

## Data analysis of weather data of Pune district

D. C. Vhatkar<sup>1</sup>, P. R. Kolhe<sup>2</sup>, V. D. Jadhav<sup>3</sup>, A. S. Waghmare<sup>5</sup>

<sup>1</sup>M. Tech, College of Agriculture Engineering and Technology, Dr. BSKKV, Dapoli Maharashtra, India

<sup>2</sup>Associate Professor (CAS), College of Agriculture Engineering and Technology,  
Dr. BSKKV, Dapoli Maharashtra, India

<sup>3</sup>Technical Assistant AKMU, Dr. BSKKV, Dapoli Maharashtra, India

<sup>5</sup>M. Tech, College of Agriculture Engineering and Technology, Dr. BSKKV, Dapoli Maharashtra, India

Submitted: 15-08-2023

Accepted: 25-08-2023

### ABSTRACT

Weather is an important aspect in every human's life. The direct impact on human life was seen in recent years. The agriculture industry, the tourism industry and government organization, are dependent on weather in some or another aspects. For humans which are to be ready for unfavourable climate condition, the prior weather knowledge can be highly helpful. Extreme changes in the weather are referred to as climate change. It represents a serious hazard to human life. People are finding challenges to predict the climatic conditions because of the unexpected changes occurring in the weather. Numerous meteorological stations helping to get variables such as temperature, humidity, rainfall, wind speed etc., are crucial in the test of weather conditions. Big data analysis is a study used for sort your data by date, concealed patterns and applicable information that can yield better results. By collecting and analyzing large datasets related to weather data and make data-driven decisions gives knowledge about weather related aspects also the huge amounts of weather data from numerous sensors for different weather characteristics are being collected and studied under big data analysis. The maximum temperature recorded was 43<sup>0</sup>C in 2013 and minimum temperature recorded was 12<sup>0</sup>C in year 2010, 2014 and 2019. The maximum sun hours recorded was 11.25 hours in 2015. The maximum precipitation recorded was 2778.2 mm in 2019. The maximum windspeed recorded was 11.46 kmph in 2019. The maximum humidity recorded was 65.26 % in 2010.

**Key words** – Data analysis, Pune, Temperature, Dashboard

### I. INTRODUCTION

Maharashtra State is situated north of 14° N and south of 22°N in the north of peninsular India. The Arabian Sea borders it on its western side. Gujarat is to its north, while Madhya Pradesh and Chhattisgarh are to its east and north, respectively,

of Maharashtra. Karnataka and Andhra Pradesh form its southern and western borders. The coastal districts of Thane, Mumbai City, Mumbai Suburban, Raigad, Ratnagiri, and Sindhudurg are divided from the rest of Maharashtra by the Western Ghats (Sahyadri), which stretch from north to south. The range is roughly 1 kilometer tall on average.

Pune has a hot semi-arid climate with average temperatures ranging between 19 to 33 °C. Since the rapid expansion of the industries the climate of Pune has changed during the past 3 decades. Pune experiences three seasons in a year viz. summer, monsoon, and winter. Typical summer months are from March to June often extending till 15 June, with maximum temperatures sometimes reaching 42 °C. The hottest month in Pune is roughly between 20 April and 20 May. The summer doesn't end until May, the city often receives heavy dusty winds in May and humidity remains always high. Even during the hottest months, the nights are usually cool it might be due to high altitude of Pune.

The highest temperature ever recorded was 43.3 °C on 30 April 1897. June to October are the monsoon months in Pune with moderate rainfall and temperatures ranging from 22 to 28 °C. Most of the 722 mm of annual rainfall in the city falls between June and September, and July is the wettest month of the year.

Winter traditionally begins in November and cold can be experienced typically during the festival of Diwali. The daytime temperature is around 26 °C while night temperature is below almost 9 °C for most of December and January. The lowest temperature ever recorded was 1.7 °C (35 °F) on 17 January 1935.

### II. REVIEW OF LITERATURE

Krishna 2015 stated that finding new knowledge might be difficult when the data being gathered is dense with irrelevant information. The processes of data preparation includes cleaning,

normalizing, transforming, choosing features, and selecting the data.

Mishra and Krishna, 2017 reported that a time-series is a collection of data that is recorded on an hourly, daily, weekly, monthly, and annual basis. In order to predict future events in fields like climate, education, stock market, and other sectors, data mining algorithms use time-series data that has been gathered in this way over time.

A. Gayathri et al., 2016 investigated data mining for weather forecasting. There are many different types of forecasting, including short-term, medium-term, and long-term forecasting. It is possible to extract value from several types of weather big data, which can be described as a huge range of factors and distinct data mining methodologies, by applying techniques like decision trees and reverse analysis. Big data is more often used while predicting the weather.

