



Project Title

Predictive Analysis of Temperature and Regional Comparisons of weather in India

Sardar Patel University

Vallabh Vidyanagar

Anand

Guided by

Dr. (Mrs.) Khimya Tinani

Submitted by

Ganesh Tatoba Bhosale

Tejas Atmaram Sawant

MSc Statistics (semester -4)

Content

1. Introduction

2. Objectives

3. Review of literature

4. What data we have

5. Graphical Analysis

6. Statistical Analysis

7. Conclusion

8. References

INTRODUCTION

Weather forecasting is a critical aspect of our daily lives, influencing decisions ranging from what to wear to how to plan transportation and outdoor activities.

This project aims to delve into the fascinating world of weather forecasting, employing data-driven approaches to enhance prediction accuracy and reliability. By leveraging advanced technologies and analytical techniques, the project aims to contribute to the improvement of weather forecasting models.

The project will focus on analysing historical weather data of Ahmedabad city in Gujrat to identify patterns and trends. The dataset will encompass a wide range of variables, including temperature fluctuations, precipitation levels.

For the region wise comparison, we have Data from August 29 2023 to till now, this contains the wide range of variables like Temperature, Humidity, Participation level, Air quality and much more mentioned below

Timely and accurate weather predictions help in preparing for and responding to natural disasters such as hurricanes, tornadoes, floods, and wildfires. Evacuation plans and emergency response strategies can be implemented based on forecasted weather conditions.

Farmers rely on weather forecasts to plan their planting and harvesting schedules. Knowing upcoming weather conditions helps optimize irrigation, pest control, and other agricultural practices, ultimately affecting crop yield and quality.

Weather forecasting is the practice of predicting atmospheric conditions for a specific time and location. It utilizes various scientific methods, data analysis techniques, and computer models to provide insights into future weather patterns. By examining factors such as temperature, humidity, air pressure, wind speed, and precipitation, meteorologists aim to generate accurate forecasts to help individuals, businesses, and governments make informed

decisions regarding activities, travel plans, agriculture, and disaster preparedness. Weather forecasting has evolved significantly with advancements in technology, allowing for more precise predictions over shorter and longer timeframes, ultimately contributing to improved safety and planning in our daily lives.

We use a weather dataset received from Ahmedabad geological station for fitting a weather prediction model and check whether which model is fit more accurately. And use a several parameters in Indian dataset to test the weather conditions in different states in India. Is it similar or not.

OBJECTIVES

- **Interrelationships:** Conduct a nuanced analysis of correlations between different weather variables, providing a holistic view of how these factors interact.
- **Temporal Dynamics:** Investigate temporal trends and cyclic patterns in meteorological variables to comprehend the seasonality of weather.
- **Regional Disparities:** Uncover regional variations in weather conditions.
- **Accurate Temperature Predictions:** Develop reliable models for predicting temperature trends
- **To identify impact of air quality on temperature and other weather factors.**
- **Conduct exploratory analysis to identify trends, seasonality, and other patterns in the weather data, using visualizations and statistical techniques to gain insights into the data's characteristics.**
- **Select appropriate time series forecasting models such as ARIMA, SARIMA, LSTM and implement them using Python or R programming languages.**

REVIEW OF LITERATURE

Nikhil Sethi et al (2014):

Used the multiple linear regression technique for predict temperature. He used Attributes for predicting weather are Rainfall, Vapour pressure, Average temperature and cloud cover

Ankita Joshi (2015):

She used Decision tree algorithm for prediction and used variables Max temperature, Rainfall, Evaporation and Wind Speed

WHAT DATA WE HAVE

We have a dataset of Ahmedabad city for predicting a temperature and rainfall containing the variable minimum temperature (TMIN), maximum temperature (TMAX), precipitation level (P RTP) and Date

From the comparison of weather region wise we have another dataset given from Kaggle. This contain several variable mentioned in next slide.

Location Name: Name of the specific location.

- Information: Identifies the exact place where weather measurements are taken.
- Measurement Instrument: N/A (Categorical variable).

3. Region: Region or state in India.

- Information: Specifies the state or region within India.
- Measurement Instrument: N/A (Categorical variable).

4. Latitude: Latitude coordinates of the location.

- Information: Geographical coordinate that specifies the north-south position.
- Measurement Instrument: GPS or satellite-based positioning systems.

5. Longitude: Longitude coordinates of the location.

- Information: Geographical coordinate that specifies the east-west position.
- Measurement Instrument: GPS or satellite-based positioning systems.

6. Time zone: Time zone of the location.

- Information: The time zone in which the location is situated.
- Measurement Instrument: N/A (Categorical variable).

