7PAM2000 Applied Data Science 1

Assignment 1: Visualisation

Student Name: VASANTH KUMAR MEKALA

Student ID: 22058159

Course: Applied Data Science with Advanced Research

Submission Date: 10 November 2023

**Task: Visualizations**

***Overview***

The main focus of the study is to apply three types of visualization to extract meaningful information from raw data. To achieve the study objective. The weather data retrieved online from the url: <https://www.kaggle.com/datasets/ananthr1/weather-prediction>.

***Visualizations***

The visualization plots considered were.

1. Line plot

To visualize the distribution of the weather elements over time. The choice of the graphs was motivated by the fact that it can show variation over time.

The Figure 1 shows the daily Precipitation in mm, maximum and minimum daily temperature in degrees Celsius and the wind speed in kilometre. Based on the results there have been irregular spikes in the precipitations over the study duration. However, both minimum and maximum temperatures have seen to show regular cyclic seasonality’s. The, maximum were found to be reached during the august of every year and the lowest were experienced at around Januarys. Noticeably, the wind speed was found to have almost stationary distribution across the study period.

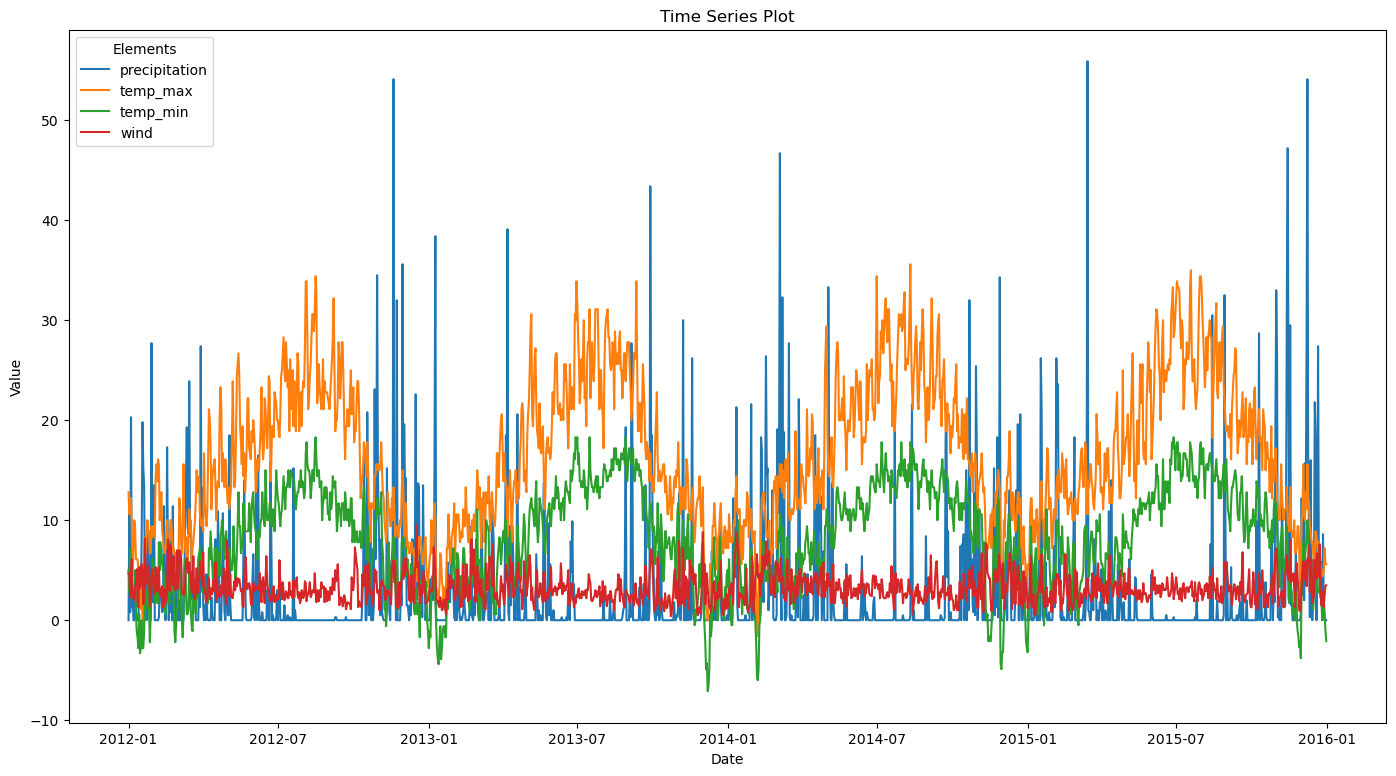


Figure 1 The distribution of the Precipitation, maximum and minimum temperature and the wind speed over time