Veershettyet al. 2015 worked on creating a framework for utilizing Hadoop to analyse weather data. Temperature and yearly precipitation served as the meteorological factors for extraction and analysis. Performance comparisons between Pig and Hive for weather data are shown. It has been shown that HIVE's efficacy results in better outcomes. On a Hadoop cluster, the recommended analytical engine scales more successfully.

### III. MATERIAL AND METHOD

#### A. Materials

Despite not being traditionally regarded as a Big Data Analytics tool, Microsoft Excel can nonetheless perform some big data tasks, particularly with the introduction of Power Query

and Power Pivot. Excel can now connect to, clean up and analyse many more datasets, than it was previously able to. Excel has limits when working with large datasets, though, as compared to specialised Big Data tools like Hadoop or Spark. Here are few examples of how Excel can be applied to Big Data Analytics: Power Query, Power Pivot, Data Analysis and Data Visualization.

#### B. Methodology

Performing big data analytics on weather data involves several steps to collect, process, analyze, and interpret the data effectively. Here's a generalized process you can follow:

1. Define Objectives: Clearly outline the objectives of your analysis, such as understanding at current conditions, wind and clouds, pressure changes, rainfall data, humidity temperature to determine how the atmosphere involves in the future. What is the problem face in future?
2. Data Collection: The kaggle website was utilised to gather the data for this project. The information covered the 30 -month period from May 2020 to December 2022.
3. Data Cleaning and Preprocessing: It is important pre-processing step in data mining process. In order to enhance the quality of the data and to produce better results. The characteristics are examined and any null values are eliminated. Duplicate records are also eliminated and outliers are imputed.
4. Data Selection: At this point, the dataset's data that would be useful for the analysis was selected. The meteorological dataset contained attributes. The meteorological dataset had nine (9) attributes, their type and description in represented in Table 1.

Table 1: Attributes of Metrological Dataset

Attribute	Type	Description
Year	Numerical	Year Considered
Month	Numerical	Month Considered
Min. Temp.	Numerical	The Minimum Temperature
Max. Temp.	Numerical	The Maximum Temperature
Wind Speed	Numerical	Average Windspeed in km/hr
precipitation	Numerical	The yearly Precipitation
Humidity	Numerical	The yearly Humidity
Sun hours	Numerical	The yearly sun hours (hours)

5. Data Integration, Exploration, Visualization and Analysis: It integrate data from multiple sources into a centralized data repository making it accessible for further analysis. Explore the data through descriptive statistics and visualization to pressure changes, rainfall data, humidity and temperature. Apply various data analytics techniques, such as clustering, classification, regression to extract meaningful pattern and trends from the weather data.

#### IV. RESULT AND DISCUSSION

In this chapter the results which are obtained from the raw weather data of Pune district is discussed. The discussion and results is included of maximum and minimum temperature, yearly precipitation, sun hours, humidity, wind speed which are showed in the form of pivot table charts.

According to pivot table generated form the raw weather data the results are recorded are as follows:

##### 1. Temperature

The fluctuation of temperature is seen in several years the bar chart is shown below which gives information of maximum and minimum temperature form 2009 to 2021.

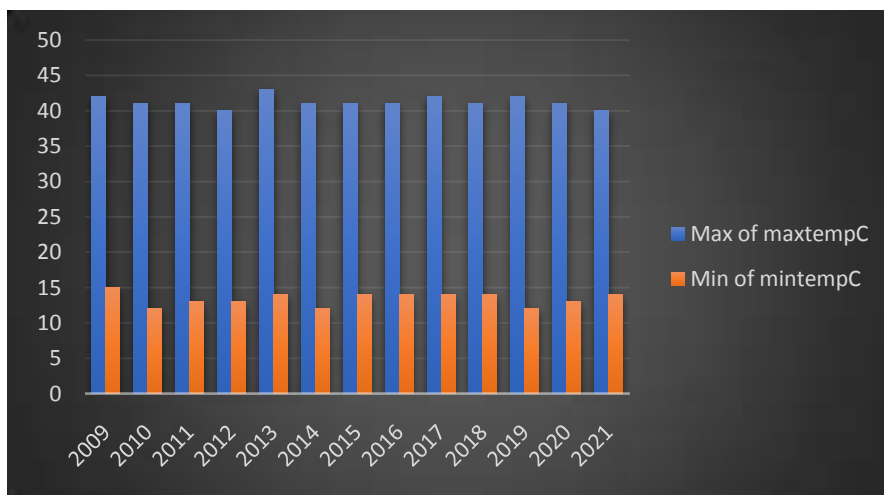


Fig No.4: Pivot table of maximum and minimum temperature

##### 2. Sun hours

The amount of total sun hours yearly is recorded and shown in the chart below.

