This helps to see the trend of temperature change across different years.

2. ***Plot 2: Correlation Heatmap (Statistical Plot)***
Description: There is a strong positive correlation between the monthly and yearly anomalies.
Insights:
   - *It is the same as one rising and the other as well.*
   - *It is logical because they measure the same temperature trend but at different time scales*

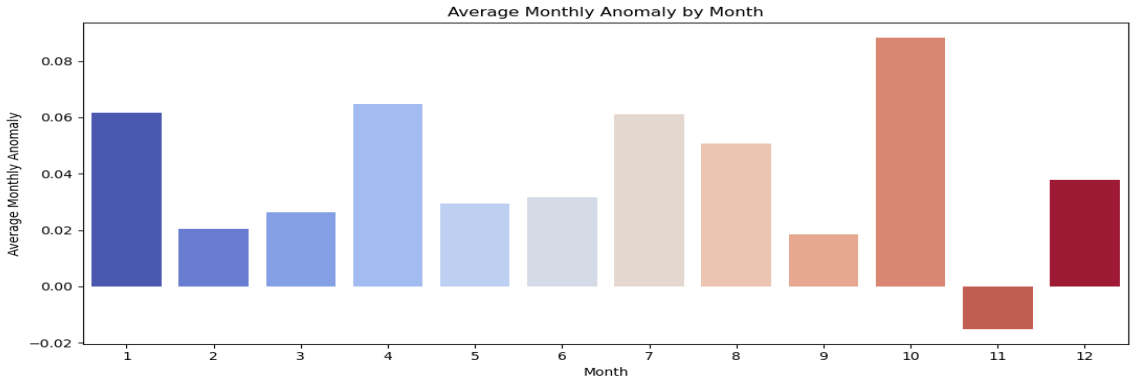


3. ***Plot 3: Bar Chart (Categorical Plot)***

Description: There are a few months when it is a little warmer or a little colder, but the variations are not too significant.

Insights:

- This tells us that the trend of the anomaly in temperature is not just in one specific season but **also in all the months.**



DISCUSSION

Observation:

Based on the findings, we observe that the global temperature anomalies have had a clear upward trend.

Visual Insights:

- The bar chart shows the pattern is similar across all months.
- The heatmap supports that different anomaly measurements are related and consistent.

CONCLUSION

In this assignment, I have applied **Python, Pandas, Seaborn, and Matplotlib** to clean and analyse a real-world dataset. I created three plots line, bar, and heatmap and calculated the four main statistical moments.

It shows that data science could assist us to understand large datasets and discovering real world trends such as climate change.