Project 9: Air Quality Analysis and Prediction in Tamil Nadu

Problem Statement:

- The objective of this project is to analyze and visualize air quality data from various monitoring stations in Tamil Nadu. The dataset contains measurements of Sulfur Dioxide (SO2), Nitrogen Dioxide (NO2), and Respirable Suspended Particulate Matter/Particulate Matter 10 (RSPM/PM10) levels in different cities, towns, villages, and areas. The project aims to gain insights into the air pollution trends, identify areas with high pollution levels, and create a predictive model to estimate RSPM/PM10 levels based on SO2 and NO2 levels.

Air pollution in Tamil Nadu, India:

- Air pollution is a big concern of humanity over the last few decades. The increase of mass production, and numbers of vehicles obviuosly not slowing down. Lets see whats going on.
- This notebook is a starter for visualizations of air pollution data in India state Tamil Nadu. I call it a starter, because there is still lots of work to do. Feel free to use this notebook as a starting point.
- # some dynamic folium features would work on the latest versions only, uncomment and run the line below in order to upgrade your
- # kernels folium package to the latest version:
- # !pip install folium --upgrade

Dataset link:

- https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014

Air pollution in Tamil Nadu, India. Starter.

Notebook Input Output Logs Comm

In [5]:

```
import numpy as np
import pandas as pd
import geopandas as gpd
import math

import folium
from folium import Choropleth, Circle,
Marker, plugins
from folium.plugins import HeatMap, Ma
rkerCluster, TimestampedGeoJson

from geopandas.tools import geocode
from geopy.geocoders import GoogleV3
geocoder = GoogleV3(api_key = 'your_go
ogle_maps_API_key') #I have uploaded th
e processed dataset with already geocod
ed properties.
# if you want to use this feature, crea
te an account at google maps, and get t
heir API key
```

Air pollution in Tamil Nadu, India. Starter.

Notebook Input Output Logs Comm

```
import datetime

import seaborn as sns
```

Reading data:

Geocoded dataset:

In [6]:

```
df = pd.read_csv('../input/tamil-nadu-
air-pollution-latlon/India_Air_LatLOn.
csv')
```

.shp file for districts of Tamil Nadu state.

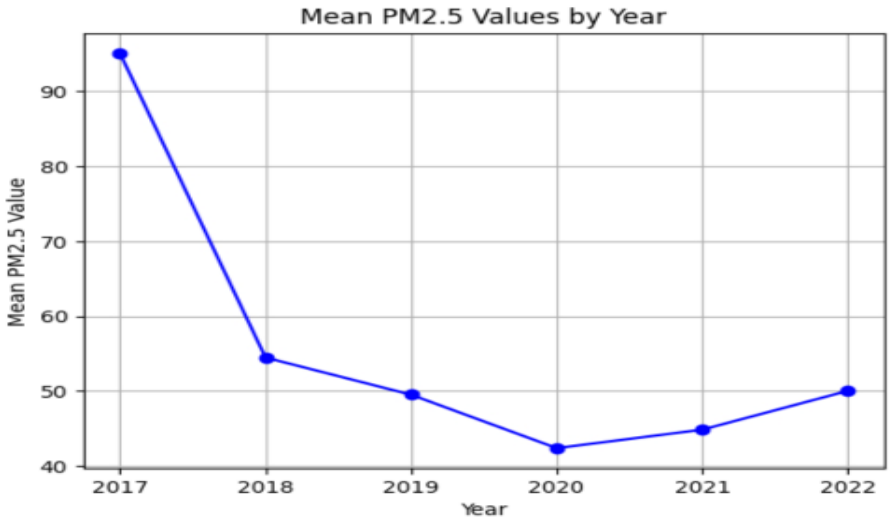
Project Steps

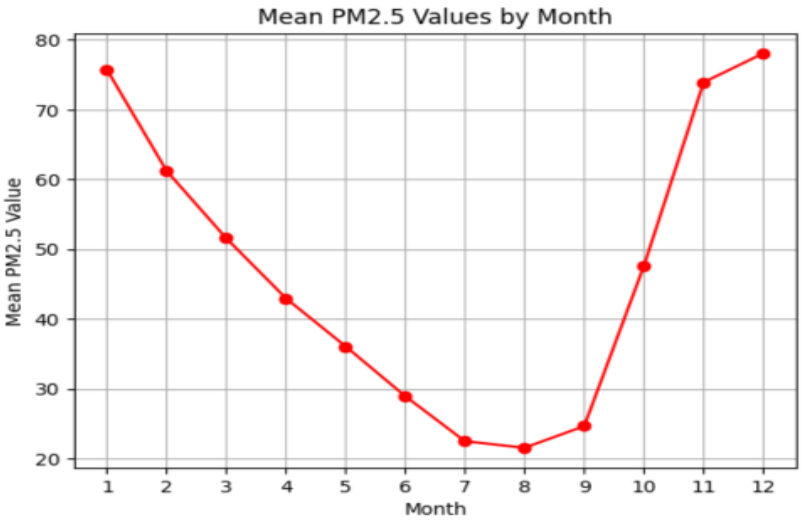
Phase 1: Problem Definition and Design Thinking

Problem Definition: The project aims to analyze and visualize air quality data from monitoring stations in Tamil Nadu. The objective is to gain insights into air pollution trends, identify areas with high pollution levels, and develop a predictive model to estimate RSPM/PM10 levels based on SO2 and NO2 levels. This project involves defining objectives, designing the analysis approach, selecting visualization techniques, and creating a predictive model using Python and relevant libraries.