7. Last Updated Epoch: Timestamp in epoch format.

- Information: Time of the last update in epoch format (seconds since 1970-01-01 00:00:00 UTC).

- Measurement Instrument: N/A (Temporal variable).

8. Last Updated: Timestamp in a readable format.

- Information: Time of the last update in a human-readable format.

- Measurement Instrument: N/A (Temporal variable).

9. Temperature (Celsius, Fahrenheit): Current temperature.

- Information: Measurement of the current air temperature.

- Measurement Instrument: Thermometer.

10. Condition Text: Description of weather conditions.

- Information: Textual description of the current weather conditions.

- Measurement Instrument: N/A (Categorical variable).

11. Wind Speed (mph, kph): Wind speed.

- Information: Speed of the wind.

- Measurement Instrument: Anemometer.

12. Wind Degree: Wind direction in degrees.

- Information: The direction from which the wind is blowing in degrees.

- Measurement Instrument: Wind Vane.

13. Wind Direction: Wind direction.

- Information: Cardinal or ordinal direction from which the wind is blowing.

- Measurement Instrument: Wind Vane.

14. Pressure (mb, in): Atmospheric pressure.

- Information: Pressure exerted by the atmosphere.

- Measurement Instrument: Barometer.

15. Precipitation (mm, in): Amount of precipitation.

- Information: Quantity of water (rain, snow, etc.) fallen or falling.

- Measurement Instrument: Rain Gauge.

16. Humidity: Humidity percentage.

- Information: The amount of moisture in the air.

- Measurement Instrument: Hygrometer.

17. Cloud: Cloud cover percentage.

- Information: The fraction of the sky obscured by clouds.

- Measurement Instrument: Ceilometer or satellite imagery.

18. Feels Like Temperature (Celsius, Fahrenheit): "Feels like" temperature.

- Information: Subjective measure of how the weather feels.

- Measurement Instrument: N/A (Derived from other variables).

19. Visibility (km, miles): Visibility.

- Information: Distance one can see clearly.

- Measurement Instrument: Visibility Sensor.

20. UV Index: UV index value.

- Information: Intensity of ultraviolet (UV) radiation in sunlight.
- Measurement Instrument: UV Index Sensor.

21. Gust Speed (mph, kph): Gust speed.

- Information: Sudden increase in wind speed.
- Measurement Instrument: Anemometer.

22. Air Quality Parameters (CO, O3, NO2, SO2, PM2.5, PM10): Concentrations of air pollutants.

- Information: Levels of different air pollutants in the atmosphere.
- Measurement Instrument: Air Quality Monitoring Stations.

23. Air Quality Indices (US EPA Index, GB DEFRA Index): Air quality indices.

- Information: Aggregated indices representing overall air quality.
- Measurement Instrument: Air Quality Monitoring Stations.

24. Sunrise, Sunset, Moonrise, Moonset: Timings for celestial events.

- Information: Timing of sun and moon-related events.
- Measurement Instrument: N/A (Derived from astronomical calculations).

25. Moon Phase: Current moon phase.

- Information: Appearance of the moon from Earth.
- Measurement Instrument: N/A (Derived from astronomical calculations).

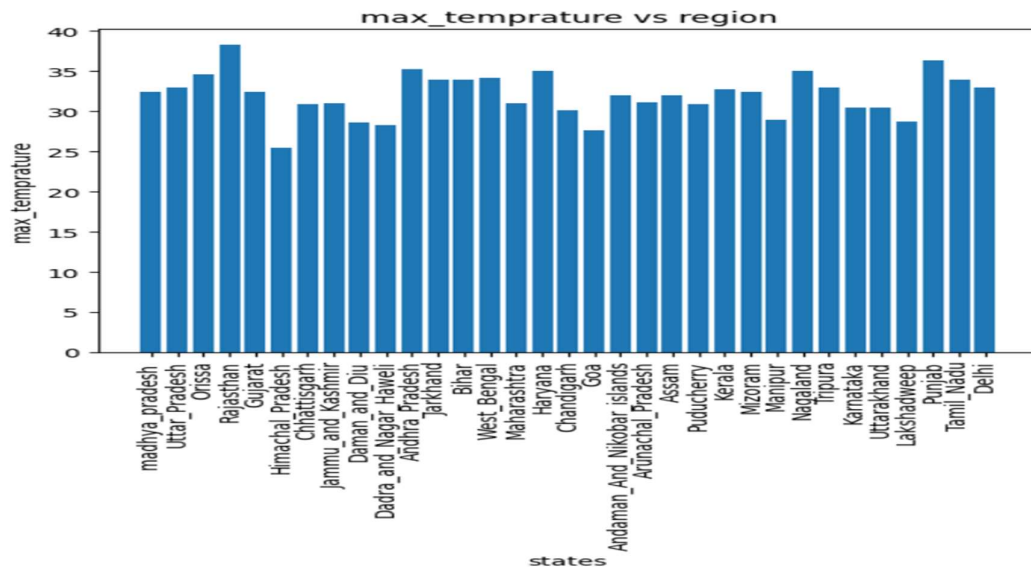
26. Moon Illumination: Illumination percentage of the moon.

