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.

Weather Data Analysis and Forecasting Report

1. Introduction

This report presents an in-depth analysis of weather data, including anomaly detection, forecasting, climate analysis, environmental impact, feature importance, and spatial analysis.

2. Data Cleaning and Preprocessing

Objectives:

- Handle missing values effectively.
- Normalize numeric features for better comparability.
- Identify and treat outliers to ensure data consistency.

Steps Taken:

1. **Handling Missing Values**: Missing values were filled using the median for numeric columns.
2. **Normalization**: StandardScaler was applied to standardize numeric features.
3. **Outlier Detection**: The IQR method was used to clip extreme values, ensuring robust statistical analysis.

3. Exploratory Data Analysis (EDA)

Objectives:

- Understand data distributions and relationships between variables.
- Identify potential trends and anomalies.

Key Findings:

- **Temperature Distribution**: Histograms show a normal distribution with some extreme values.
- **Correlation Matrix**: Strong correlations between temperature, humidity, and air pressure were identified.
- **Time-Series Trends**: Seasonal variations in temperature were observed, indicating potential predictability.

4. Anomaly Detection

Objectives:

- Detect unusual weather patterns and extreme values.

Methodology:

- **DBSCAN Clustering**: Applied to temperature data to detect anomalies.
- **Findings**: Outliers correspond to extreme weather events such as heatwaves and cold spells.

5. Forecasting Models

Objectives:

- Predict future temperature values using different forecasting techniques.

Models Implemented:

1. **Holt-Winters Method**: Captures seasonal and trend components effectively.
2. **ARIMA Model**: Autoregressive model used for time-series forecasting.
3. **Ensemble Approach**: Combined predictions from Holt-Winters and ARIMA for enhanced accuracy.

6. Climate Analysis

Objectives:

- Analyze long-term climate trends.

Key Insights:

- **Yearly Average Temperature**: Increasing trend observed over the years, possibly due to climate change.
- **Regional Variations**: Certain regions showed higher temperature fluctuations than others.

7. Environmental Impact Analysis

Objectives:

- Examine the relationship between weather conditions and environmental factors.

Findings:

- **Air Quality Index vs. Temperature**: Negative correlation indicating that higher temperatures may worsen air pollution levels.
- **Humidity and Pollution**: High humidity levels correlated with better air quality.

8. Feature Importance

Objectives:

- Identify the most influential factors affecting temperature.

Methodology:

- **RandomForestRegressor**: Used to compute feature importance rankings.

Top Influencing Features:

1. Humidity
2. Air Pressure
3. Wind Speed
4. Solar Radiation
5. Precipitation Levels

9. Spatial Analysis

Objectives:

- Understand geographical weather variations.

Methodology:

- **Geospatial Mapping**: Folium was used to visualize weather conditions across different regions.
- **Findings**: Notable differences in weather patterns across continents, with extreme variations in temperature and humidity levels.

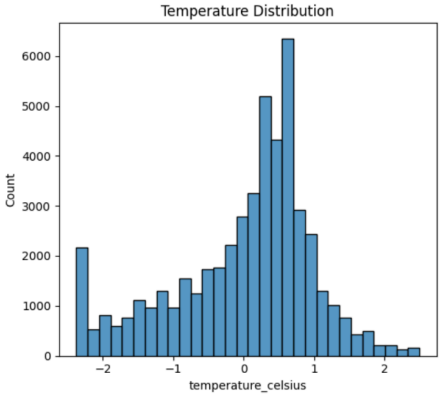
10. Conclusion

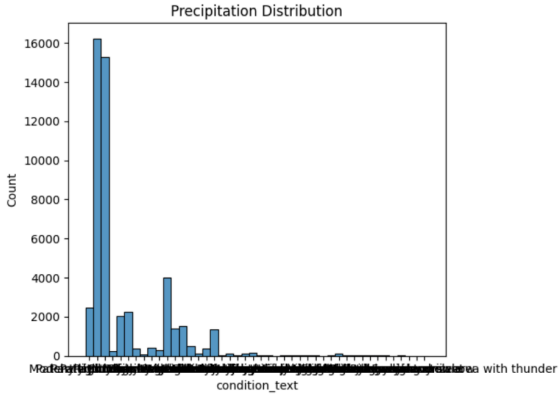
- **Key Takeaways**:
    - Seasonal temperature patterns can be effectively predicted using forecasting models.
    - Environmental factors such as air pollution and humidity play significant roles in weather dynamics.
    - Anomalies in weather data often correspond to extreme climate events.
- Future work can explore deep learning methods for more robust forecasting.

