

PROJECT 3

WEATHER DATA ANALYSIS REPORT

Domain: Weather & Climate Analytics

1. Introduction

Weather and climate data play a critical role in decision-making across multiple sectors including agriculture, urban planning, water resource management, transportation, and disaster preparedness. Accurate analysis of historical weather data helps identify long-term trends, seasonal patterns, and extreme events that can have significant social and economic impacts.

This project focuses on analyzing daily weather data to understand temperature trends, rainfall distribution, humidity behavior, and the occurrence of extreme weather conditions. By applying data analysis and visualization techniques, the project demonstrates how raw meteorological data can be transformed into meaningful insights for planning and forecasting purposes.

2. Problem Statement

Weather patterns are becoming increasingly unpredictable due to climate change. Organizations and governments face several challenges, such as:

- Identifying rising temperature trends
- Preparing for irregular rainfall and flooding
- Managing water resources efficiently
- Predicting and responding to extreme weather events

Without systematic data analysis, it becomes difficult to anticipate these changes and take preventive action. This project addresses these challenges by analyzing historical weather data and extracting actionable insights.

3. Objectives

The main objectives of this project are:

1. To analyze temperature trends over time
 2. To study rainfall distribution and variability
 3. To examine humidity patterns and seasonal behavior
 4. To identify extreme weather events such as heatwaves and heavy rainfall
 5. To provide insights useful for environmental planning and risk management
-

4. Dataset Description

The dataset used in this project contains daily weather observations. Each record represents weather conditions for a specific date.

Key Attributes:

- **Date:** Observation date
- **Temperature (°C):** Average daily temperature
- **Rainfall (mm):** Daily rainfall measurement
- **Humidity (%):** Average daily humidity

The dataset captures both continuous and discrete meteorological variables, making it suitable for trend, distribution, and correlation analysis.

5. Data Cleaning and Preparation

Before analysis, the dataset was cleaned and prepared to ensure accuracy and consistency.

5.1 Column Standardization

All column names were converted to lowercase and formatted with underscores to maintain coding consistency.

5.2 Date Conversion

The date column was converted into a datetime format to enable time-series analysis.

5.3 Handling Missing Values

Missing values were handled using forward-fill techniques to maintain continuity in time-series data.

5.4 Feature Engineering

Additional features were created to enhance analysis:

- Month
- Day
- Day of the week

These features helped identify seasonal and temporal patterns.

6. Exploratory Data Analysis

6.1 Temperature Trend Analysis

Temperature data was visualized over time using line plots. The analysis revealed a **gradual upward trend**, indicating increasing average temperatures during the observed period. This pattern may reflect seasonal warming or broader climate influences.

Interpretation:

Rising temperature trends can increase energy demand, affect crop cycles, and contribute to heat-related health risks.

6.2 Rainfall Distribution Analysis

Rainfall data was analyzed using histograms to study its distribution. The results showed that rainfall is **highly uneven**, with:

- Many days receiving little or no rainfall
- A few days experiencing very high rainfall

Interpretation:

Such skewed rainfall distribution increases the risk of flooding while also causing dry spells, which affects agriculture and water storage planning.

6.3 Monthly and Seasonal Patterns

Weather variables were aggregated at a monthly level to identify seasonal trends:

- Higher temperatures during later months
- Increased rainfall and humidity during specific periods

Interpretation:

Seasonal analysis is essential for agricultural planning, reservoir management, and infrastructure design.

6.4 Humidity Analysis

Humidity levels were analyzed using box plots to understand monthly variation. Higher humidity levels were often associated with rainy periods.

Interpretation:

High humidity combined with high temperature can increase heat stress and discomfort, impacting public health.

7. Extreme Weather Event Detection

Extreme weather events were identified using percentile-based thresholds:

- Days with temperatures above the 90th percentile were classified as **high-temperature events**

- Days with rainfall above the 90th percentile were classified as **heavy rainfall events**

Interpretation:

Although extreme events occur infrequently, they have a disproportionate impact on infrastructure, agriculture, and public safety.

8. Correlation Analysis

Correlation analysis was conducted between temperature, rainfall, and humidity:

- Rainfall and humidity showed a positive correlation
- Temperature showed weaker correlation with rainfall

Interpretation:

Understanding these relationships helps in building predictive weather models and early warning systems.

9. Key Insights

1. Temperature shows a gradual increasing trend over time
 2. Rainfall distribution is highly uneven, with few intense rainfall days
 3. High humidity levels often coincide with heavy rainfall
 4. Extreme weather events, though rare, have high potential impact
 5. Seasonal patterns strongly influence all weather parameters
-

10. Practical Implications

The insights from this analysis can be applied in:

- Urban flood risk management
 - Agricultural crop planning
 - Water resource allocation
 - Heatwave preparedness strategies
 - Climate monitoring and forecasting
-

11. Recommendations

1. Authorities should monitor extreme rainfall days for flood preparedness
2. Heat alerts should be issued during periods of high temperature and humidity

3. Seasonal trends should guide agricultural and irrigation planning
 4. Long-term temperature trends should be incorporated into climate adaptation policies
-

12. Limitations

- Dataset covers a limited time period
- Regional variation is not included
- External climate drivers were not analyzed

Future work can expand the dataset and include predictive modeling techniques.

13. Conclusion

This project demonstrates how weather data analysis can uncover meaningful trends, seasonal behaviors, and extreme events. By applying data cleaning, visualization, and statistical analysis techniques, the project provides insights that support environmental planning, risk mitigation, and climate-informed decision-making.