# Global Weather Analysis and Forecasting Report

## PM Accelerator Mission

By making industry-leading tools and education available to individuals from all backgrounds, we level the playing field for future PM leaders. This is the PM Accelerator motto, as we grant aspiring and experienced PMs what they need most – Access. We introduce you to industry leaders, surround you with the right PM ecosystem, and discover the new world of AI product management skills.  
  
This project aligns with the PM Accelerator mission by leveraging advanced data analytics, machine learning, and forecasting models to enhance decision-making and predictive capabilities.

## 1. Introduction

\*\*Objective\*\*  
The goal of this project is to analyze a global weather dataset, perform exploratory data analysis (EDA), detect anomalies, and build predictive models for temperature forecasting using machine learning techniques.

\*\*Scope\*\*  
- Data Cleaning and Preprocessing  
- Exploratory Data Analysis (EDA)  
- Time-Series Forecasting using ARIMA and Prophet models  
- Anomaly Detection  
- Feature Importance Analysis using Machine Learning  
- Geospatial Weather Visualization

## 2. Data Processing and Cleaning

\*\*Data Source\*\*  
The dataset contains global weather records, including temperature, humidity, wind speed, and geographic information.

\*\*Data Cleaning Steps:\*\*  
- Handling Missing Values: Forward fill method (`ffill`) was used.  
- Outlier Detection: Interquartile Range (IQR) method was applied.  
- Feature Scaling: StandardScaler was used.  
- Datetime Conversion: Converted `last\_updated` column.

## 3. Exploratory Data Analysis (EDA)

\*\*Key Insights:\*\*  
- Temperature trends visualized over time.  
- Correlation analysis via heatmaps.  
- Moving averages for seasonal trends.

\*\*Visualizations Used:\*\*  
- Line plots  
- Correlation heatmaps  
- Seasonal rolling averages

## 4. Forecasting Models

\*\*ARIMA Model\*\*  
- Used to model temperature trends.  
- Training and testing split (80-20).  
- Evaluated with Mean Absolute Error (MAE) and Mean Squared Error (MSE).

\*\*Prophet Model\*\*  
- Used for long-term forecasting.  
- Forecasted temperature trends for the next 30 days.

## 5. Anomaly Detection

- Z-score method used to detect anomalies.  
- Any point with a Z-score > 3 was flagged as an anomaly.

## 6. Machine Learning - Feature Importance Analysis

- Random Forest Regressor trained to analyze feature importance.  
- Identified key weather attributes affecting temperature.

## 7. Geospatial Weather Visualization

- Used Folium maps to visualize temperatures across locations.  
- Interactive map created and saved as HTML.

## 8. Project Submission

\*\*GitHub Repository:\*\*  
- Includes Python scripts, visualizations, and a README file.  
- https://github.com/prabhatadvait/Weather\_forecasting

## 9. Conclusion and Key Takeaways

- Data preprocessing improved model performance.  
- ARIMA and Prophet models captured seasonal trends.  
- Machine learning analysis identified key weather attributes.  
- Geospatial visualization helped in climate studies.

## 10. Next Steps

- Extend forecasting to additional weather attributes.  
- Implement deep learning models like LSTMs.  
- Integrate real-time weather data.