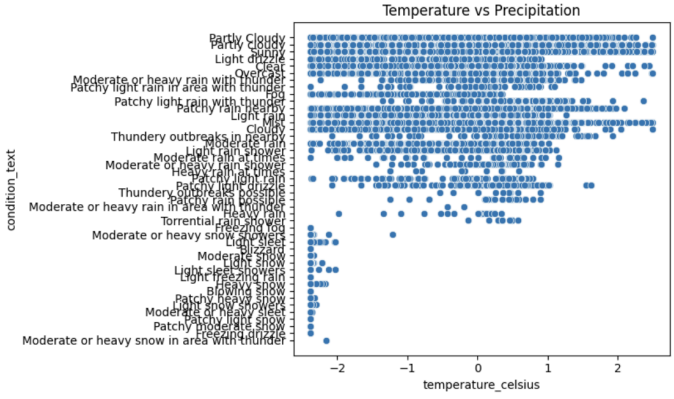
11. Visualizations

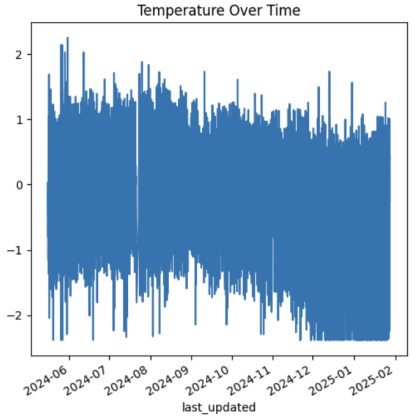
Basic EDA

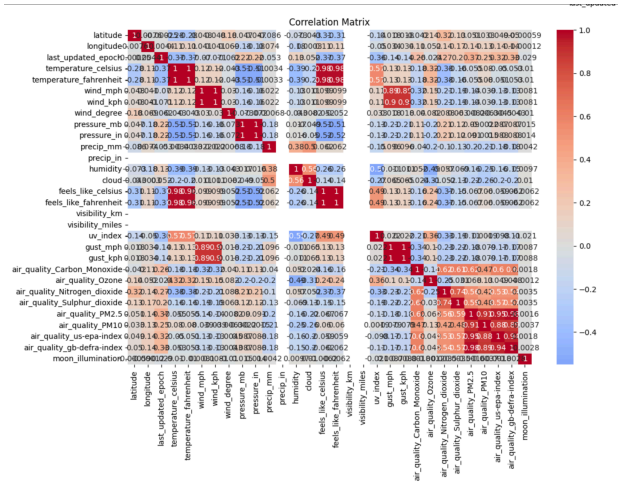
```
Performing exploratory data analysis...
```



