

Title: Weather Data Analysis

Data Overview

The CSV file contains various weather variables such as MinTemp, MaxTemp, Rainfall, Evaporation, Sunshine, Wind direction and speed measurements, Humidity, Pressure, Cloud cover, Temperature readings, and indicators for rain.

Data Preprocessing

Forward fill operations were performed on the WindDir9am and WindDir3pm variables to handle missing values by filling them with the last observed value. The code to check for null values in specific variables was provided, allowing for data cleaning and preprocessing tasks.

Descriptive Statistics:

- A table was generated displaying the mean, median, and standard deviation for the variables MinTemp, MaxTemp, Rainfall, and Evaporation.
- This table provided a concise summary of the central tendency and dispersion of these key weather variables.

	MinTemp	MaxTemp	Rainfall	Evaporation
mean	7.265574	20.550273	1.428415	4.521858
median	7.450000	19.650000	0.000000	4.200000
Std	6.025800	6.690516	4.225800	2.669383

Distribution of Rainfall

- The distribution of rainfall was visualized using a histogram and a kernel density plot.
- The plots highlighted common rainfall levels around 0 mm and identified potential outliers with higher rainfall values.

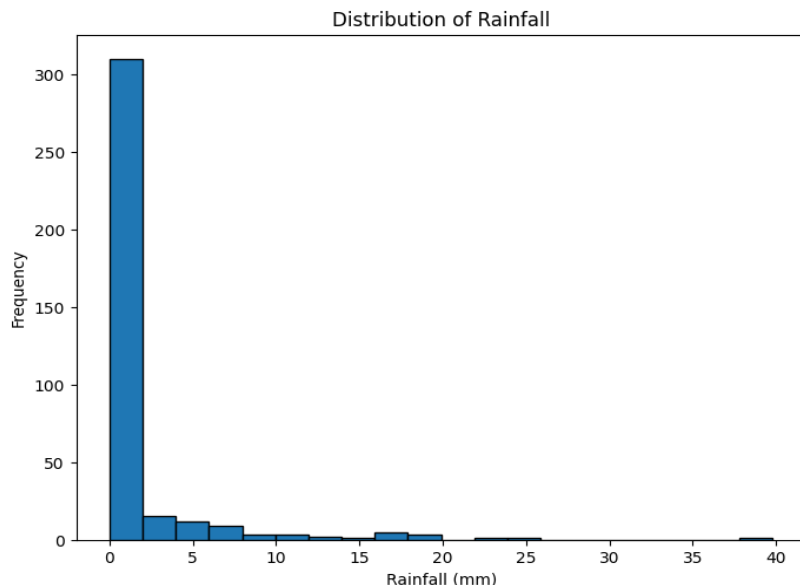


Figure 1: Distribution using Histogram

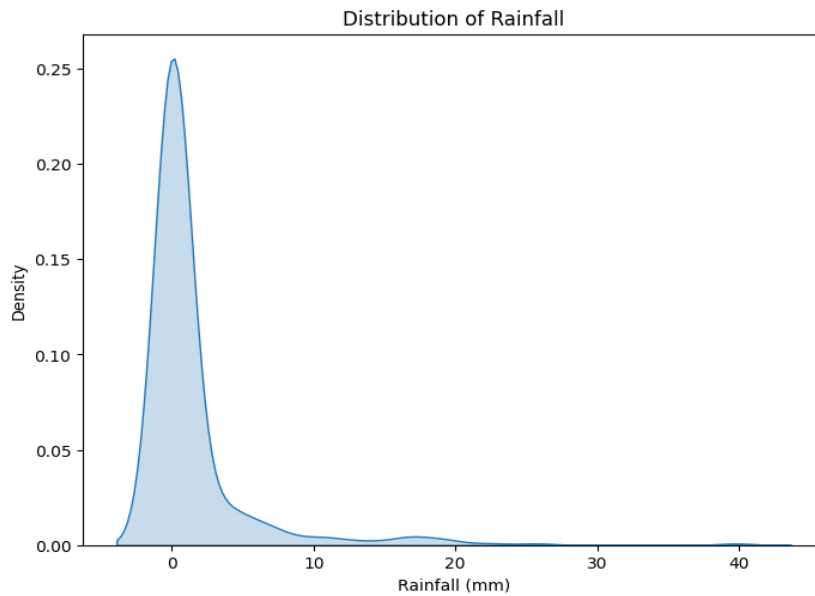


Figure 2: Distribution using Kernel Density

Line Charts

- Line charts were generated to show the variations in weather variables over time, allowing for the identification of trends or patterns.
- These charts provided visual representations of how the variables changed across the dataset.

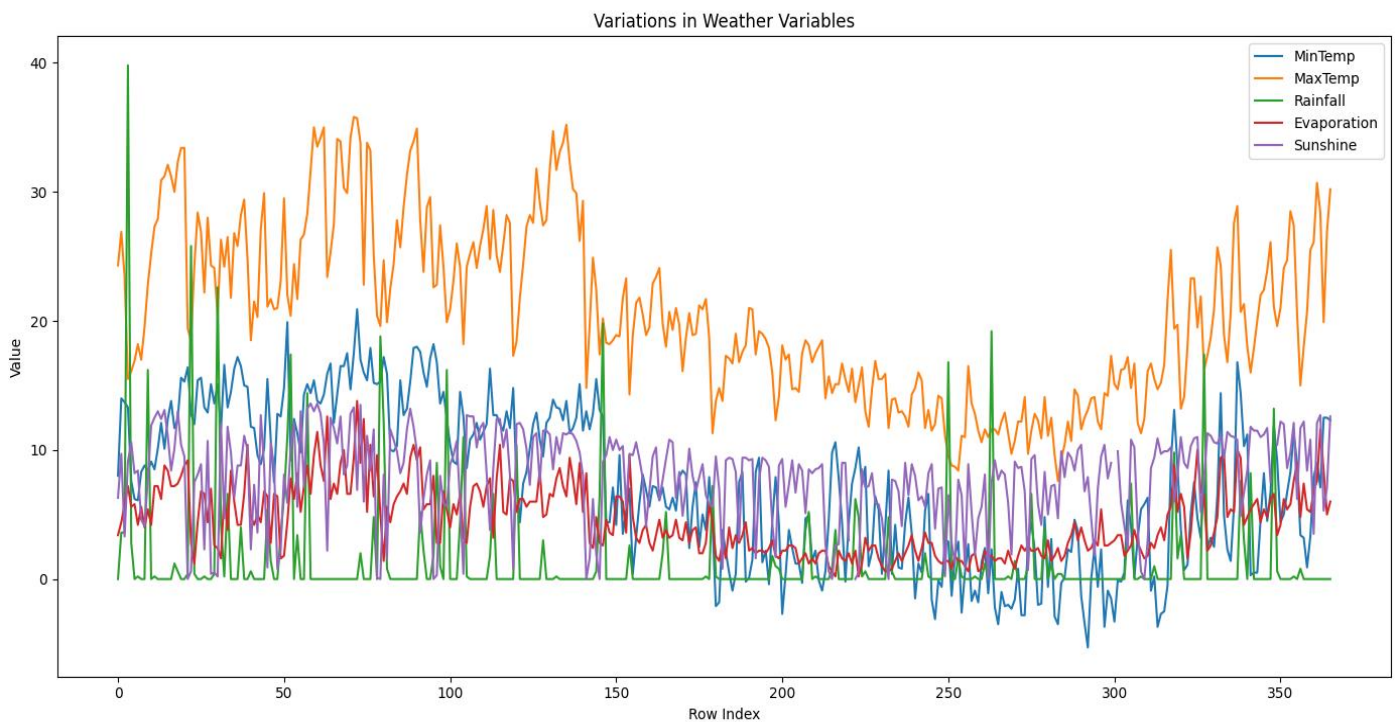


Figure 3: variations in weather variables over time

Correlation Analysis

- A heatmap was created to visualize the correlations between MinTemp, MaxTemp, Rainfall, and Evaporation.
- The heatmap used a color scale to represent the strength and direction of the correlations, helping identify strongly correlated variables.

