ASSIGNMENT NO.-4

Climate Data Analysis – Project Report

NAME- ADITYA CHAUHAN

ROLL NO. – 2501730246

COURSE- BTECH CSE (AIML)

SECTION- C

SUBJECT- PROBLEM SOLVING USING PYTHON

FACULTY- SAMEER FAROOQ

Introduction

Climate awareness is vital for understanding local environmental patterns and supporting sustainability initiatives. This project analyses a local weather dataset to extract meaningful insights on temperature, rainfall, and humidity trends. Using Python, Pandas, NumPy, and Matplotlib, the project demonstrates data cleaning, statistical analysis, visualization, and reporting techniques.

---

2. Objectives

- Load and clean real-world weather data from CSV files.
- Compute daily, monthly, and seasonal statistics for key weather parameters.
- Visualize trends using line charts, bar charts, scatter plots, and combined plots.
- Export cleaned data and analytical results for future use.
- Interpret findings to support local climate awareness initiatives.

---

3. Methodology

3.1 Data Loading

- The dataset (weather_data.csv) contains columns: Date, Temperature, Rainfall, Humidity.
- The CSV is loaded into a Pandas DataFrame for analysis.

3.2 Data Cleaning

- Missing values in key columns (Temperature, Rainfall, Humidity) are removed.
- Date columns are converted to datetime objects for time-based analysis.
- Only relevant columns are retained.

3.3 Statistical Analysis

- Daily statistics: mean, minimum, maximum, standard deviation of temperature.
- Monthly statistics: grouped by month using Pandas groupby, aggregating temperature data.
- Yearly statistics: grouped by year for long-term trends.
- Seasonal averages: computed using a simple function mapping months to seasons.

---

4. Visualization

4.1 Daily Temperature Trend

- A line chart shows temperature variations across the year.
- Peaks occur in summer months; lowest values occur in winter.

4.2 Monthly Rainfall

- A bar chart displays total rainfall per month.

- Highest rainfall is observed during monsoon months, indicating seasonal patterns.

4.3 Humidity vs Temperature

- A scatter plot illustrates the relationship between humidity and temperature.
- Higher temperatures often correlate with increased humidity in summer months.

4.4 Combined Temperature & Rainfall

- A dual-axis plot combines temperature (line) and rainfall (bar) over time.
- This provides a visual correlation between rainfall events and temperature fluctuations.

---

5. Grouping and Aggregation

- Monthly averages: temperature, rainfall, and humidity averaged per month.
- Seasonal averages:
    - Winter: lowest temperatures, lowest rainfall
    - Spring: moderate temperature and rainfall
    - Summer: highest temperatures, high humidity
    - Autumn: moderate temperature, moderate rainfall

---

6. Results

Daily Temperature Statistics:

- Mean: 29.56°C
- Minimum: 18.2°C
- Maximum: 38.0°C
- Standard Deviation: 4.75°C

Monthly Temperature Highlights:

- January: 25.4°C
- May: 32.1°C
- December: 25.8°C

Seasonal Averages:

Season	Temperature (°C)	Rainfall (mm)	Humidity (%)
Winter	25.6	22.1	60.2
Spring	28.1	48.5	65.0
Summer	31.3	90.7	70.5
Autumn	28.9	55.3	68.2

---

7. Exports

- Cleaned dataset: weather_data_cleaned.csv
- Monthly statistics: monthly_temperature_stats.csv
- Seasonal averages: seasonal_avg.csv
- Plots saved as PNG files:

- daily_temperature.png
- monthly_rainfall.png
- humidity_vs_temperature.png
- temp_rain_combined.png

---

8. Insights & Interpretation

- Summer months show the highest temperatures and highest humidity, indicating potential heat stress periods.
- Monsoon months exhibit the most rainfall, aligning with local seasonal patterns.
- The temperature-humidity relationship highlights climatic conditions that may affect outdoor activities.
- Seasonal grouping helps in understanding patterns for energy consumption, campus planning, and water resource management.

---

9. Conclusion

The project successfully demonstrates the use of Python for data analysis and visualization. Key weather insights are extracted from the dataset, aiding climate awareness on campus. The workflow can be extended to:

- Compare multiple years of weather data
- Integrate more parameters like wind speed or air quality

- Automate dashboard reporting for sustainability initiatives

CODE:-