- Information: Percentage of the moon's visible surface illuminated by the sun.

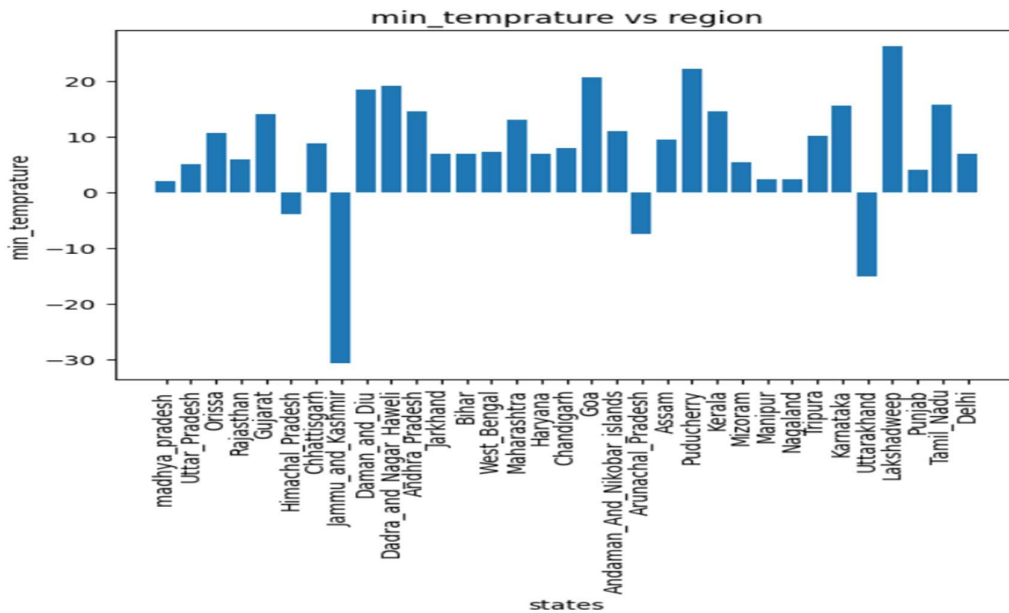
- Measurement Instrument: N/A (Derived from astronomical calculations).

These variables collectively provide a rich set of information about the weather, air quality, and astronomical conditions, measured using various instruments designed for each specific parameter.

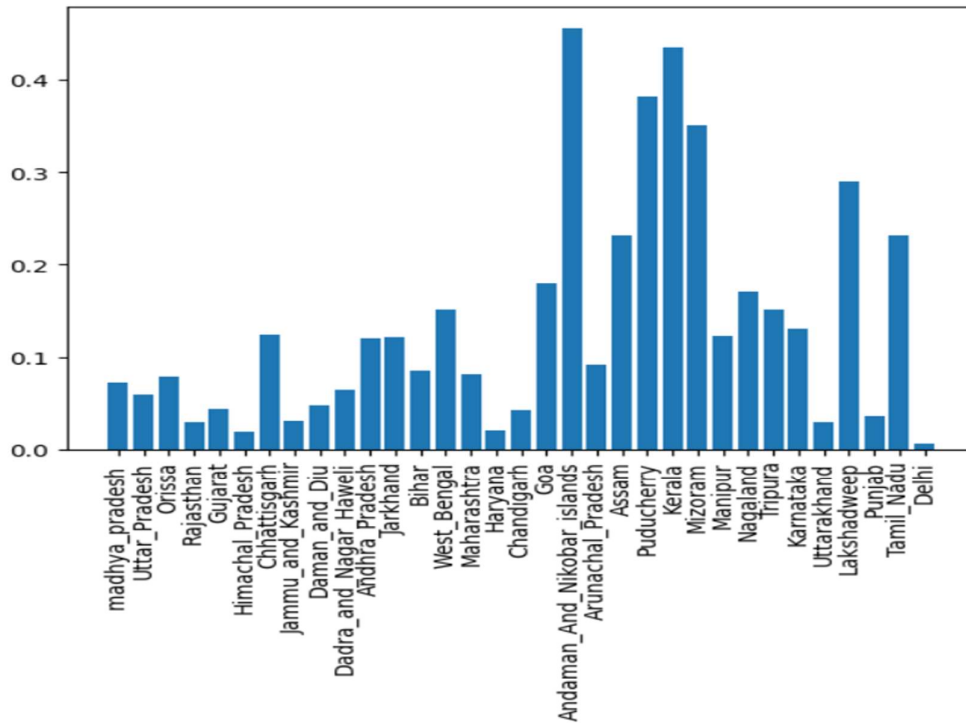
GRAPHICAL ANALYSIS



From this bar plot we compare the maximum temperature in the States. As we can see that the temperature of Rajasthan is higher than other states since 29 August 2023.