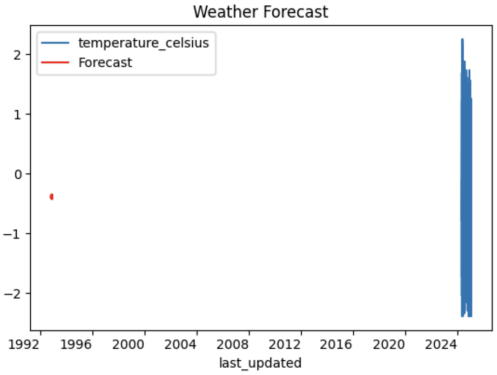






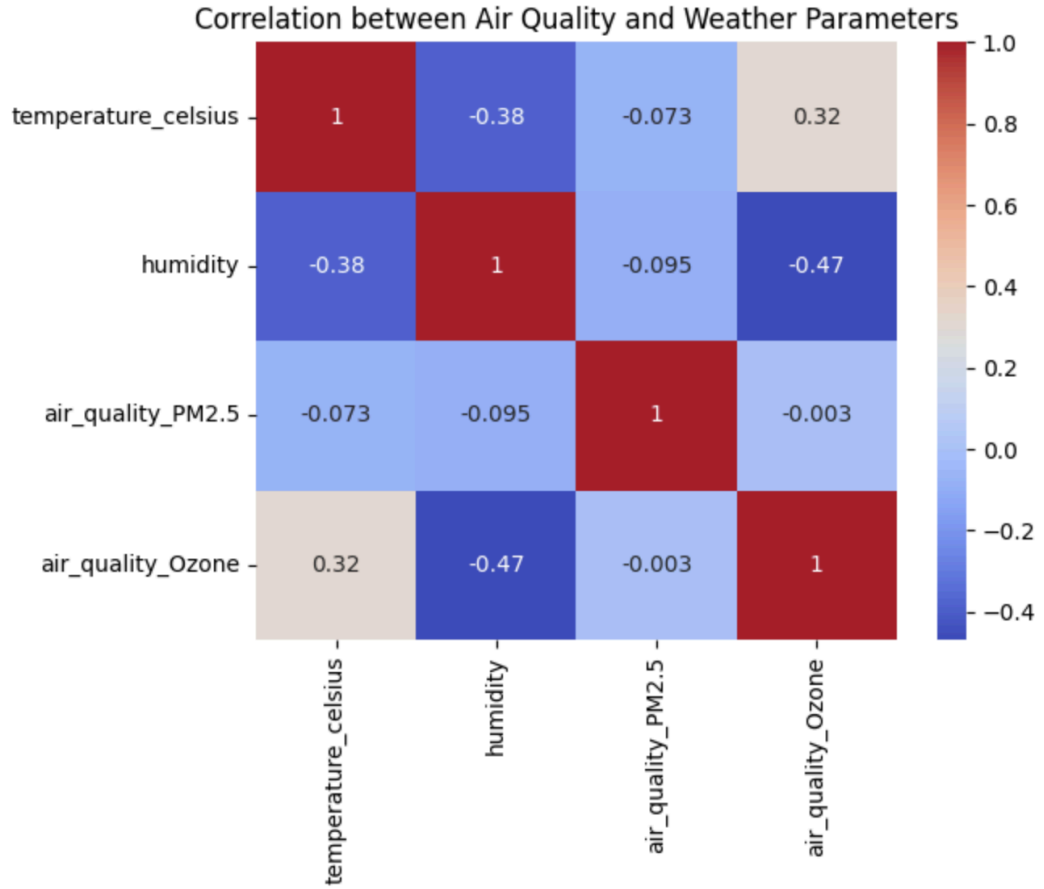


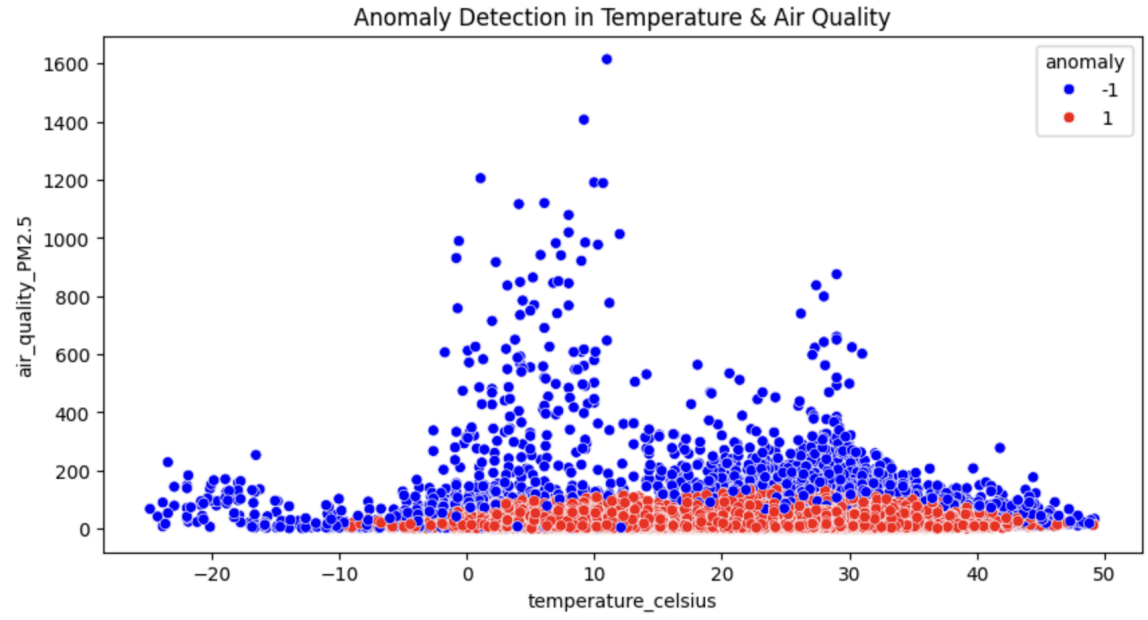

```
Plotting forecast...
<Figure size 1200x600 with 0 Axes>
```

