Weather Data Analysis Report

Weather Data Analysis Report

Project Overview

This project analyzes historical weather data to uncover patterns and insights such as monthly and yearly temperature trends, common weather conditions, and extreme weather events. It utilizes Python libraries including Pandas, NumPy, and Matplotlib for data manipulation and visualization.

Dataset Used

- Source: Kaggle - Weather History Dataset
- Link: https://www.kaggle.com/datasets/muthuj7/weather-dataset
- File: weatherHistory.csv
- Size: Over 96,000 records
- Key Columns:
  - Formatted Date
  - Temperature (C) (renamed to Temp_C for clarity)
  - Summary
  - Daily Summary

Key Analyses Performed

1. Data Preprocessing

- Converted the Formatted Date column to datetime format with UTC.
- Extracted Year, Month, and Day from the datetime column.
- Renamed columns for clarity (e.g., Temperature (C) to Temp_C).

Weather Data Analysis Report

2. Monthly Temperature Analysis

- Grouped data by month to calculate average temperatures.

- Created a bar chart showing average monthly temperatures.

3. Yearly Temperature Trends

- Calculated the average temperature for each year.

- Visualized yearly trends using a bar chart.

4. Extreme Temperature Days

- Identified the days with the highest and lowest temperatures.

- Displayed these dates with their corresponding temperatures.

5. Weather Summary Analysis

- Used value_counts() to determine the five most common weather conditions.

- Created a pie chart to visualize their distribution.

6. Seasonal Temperature Analysis

- Mapped months to seasons (Winter, Spring, Summer, Autumn).

- Calculated the average temperature for each season.

- Visualized seasonal trends using a bar chart.

Charts and Visualizations

All visualizations were saved in a charts/ folder using plt.savefig() for organized access and easy reporting.

Weather Data Analysis Report

Tools Used

- Python: For scripting and logic
- Pandas: For data manipulation
- NumPy: For numerical operations
- Matplotlib: For data visualization

Future Improvements

- Implement predictive modeling using Scikit-learn (e.g., future temperature forecasting).
- Apply anomaly detection to identify extreme weather events.
- Include more features such as humidity, pressure, and visibility for deeper insights.