**Weather Data Analysis - Documentation**

**1. Introduction**

This project focuses on analyzing weather data, particularly examining temperature and rainfall trends. Additionally, a predictive model is implemented to estimate rainfall based on temperature variations.

**2. Requirements**

**2.1 Software Requirements**

* Python (>=3.6)
* Required libraries:
  + pandas
  + matplotlib
  + seaborn
  + scikit-learn

**2.2 Installation**

Install the required libraries using the following command:

pip install pandas matplotlib seaborn scikit-learn

**3. Dataset**

* The dataset used for this analysis is weather.csv.
* It contains columns: MinTemp, MaxTemp, and Rainfall.
* Missing values in these columns are removed before proceeding with the analysis.

**4. Steps in the Analysis**

**4.1 Data Loading**

* The dataset is loaded using Pandas.
* Missing values in essential columns (MinTemp, MaxTemp, Rainfall) are dropped.

**4.2 Data Exploration**

* Displays the first few rows of the dataset.
* Prints dataset information, including column names, data types, and missing values.
* Computes statistical summaries (mean, standard deviation, etc.).

**4.3 Data Visualization**

* A pair plot is generated using Seaborn to visualize relationships among MinTemp, MaxTemp, and Rainfall.
* A histogram is plotted to analyze the distribution of MaxTemp and Rainfall.

**4.4 Statistical Analysis**

* Computes overall average values:
  + **Average Max Temperature:** 20.55°C
  + **Average Rainfall:** 1.43 mm

**4.5 Predictive Analysis**

* A linear regression model is trained to predict Rainfall based on MinTemp and MaxTemp.
* The dataset is split into training (80%) and testing (20%) subsets.
* The model is trained and evaluated using the Mean Squared Error (MSE):
  + **MSE:** 37.0768

**4.6 Insights**

* **Highest Rainfall Recorded:** 39.80 mm
* **Lowest Rainfall Recorded:** 0.00 mm

**5. Output and Results**

The analysis results are saved to results.txt containing:

Average Max Temperature: 20.55°C

Average Rainfall: 1.43 mm

Mean Squared Error for Rainfall Prediction: 37.0768

Highest Rainfall Recorded: 39.80 mm

Lowest Rainfall Recorded: 0.00 mm

**6. Execution**

Run the script using:

python weather\_data\_analysis.py

**7. Future Work**

* Enhance predictive accuracy by incorporating additional weather features.
* Explore advanced machine learning models (e.g., Decision Trees, Random Forests).
* Perform time-series analysis to detect seasonal trends in rainfall and temperature.