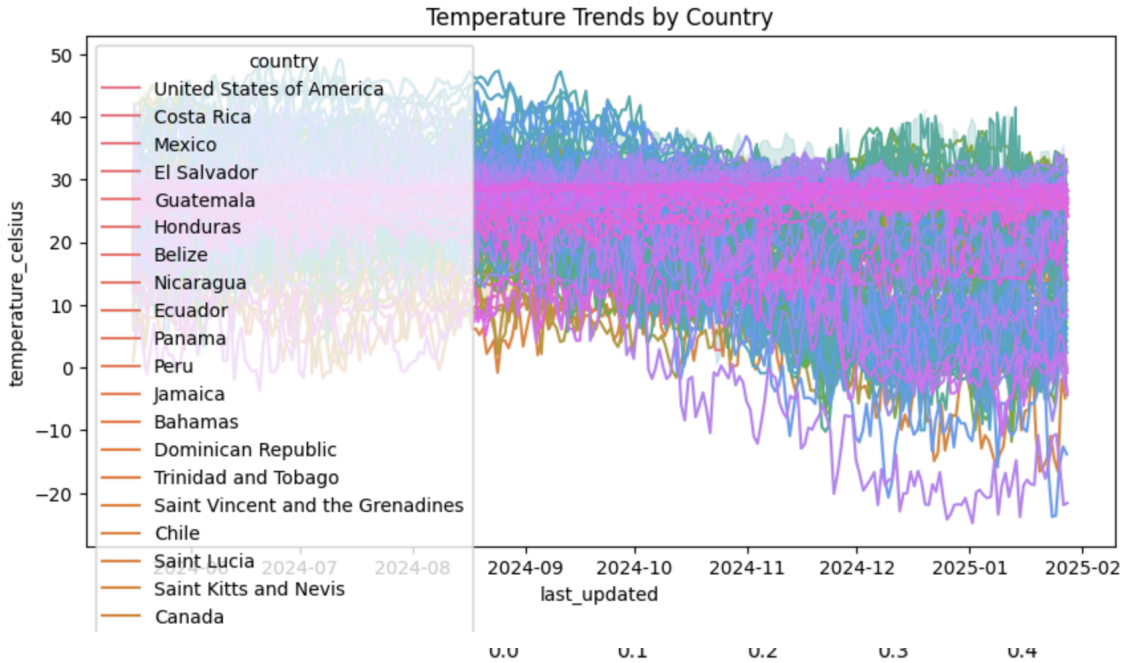


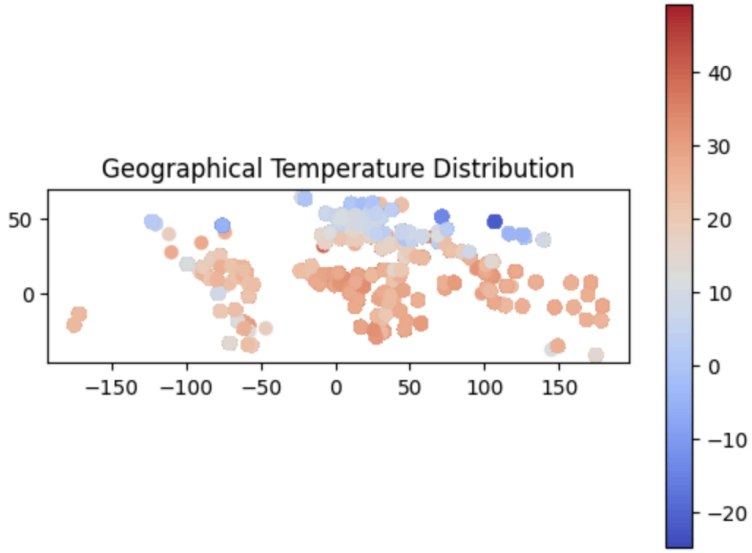

```
Model Performance Metrics:
MAE: 0.6448
RMSE: 0.8305
R2: 0.0043
```

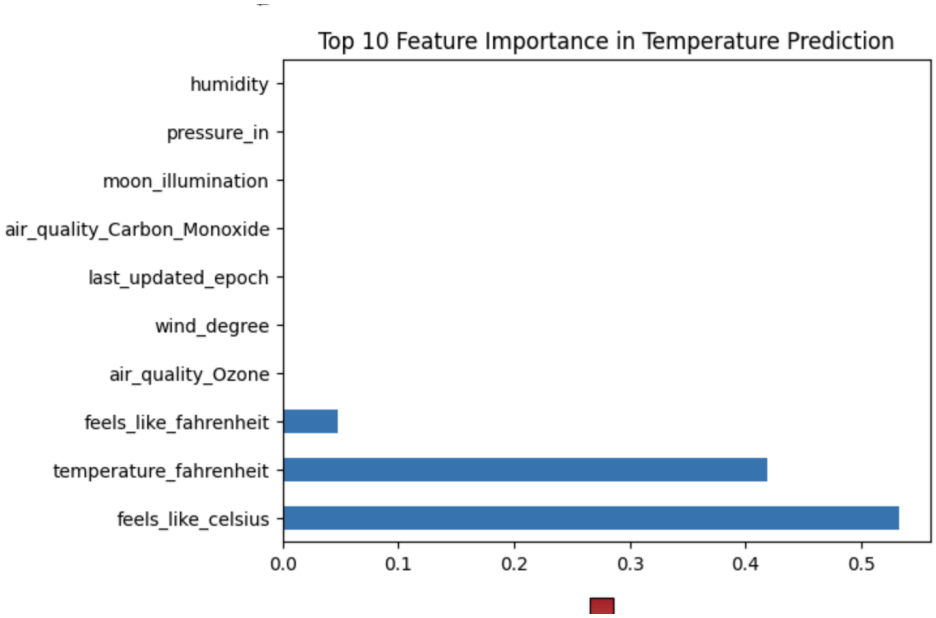
Advanced EDA






Temperature Trends by Country
country
United States of America
Costa Rica
Mexico
El Salvador
Guatemala
Honduras
Belize
Nicaragua
Ecuador
Panama
Peru
Jamaica
Bahamas
Dominican Republic
Trinidad and Tobago
Saint Vincent and the Grenadines
Chile
Saint Lucia
Saint Kitts and Nevis
Canada
temperature_celsius
50
40
30
20
10
0
−10
−20
2024-07
2024-08
2024-09
2024-10
2024-11
2024-12
2025-01
2025-02
last_updated


Geographical Temperature Distribution
50
0
−150
−100
−50
0
50
100
150
40
30
20
10
0
−10
−20


Top 10 Feature Importance in Temperature Prediction
humidity
pressure_in
moon_illumination
air_quality_Carbon_Monoxide
last_updated_epoch
wind_degree
air_quality_Ozone
feels_like_fahrenheit
temperature_fahrenheit
feels_like_celsius
0.0
0.1
0.2
0.3
0.4
0.5