# Global Weather Analysis Report

**My Product Management Journey: Accelerating My Career**

Progress Report Towards Achieving PM Accelerator's Mission:

1. **Introduction:**

It's an analysis project to predict the future weather for "Global Weather Repository.csv." There is provided daily weather data for cities of almost all cities worldwide, and that data includes over 40 features. Perform EDA and predictive models developed on it. A complex analysis has been carried out to draw inferences that present global trends insights for the future weather.

1. **Data cleaning and preprocessing:**

Several preprocessing steps are performed on the dataset before analysis. For temperature and precipitation, mean/mode imputation was done to fill missing values to make data complete. Both Z-score and IQR methods have identified outliers, and they have been removed from the dataset because extreme values would skew the analysis. Min-Max scaling is also performed on the numerical features where all values are made to fall in a uniform range for comparison purposes of different features. The above preprocessing steps enhanced the quality and accuracy of data further before further analysis.

1. **Exploratory Data Analysis (EDA):**

Through exploratory data analysis, the trends, correlations, and patterns that exist in the dataset were determined. Seasonal variations in temperature over time were also portrayed using line plots. Histogram and KDE plot were used to study the aspect patterns of precipitation because these will enable understanding of what is going on in terms of the underlying distribution and change that is happening. To check whether there exists any correlation among the concerned factors of air quality, temperature, and more, correlation analysis was done using a heat map. Further, time series analysis was conducted, which incorporated rolling average to detect all forms of fluctuations or long trends that may exist in the graph.

1. **Forecasting Models:**

last\_updated feature was utilized to develop time series forecasting models. This model was used for short term trend analysis. The model used previous values as predictions to be secured for correct forecasts. Seasonality components were incorporated in the case of long-term forecasts, using periodic variations. Besides, Facebook Prophet was utilized along with the approach of deep learning by modelling complicated temporal dependencies through LSTM networks within the fluctuation of the data. In turn, it utilizes RMSE, MAPE, and R² scores in judging the model. From the results, accuracy and efficiency comparisons between different techniques used in the forecasting will be made.

1. **Advanced Analysis:**

Further analyses were performed on the dataset. Anomaly detection was performed using Isolation Forest and DBSCAN algorithms to identify the abnormal weather patterns and the outliers. Climatic analysis was performed to determine long-term trends over the regions to identify changing conditions of temperature, rainfall, and air quality with time. Feature importance analysis was adopted through Random Forest and from SHAP values to identify what parameters are contributing the most in determining the climate. It yielded information regarding how the most impact factors affect the climate condition. Geographical patterns of weather have been visualized through Folium and GeoPandas for doing spatial analysis about how climate conditions may vary across the regions. Some of the important findings regarding the shift in weather characteristics and predictability were realized.

1. **Key Insights & Conclusion:**

Trending revealed yearly cycles of high variability for both temperatures and precipitations, thereby predictability during other times of the year. This was strongly interrelated with both temperature fluctuations being interrelated, implying the causative effects by changes in the climatic situations. Prophet outperformed ARIMA results for trend forecasts over longer periods, and LSTM succeeded mostly with complex relationships in the variability and trends at time dependencies. These showed substantial weather variance across the geographical areas, in turn emphasizing region-specific differences across continents.

1. **Demo video link:**

<https://www.loom.com/share/bc3b2203a62943d2804d46d298a0ccba?sid=ad4c2455-d98e-43fd-b71f-0a7efd5b6216>