Name : Shyam Keshav rao yerra

Student ID : AF0427835

Batch code : ANP-C8256

Project Name : Weather data analysis

Project Guide : Ms. Maseera Jamal Shaikh



Objective:

The objective of weather data analysis is to extract meaningful insights, patterns, and trends from historical data. Specific goals may include:

**Exploratory Analysis:**

* Understanding weather trends over time (temperature, precipitation, etc.).
* Identifying seasonal or geographic variations**.**

**Data Visualization:**

* Representing data through graphs, charts, and maps for better interpretation.
* Creating graphs or trend lines to showcase patterns.

**Anomaly Detection**:

* Identifying unusual weather events, like extreme temperatures or rainfall.

**Automation**:

* Automating data collection, cleaning, and analysis using Python libraries like Pandas, NumPy and Matplotlib.

Problem statements:

Weather plays a critical role in various aspects of life, including agriculture, transportation, disaster management, and daily decision-making. However, analyzing weather data to identify trends, predict future conditions, and provide actionable insights is often challenging due to the vast and complex nature of meteorological datasets.

This project aims to leverage Python to analyze historical and real-time weather data, addressing the following challenges:

1.Collecting and preprocessing raw weather data from multiple sources, ensuring it is clean, structured, and ready for analysis.

2. Identifying patterns and trends in key weather parameters such as temperature, wind speed, latitude, and longitude over time and across locations.

3. Visualizing weather data to provide intuitive insights for stakeholders.

By implementing this project, we aim to provide a Python-based framework for weather data analysis that is scalable, accurate, and beneficial for real-world applications.

Data Description :

• Source: The dataset was retrieved from Kaggle and contains information about weather data

Data models:

A data model defines the structure and relationships within the weather dataset, ensuring efficient data storage, processing, and analysis.

**1. Data Model Overview**

The data model follows a relational schema, organizing weather data into multiple tables or a structured dataset. For simplicity, attributes can also be stored in a flat file format (e.g., CSV).

**Entities and Attributes**

* **Attributes**

| **Attribute Name** | **Data Type** | **Description** |
| --- | --- | --- |
| City | String | Cities |
| Temperature | Float | Recorded temperature in °C |
| Wind\_Speed | Float | Wind velocity in (m/s). |
| latitude | Float | Latitude of the city. |
| longitude | Float | Longitude of the city |
| Discription | String | Description of weather (e.g., Cloudy, Clear). |
| Country | String | we see city belong to which country. |

**2. Data Relationships**

* **Historical Data vs. Real-time Data**: Historical weather data is used for trend analysis, while real-time data helps in forecasting.
* **Spatial and Temporal Dimensions**: Weather data is linked to specific geographic locations, forming spatiotemporal relationships.
* **Meteorological Variables**: Variables like temperature, wind speed, latitude and longitude are interconnected (e.g., temperature changes affect humidity levels).

Approach

**1. Data Import and Libraries:**

**Libraries Used:** Panda , Matplotlib , Seaborn

**Process:**

* Import relevant libraries.
* Load the dataset and inspect its structure for initial insights**.**

**2. Data Cleaning:**

* checks missing values, duplicates, or any inconsistencies in the dataset.
* Standardize values for clarity.

**3. Exploratory Data Analysis (EDA):**

* Use visualizations (e.g., bar charts, line plots) to identify trends over time and geographical patterns

Project Result :

**1. Data Collection and Preparation**

* **Input Data**: The data could be collected from CSV files. It may include variables like city, temperature, wind speed, latitude, longitude and weather conditions.
* **Cleaning and Preprocessing**: Missing values, duplicates, or outliers might be handled before analysis. This ensures the data is ready for further processing.

**2. Scatter Plot (Temperature vs wind speed)**: Helps analyze the relationship between temperature and **wind speed**.

* Result: Points plotted on a 2D grid could show a linear or random distribution, helping identify correlation patterns.

**3.Bar Chart (city vs temperature)**: Result: A bar chart showing temperature of various city.

**4.Pie chart (Weather condition)**: Helps analyze weather condition of the different city.

**5.Kde plot (temperature distribution and wind speed distribution):**

.

Conclusion: This analysis provides a comprehensive overview of the weather conditions in major cities, highlighting the temperature, wind speed, and other key weather variables. By using visualizations like kde chart, bar chart, scatterplot charts and pie charts .we were able to extract meaningful insights into the data. Further analysis could involve comparing these trends with historical weather patterns or exploring predictive modeling to forecast future weather trends.