

## Team:- Decimal Binary (EN43) (Round 2)

### Solution Planning Document

**Problem Statement:** To analyse and predict weather conditions, such as temperature and rainfall, for Maharashtra's 16 cities using a dataset of December's weather data comprising 400+ queries. The solution will help in planning for agricultural or travel needs.

---

- **Background:**

- Weather prediction plays a vital role in various sectors, including agriculture and travel. By analysing historical weather data, actionable insights can be derived to improve decision-making.

Preprocessing:-

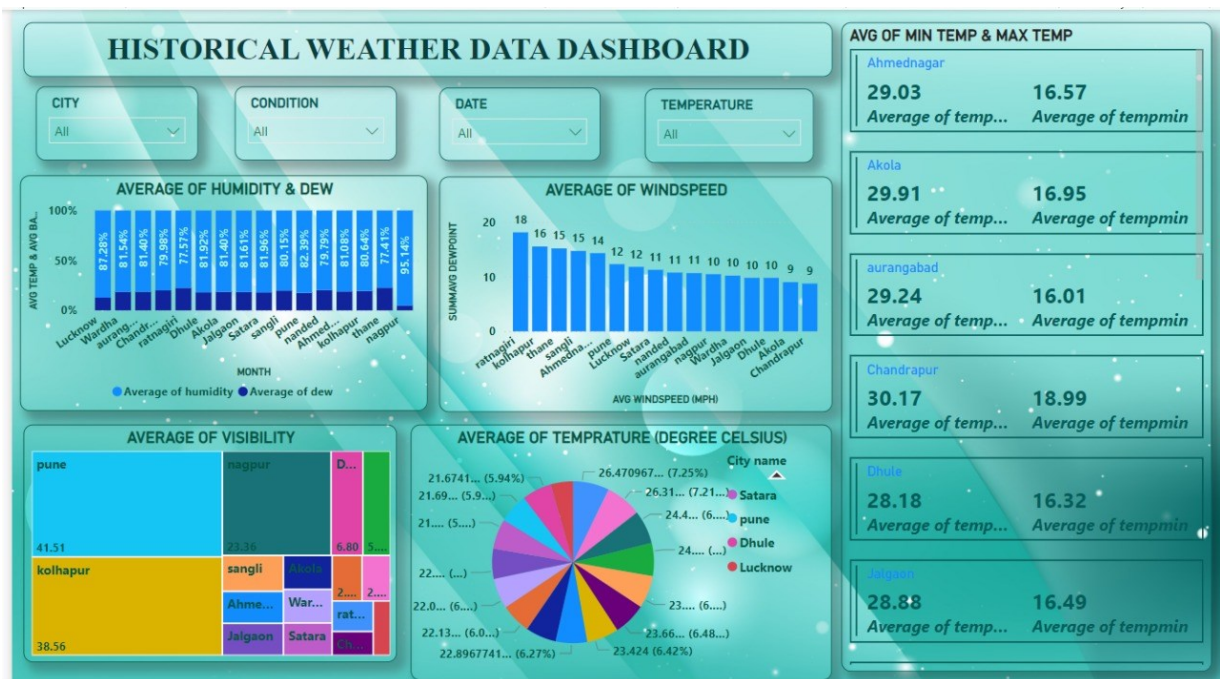
```
[21]: print(df.isnull().sum())
      # Check for missing values
      df = df.dropna()

cityID          0
datetime        0
tempmax         0
tempmin         0
temp            0
feelslikemax    0
feelslikemin    0
feelslike       0
dew             0
humidity        0
precip          0
precipprob      0
precipcover     0
windgust        0
windspeed       0
winddir         0
sealevelpressure 0
cloudcover      0
visibility      8
solarradiation  0
solarenergy     0
uvindex         0
conditions      0
description     0
icon            0
stations        14
dtype: int64

[22]: df.to_csv('Total_weather_data.csv', index=False)
```

In this step we have checked for the missing values and drop that rows for the same .

- **Scope:**
  - Includes: Data preprocessing, exploratory data analysis (EDA), and predictive modeling.
  - Normalize and scale features.
  - Excludes: Real-time data integration and predictions for other months or states.
- **Libraries used for preprocessing:**
  - Pandas
  - NumPy
  - Matplotlib
  - Seaborn
  - sklearn (for preprocessing)
- **Visualization Using Power-BI**



**Build visual tools used in PowerBi dashboard:**

1. Slicer
2. Stacked column chart
3. Line and Clustered Column Chart
4. Multi Row Card
5. Pie Chart
6. Tree Map

Using this tools we have compared the average values of Temperature, Windspeed ,Due & Humidity to visualise the weather Condition of different cities .

## • Solution Plan

### Step 1: Problem Definition

**Goal:** Predict weather description (e.g., "Partly cloudy throughout the day") based on the provided weather features:

tempmax, tempmin, temp, feelslikemax, feelslikemin, feelslike, dew, humidity, snowdepth, windgust, windspeed, winddir, sealevelpressure, cloudcover, visibility, solarradiation, solarenergy, uvindex.

### Step 2: Data Preparation

**Load the Dataset:**

Load the CSV file and examine its structure.

**Feature Selection:**

Use provided weather-related features as input.

Encode description into numeric labels using LabelEncoder.

**Train-Test Split:**

Split the data into training (80%) and testing (20%) sets to evaluate model performance.

**Step 3: Model Training**

**Algorithm:** Use a Random Forest Classifier as it handles both numerical and categorical data well and is robust against overfitting.

**Hyperparameters:** Use default hyperparameters for simplicity; tune later if needed.

**Step 4: Model Evaluation**

Evaluate the model using the following metrics:

**Accuracy:** Percentage of correct predictions.

**Precision:** Proportion of true positives among predicted positives.

**Recall:** Proportion of true positives among actual positives.

**F1-Score:** Harmonic mean of precision and recall.

---

- **Output :-**

1. Process and clean the given dataset.
2. Generate descriptive statistics and visualizations for weather patterns.
3. Build a predictive model for temperature and rainfall.
4. Output predictions for specific regions based on user queries.
5. Analyze December's weather data for 16 cities in Maharashtra.
6. Develop a predictive model for temperature and rainfall.
7. Provide insights for agricultural planning and travel.

---

- **Conclusion**

This solution aims to provide reliable weather predictions for Maharashtra's 16 cities using December's data. The insights will support agricultural and travel planning. Future enhancements could include integration with live weather data and expansion to other regions or timeframes.