1. The Scatter plots

The scatter plots are used to show the distribution and association between two numeric features. This forms the main motivation for its application in this study. Besides, the plot can be used to determine the direction and strength of the associations.

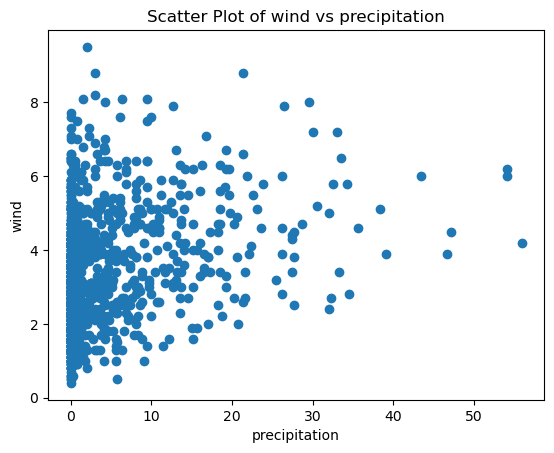
As shown in figure 2, there seems to be no significant association between precipitations and the wind speed. However, there seems to be high variability in the wind speed for low amounts of the precipitation and the variability was found to reduce as the precipitations increase. This moight suggest that under hkigh precipatation the speed of wind tends to be amost equal over time.

Figure 2 The scatter plot showing the association between the wind speed (km) against the precipitations (mm)

1. The pie chart.

The plie chart is used to compare the relative size of a variable under different conditions. The Figure 3 shows the distribution of the maximum temperature against weather conditions. Based on the results there were 27.8% chances of having sun if the temperature was on average, the weather would be 23.4% likely to be foggy , 22.3% likely to be drizzling, 18.8% rainy and 7.8% snow if the maximum temperature was on average.

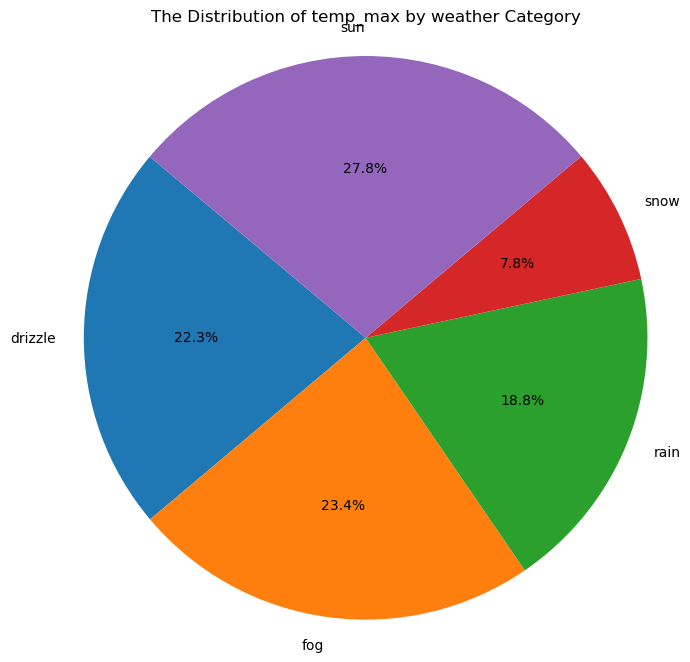


Figure 3 The Distribution of Maximum Temperature by Weather conditions