**Global Weather Forecasting & Climate Analysis 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. PM Accelerator grants aspiring and experienced PMs what they need most – *Access*. We introduce you to industry leaders, surround you with the right PM ecosystem, and help you discover the new world of AI Product Management skills. PM Accelerator is committed to empowering aspiring and experienced Product Managers through quality education, real-world opportunities, and a supportive community. In line with this mission, we also extend our support to disadvantaged youths through PMA Kids – fostering the next generation of leaders and innovators.”

**Introduction:**

This project explores global weather patterns and climate dynamics using advanced data analytics and forecasting techniques. By analyzing worldwide temperature, air quality, and atmospheric conditions, we uncover hidden anomalies and regional variations.  
Multiple forecasting models are compared to predict temperature trends and assess their accuracy. The study aims to translate complex environmental data into actionable insights aligning with PM Accelerator’s mission to democratize access to analytical and AI-driven tools for future leaders.

**1. Data Cleaning & Preprocessing**

**Steps Performed:**

* Loaded the Global Weather Repository dataset containing 98,604 records and 41 features.
* Converted date columns (last\_updated) into datetime format and extracted useful features such as year.
* Handled missing values using mean imputation for numeric data and mode imputation for categorical fields.
* Removed unrealistic outliers using the IQR (Interquartile Range) technique.
* Standardized key numerical features like temperature, humidity, wind speed, and pressure for modeling.

**Result:**

Cleaned dataset ready for analysis and modeling with consistent, complete, and noise-free weather data.

**2. Advanced Exploratory Data Analysis (EDA)**

**a. Feature Overview**

* Visualized the distribution of temperature across the globe.
* Created a correlation heatmap between weather and air-quality variables.

**Insights:**  
Temperature shows moderate negative correlation with humidity and pressure, and a mild positive correlation with PM10 air pollutants. This suggests warmer, drier conditions often coincide with poor air quality.

A graph of a temperature

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

**b. Anomaly Detection – Outliers in Global Weather:**

Used the IQR method to detect temperature anomalies worldwide.

**Insights:**  
Out of 98,604 readings, ~12% were anomalies extremely high or low temperatures beyond normal seasonal variation.  
These outliers may represent heatwaves or cold snaps, signaling climate instability.

A graph of a temperature

AI-generated content may be incorrect.

**3. Forecasting with Multiple Models**

**Models Built**

1. **ARIMA** – Captures linear time-based dependencies.
2. **Prophet** – Handles trend + seasonality with missing data robustness.
3. **XGBoost** – Learns nonlinear temporal patterns.
4. **Ensemble Model** – Combines ARIMA, Prophet, and XGBoost outputs.

**Insights:**  
ARIMA performed best (MAE: 1.089, RMSE: 1.605) on stable trends, while the ensemble smoothed short-term noise.  
This combination approach enhances forecast reliability for global temperature prediction.

A graph showing the temperature of a temperature

AI-generated content may be incorrect.

**4. Unique Analyses**

**a. Climate Analysis – Regional Heat Patterns:**

Analyzed the average temperature by country.

**Insights:**  
Countries such as Saudi Arabia, Morocco, and Turkmenistan ranked hottest.  
Heat zones align with desert and tropical climates, indicating extreme regional stress under global warming.

A graph of different colors

AI-generated content may be incorrect.

**b. Feature Importance (Exploratory)**

Used an XGBoost regressor to rank which factors most influence temperature.

**Top Features:**

1. Pressure (0.37)
2. Humidity (0.27)
3. PM10 concentration (0.12)

**Insights:**  
Pressure and humidity are the dominant factors affecting global temperature variation, while pollution adds secondary effects.

A screen shot of a graph

AI-generated content may be incorrect.

**c. Spatial Analysis – Global Weather Anomalies**

Plotted detected anomalies on a world map to visualize their spatial distribution.

**Insights:**  
Anomalies are concentrated in Africa, the Middle East, and South Asia, indicating climate stress in already hot regions.  
This global visualization shows temperature extremes are not random but geographically patterned.

A map of the world with red dots

AI-generated content may be incorrect.

**d. Environmental Impact – Weather vs Air Quality**

Analyzed how PM2.5 pollution correlates with temperature and humidity.

**Insights:**  
High PM2.5 concentrations were found in high-temperature and low-humidity zones.  
This confirms that hot, stagnant air traps pollutants, worsening air quality in developing urban areas.

A graph of red dots

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A diagram of different colors

AI-generated content may be incorrect.

**e. Geographical Patterns – Weather by Continent**

Grouped countries into continents and compared average climate metrics.

**Insights:**  
Asia and Africa recorded the highest mean temperatures.  
Oceania and Europe remain cooler, moderated by oceans and high latitudes.  
Pairplot analysis revealed temperature rises as humidity falls, confirming climate zone contrasts.

A graph showing the temperature of a temperature

AI-generated content may be incorrect.

A graph of a diagram

AI-generated content may be incorrect.

**Key Findings & Takeaways:**

* Data-driven anomaly detection highlights regions under severe heat variation.
* Air quality is strongly linked to temperature, suggesting policy focus on emissions in hot, urbanized regions.
* ARIMA-based forecasting offers a reliable foundation for near-term climate predictions.
* Geographical pattern analysis confirms that continental climate zones explain much of the world’s weather diversity.

**Conclusion:**

This project integrates data cleaning, EDA, anomaly detection, forecasting, and climate analytics to uncover global weather insights.  
It supports PM Accelerator’s mission of using data and AI to empower future innovators, demonstrating how analytical tools can drive real-world environmental intelligence.