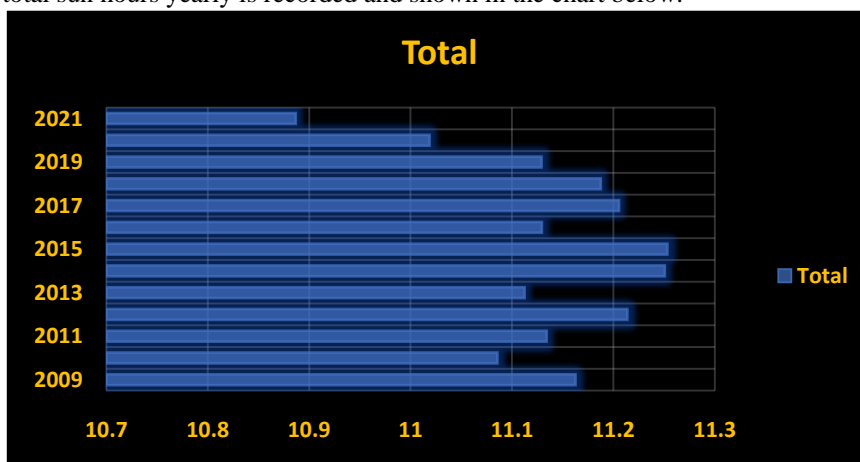
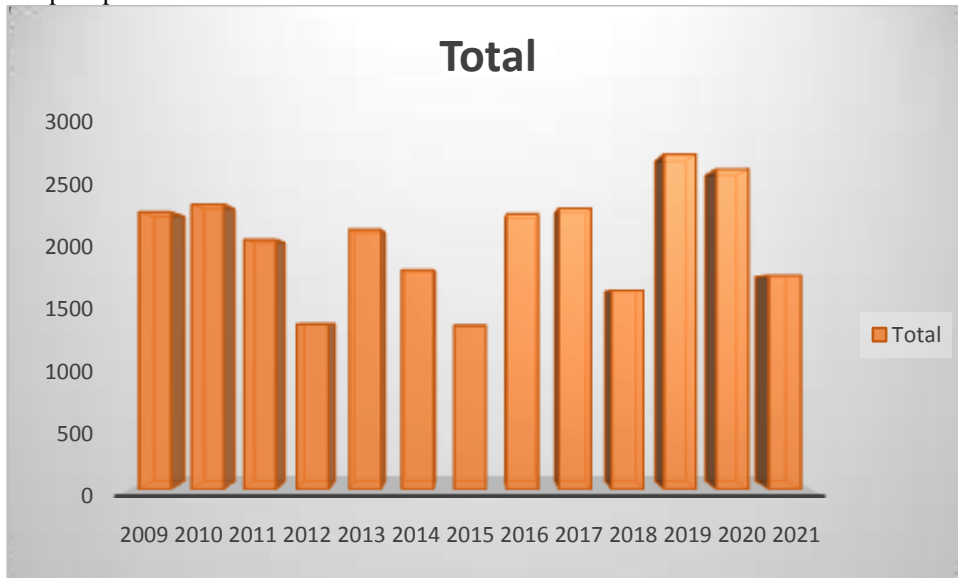


Fig No.5: Pivot table of Sun hours

### 3. Precipitation

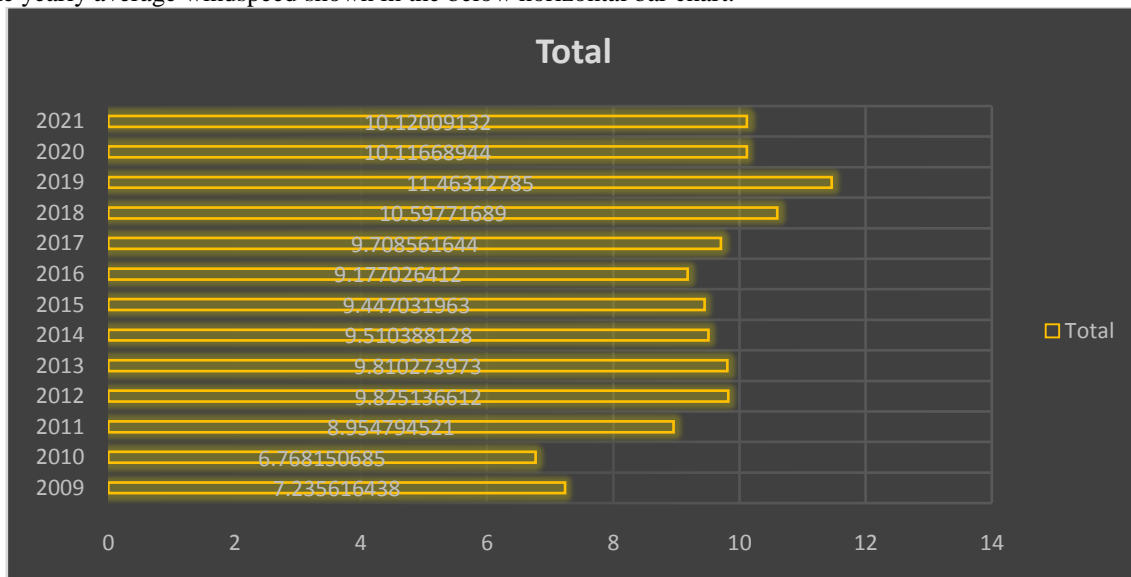
The amount of precipitation that means rainfall is shown in the below bar chart.