From this bar plot we compare the minimum temperature in the States. As we can see that the temperature of Jammu and Kashmir is lower than other states since 29 August 2023.



From this bar plot we can see that the average rain in the states. As we can see that the maximum rain occurred in Andaman and Nikobar islands

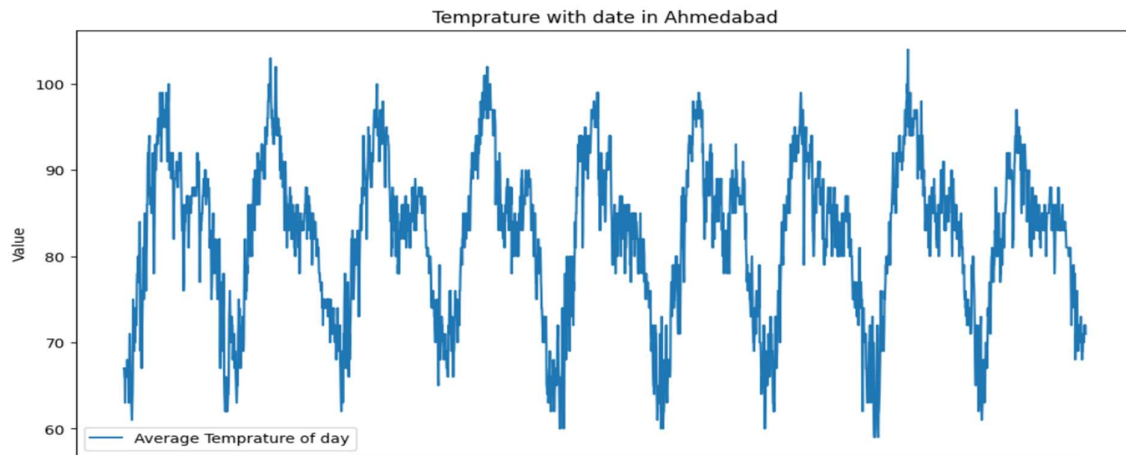
STATISTICAL ANALYSIS

Correlation matrix:

This correlation matrix shows the relationship between temperature, wind speed, participation level, humidity and cloud. Correlation coefficient varies to -1 to 1. If there is a positive correlation then correlation coefficient is positive and value near to 1. If there is a Negative correlation then correlation coefficient is negative and value near to -1.

	temp	wind	precip	humidity	cloud
temp	1	0.221136	0.097628	0.294745	0.27211
wind	0.221136	1	0.081938	-0.01195	0.131994
precip	0.097628	0.081938	1	0.1781	0.269386
humidity	0.294745	-0.01195	0.1781	1	0.542511
cloud	0.27211	0.131994	0.269386	0.542511	1

We wish to check the temperature of Ahmedabad city is stationary or not? For the testing we apply the ADF test and plot the graph with time to Temperature.



ADF Statistic:- 3.768906681682775

p-value for TAVG: 0.00324195092487071

H0 : Data is not stationary.

H1 : Data is stationary.

From the graph we can say that, the data is stationary Now for confirmation we apply ADF Test. Here from ADF test P-Val < 0.05 hence we reject H0 Hence the data is stationary.

CONCLUSION

From the data of Indian weather the maximum temperature is in Rajasthan. And the minimum temperature is noted in Jammu and Kashmir.

The maximum average rainfall happen in Andaman and Nikobar islands

From meteorological parameter we can say that there is maximum relation between cloud and humidity.

From the plot of Time to temperature we can say that our data is stationary. For conform we apply ADF test P-Val < 0.05 hence we can say that our data is stationary.

REFERENCES

- <https://www.ncdc.noaa.gov/cdo-web/search>
- <https://www.kaggle.com/datasets/nelgiryewithana/indian-weather-repository-daily-snapshot>
- Book: Fundamentals of Mathematical Statistics | Author: S.C. Gupta & V.K. Kapoor | Edition: Twelfth edition | Publication: Sultan Chand & Sons.