Design Thinking:

Project Objectives: Define objectives such as analyzing air quality trends, identifying pollution hotspots, and building a predictive model for RSPM/PM10 levels.
Analysis Approach: Plan the steps to load, preprocess, analyze, and visualize the air quality data.
Visualization Selection: Determine visualization techniques (e.g., line charts, heatmaps) to effectively represent air quality trends and pollution levels.





Data in row 2

Stn Code
38

Sampling Date
01/02/2014

State
Tamil Nadu

City/Town/Village/Area
Chennai

Location of Monitoring Station
Kathivakkam, Municipal Kalyana Mandapam, Chennai

Agency
Tamilnadu State Pollution Control Board

Type of Location
Industrial Area

Data in row 3

Stn Code
38

Sampling Date
01/07/2014

State
Tamil Nadu

City/Town/Village/Area
Chennai

Location of Monitoring Station
Kathivakkam, Municipal Kalyana Mandapam, Chennai

Agency
Tamilnadu State Pollution Control Board

Type of Location
Industrial Area

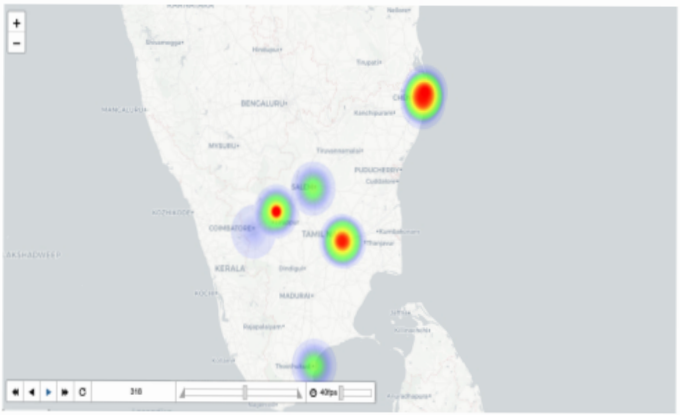
Air pollution in Tamil Nadu, India. Starter.

Notebook Input Output Logs Comm

Air pollution in Tamil Nadu, India. Starter.

Show hidden code

Out[1]:



Design Thinking:

Project Objectives: Define objectives such as analyzing air quality trends, identifying pollution hotspots, and building a predictive model for RSPM/PM10 levels.
Analysis Approach: Plan the steps to load, preprocess, analyze, and visualize the air quality data.
Visualization Selection: Determine visualization techniques (e.g., line charts, heatmaps) to effectively represent air quality trends and pollution levels.

Air pollution in Tamil Nadu, India. Starter.

Notebook Input Output Logs Comm

```
    },
    8: {'notation' : 'NO2',
        'name' :'Nitrogen dioxide',
        'bin_edges' : np.array([25,45,
60,80,110,150,200,270,400])
    },
    10: {'notation' : 'CO',
        'name' :'Carbon monoxide',
         'bin_edges' : np.array([1.4,
2.1,2.8,3.6,4.5,5.2,6.6,8.4,13.7])
    },
    20: {'notation' : 'C6H6',
        'name' :'Benzene',
         'bin_edges' : np.array([0.5,
1.0,1.25,1.5,2.75,3.5,5.0,7.5,10.0])
    }
}
```