**Fig No.6: Pivot table of Precipitation**

### 4. Windspeed

The yearly average windspeed shown in the below horizontal bar chart.



**Fig. No.7: Pivot table of Windspeed**

### 5. Humidity

The yearly average humidity shown in the pie chart below.

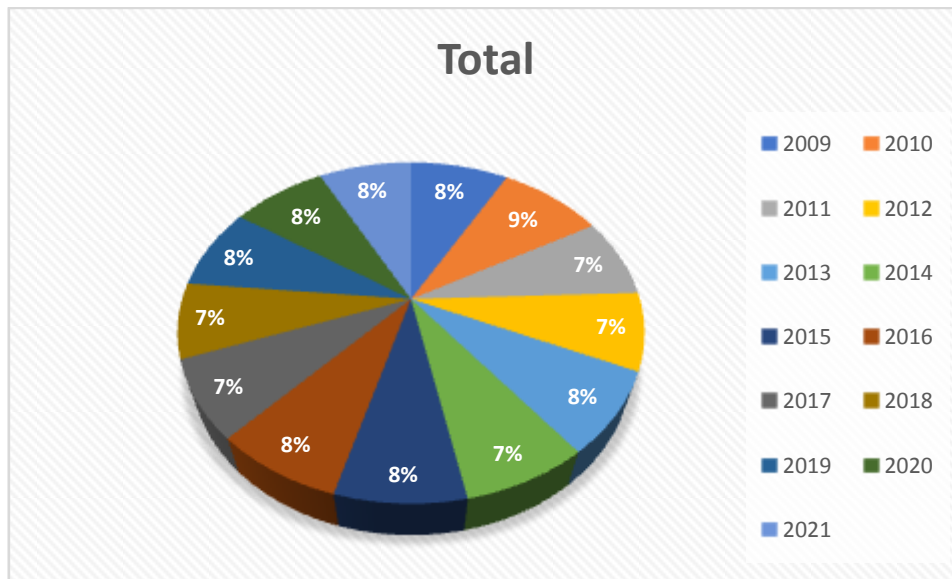


Fig. No.8: Pivot table of Humidity



Fig No.9: Dashboard of Weather Data



Fig. No.10: Slicer

## V. CONCLUSION

- [1]. The maximum temperature recorded was 43<sup>0</sup>C in 2013 and minimum temperature recorded was 12<sup>0</sup>C in year 2010, 2014 and 2019.
- [2]. The maximum sun hours recorded was 11.25 hours in 2015.
- [3]. The maximum precipitation recorded was 2778.2 mm in 2019.
- [4]. The maximum windspeed recorded was 11.46 kmph in 2019.
- [5]. The maximum humidity recorded was 65.26 % in 2010.

## REFERENCES

- [1]. A Gayathri, M. Revathi, J. Velmurugan, (2016). A Survey on Weather forecasting y Data Mining, IJARCCCE, ISSN: 2278-1021, 5(2), 298-300.
- [2]. Climate of Maharashtra, National Climate Centre Office of the Additional Director General of Meteorological, Pune.
- [3]. Krishna GV, (2015), A review of weather forecasting models-based on data mining and Artificial neural networks, Computer Science and engineering, GITAM university, 6,214-222.
- [4]. Microsoft (n.d.-b). Power Pivot: Powerful data analysis and data modeling in Excel. Retrieved from <https://support.office.com/en-us/article/Power-Pivot-Powerfuldata-analysis-and-data-modeling-in-Excel-A9C2C6E2-CC49-4976-A7D7-40896795D045>.
- [5]. N. Mishra, H. K. Soni, S. Sharma, and A. K. Upadhyay, (2017), A Comprehensive Survey of Data Mining Techniques on Time Series Data for Rainfall Prediction, J. ICT, 11, 168.
- [6]. VeershettyDagade, Mahesh Lagali, SupriyaAvadhani and PriyaKalekar, (2015). Big Data Weather Analytics Using Hadoop, IJETCSE, ISSN: 0976-1353,14(2).