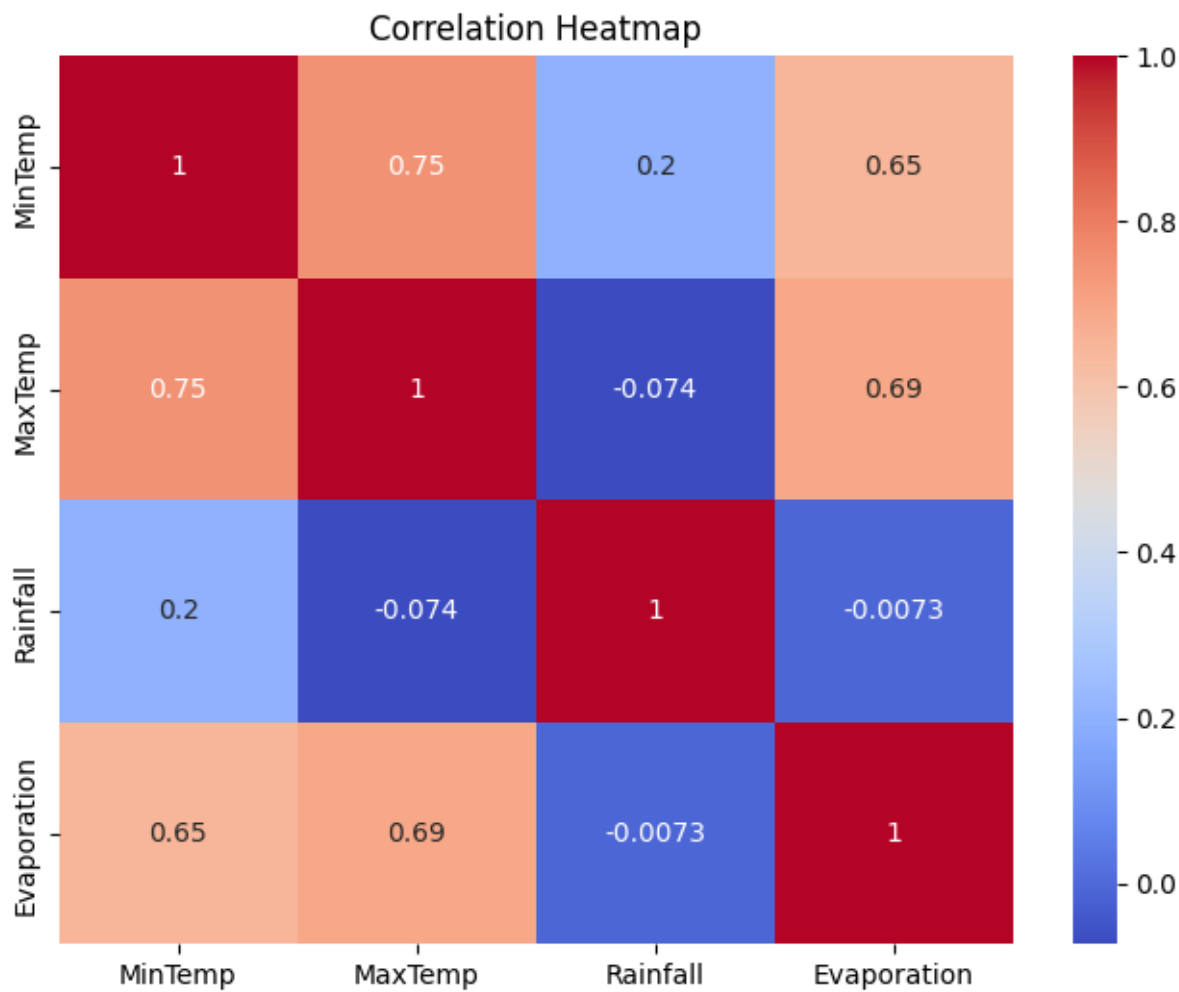
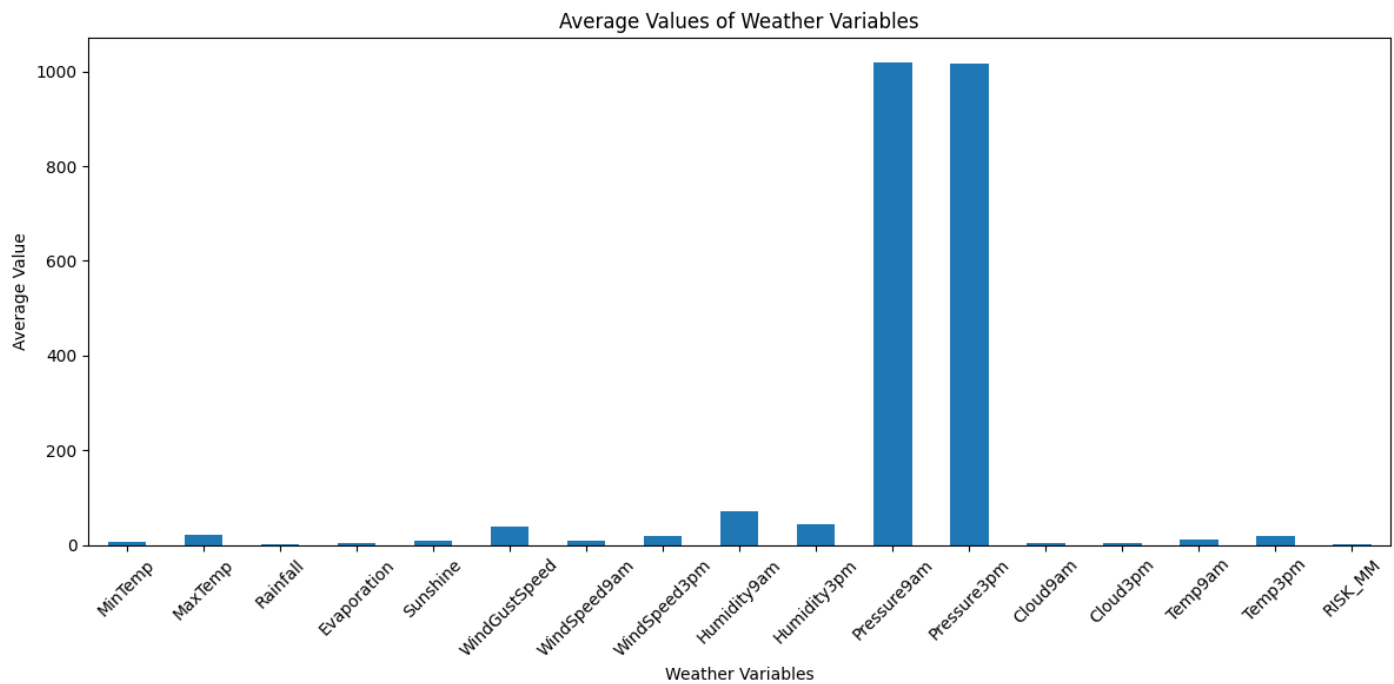


Figure 4: Heatmap

Seasonal Patterns

- Average values of weather variables were calculated across different seasons (Summer, Autumn, Winter, Spring) based on the month data.
- Bar graphs were created to visualize the seasonal patterns, allowing for comparison of average values across seasons and identification of seasonal trends.



Overall, the analysis covered various aspects of the weather data, including data exploration, visualization, summary statistics, data preprocessing, correlation analysis, and trend identification. These insights can be valuable for understanding weather patterns, identifying relationships between variables, and informing decision-making processes related to weather data analysis.