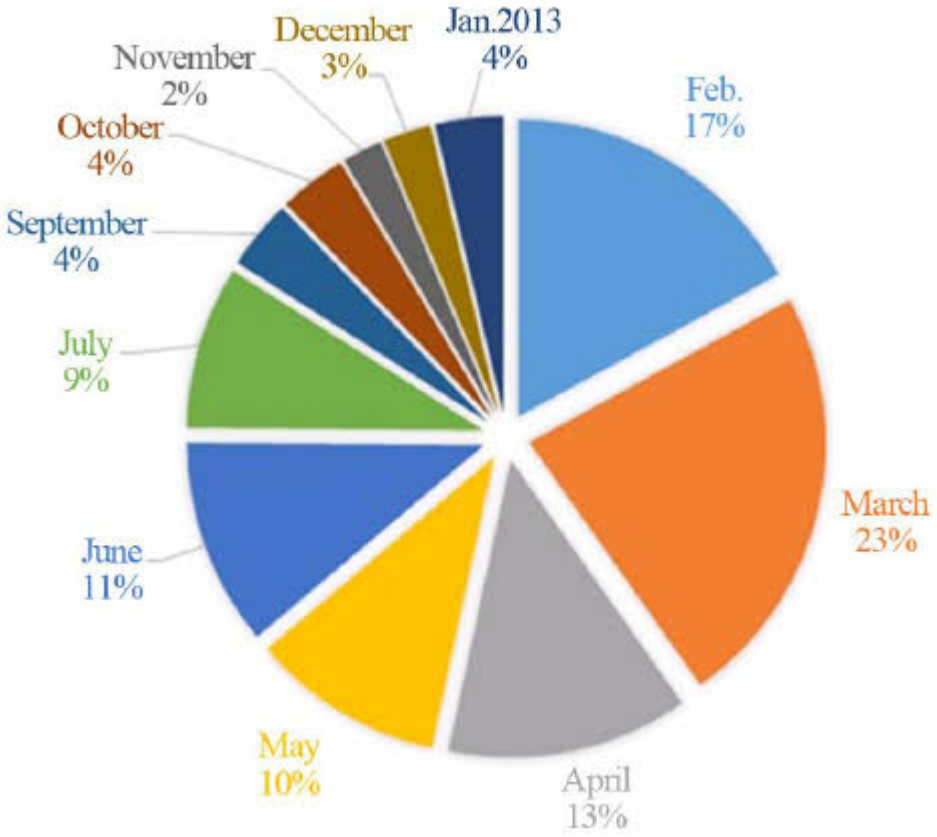
Air pollution in Tamil Nadu, India. Starter.

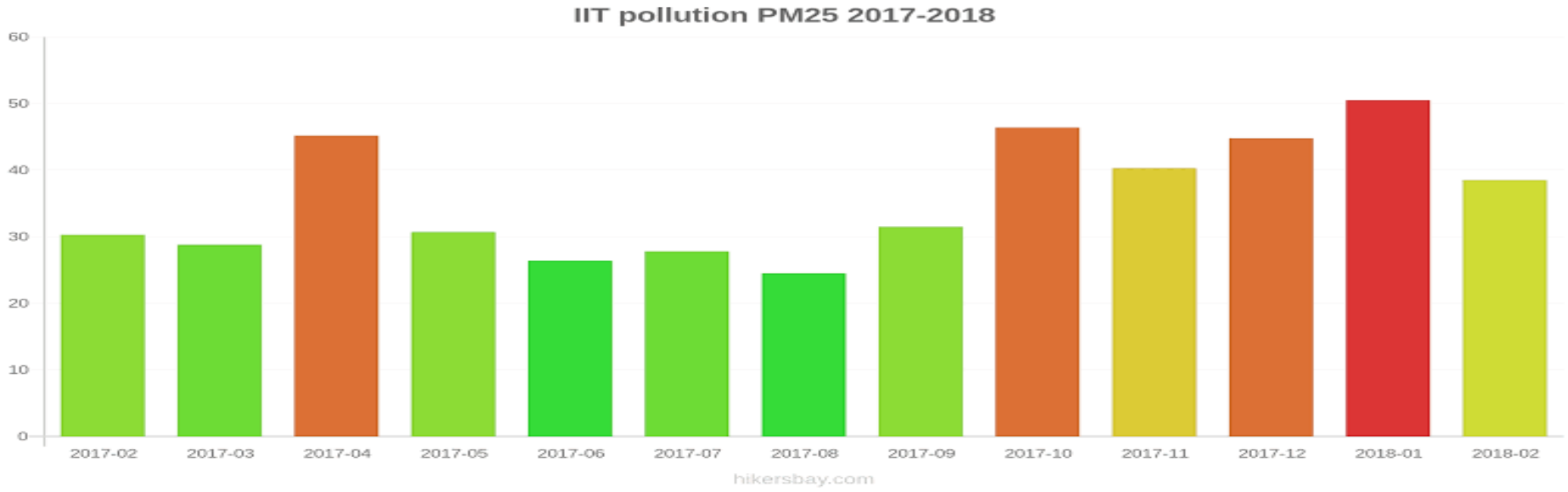
Notebook Input Output Logs Comm

Dictionary of air pollutants:

In [10]:

```
pollutants = {
    1: {
        'notation' : 'SO2',
        'name' :'Sulphur dioxide',
        'bin_edges' : np.array([15,30,
45,60,80,100,125,165,250])
    },
    5: {
        'notation' : 'PM10',
        'name' :'Particulate matter <
10 µm',
        'bin_edges' : np.array([10,20,
30,40,50,70,100,150,200])
    },
    7: {'notation' : 'O3',
        'name' :'Ozone',
        'bin_edges' : np.array([30,50,
70,90,110,145,180,240,360])
```


December
3%
Jan.2013
4%
November
2%
Feb.
17%
October
4%
September
4%
July
9%
March
23%
June
11%
May
10%
April
13%


IIT pollution PM25 2017-2018
60
50
40
30
20
10
0
2017-02
2017-03
2017-04
2017-05
2017-06
2017-07
2017-08
2017-09
2017-10
2017-11
2017-12
2018-01
2018-02
hikersbay.com

THANK YOU