DATA ANALTYICS WITH COGNOS

PROJECT: AIR QUALITY ANLYSIS TAMIL NADU

PHASE 5 PROJECT

TEAM MEMBERS:

1. S. BALAKANDAN – 820421205012 – balakandan50@gmail.com
2. R. BHARATH RAJ – 820421205015 – bharathrajravi2004@gmail.com
3. V. ARUN KUMAR – 820421205008 – arurarun0088@gmail.com
4. M. MUKILAN – 820421205042 – mukilmuruga07@gmail.com
5. T. LOGESH – 820421205039 – logeshselvan6@gmail.com

Table of Contents:

1. Introduction:

The project aims to analyze and visualize air quality in Tamil Nadu using IBMCognos.The objective is to create Reports, Dashboard based on the dataset. To analyze and gain insights into air pollution trends, identifying areas with high pollution level and to develop a predictive model to estimate RSPM/PM10 levels based onSO2 and NO2 levels. This project encompasses defining objective, designing the analysis approach, selecting a relevant visualization technique using IBM Cognosand creating a predictive model using Python and relevant libraries.

Chennai beats Delhi in pollution, records 'very poor' air quality

Velachery, Ramapuram, Manali, Kodungaiyur, Anna Nagar, Chennai Airport clocked pollution levels as high as 341, while Delhi stood at 254.

Published: 07th November 2019 11:51 AM | Last Updated: 07th November 2019 09:43 PM

A+ A A-

Many areas in the Tamil Nadu capital recorded 'very poor' air quality. | (Photo | Martin Louis/EPS)

2.Project Objectives:

- Analyze air quality trends**.**
- Identifying pollution hotspots using visualization technique.
- To build a predictive model for RSPM/PM10 levels based on SO2 and NO2 levels using Python and relevant libraries.

3.Analysis Approach:

1. Data Collection:

- The dataset containing location-wise daily ambient air quality records for Tamil Nadu in the year 2014 has been obtained from the below datalink.
- Dataset Link: https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014
- Describe the data sources and methods used to collect air quality data in Tamil Nadu.
- Mention any data preprocessing steps, data cleaning, and data format (e.g., CSV files, API calls).

Dataset:

	A	B	C	D	E	F	G	H	I	J	K
1	Stn Code	Sampling D	State	City/Town	Location o	Agency	Type of Lo	SO2	NO2	RSPM/PM:	PM 2.5
2	38	1/2/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	11	17	55	NA
3	38	1/7/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	13	17	45	NA
4	38	21-01-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	12	18	50	NA
5	38	23-01-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	16	46	NA
6	38	28-01-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	13	14	42	NA
7	38	30-01-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	18	43	NA
8	38	2/4/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	12	17	51	NA
9	38	2/6/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	13	16	46	NA
10	38	########	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	10	19	50	NA
11	38	13-02-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	14	48	NA
12	38	18-02-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	16	32	NA
13	38	20-02-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	14	29	NA
14	38	25-02-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	13	17	17	NA
15	38	27-02-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	16	44	NA
16	38	3/4/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	12	17	25	NA
17	38	3/6/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	13	16	29	NA
18	38	########	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	11	18	29	NA
19	38	13-03-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	16	41	NA
20	38	18-03-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	17	43	NA
21	38	20-03-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	14	42	NA
22	38	25-03-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	17	54	NA
23	38	27-03-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	19	62	NA
24	38	4/1/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	15	66	NA
25	38	4/3/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	11	16	40	NA
26	38	4/8/2014	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	14	17	56	NA
27	38	########	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	17	50	NA
28	38	15-04-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	12	14	49	NA
29	38	17-04-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	16	63	NA
30	38	22-04-14	Tamil Nad	Chennai	Kathivakka	Tamilnadu	Industrial A	15	18	42	NA

cpcb-dly-aq-tamil-nadu-2014 (2)

2. **Data Cleaning and Preprocessing:**

- Data cleaning is one of the important parts of machine learning. It plays a significant part in building a model.

- Data pre-processing is a crucial step in preparing the dataset for machine learning.

- Data Cleansing involves:
    - Identifying and removing any missing, duplicate or irrelevant data.
    - Handling missing data.
    - Removing outliers.
    - Scaling the data.

- Using Methods Such as:
    - dropna()
    - drop()
    - drop_duplicates()
    - scale()
    - get_dummies()

- Handling Missing Data:
    - Checked for missing values in the dataset and decided an appropriate strategy for handling them (e.g., imputation or removal).

- Data conversion:
    - The datatypes of every columns are changed or converted for the ease of analysis.

3. Exploratory Data Analysis:

- Descriptive Analysis:
    - The mean / average of the pollutants like SO2,NO2 and RSPM/PM10 are calculated to understand the central tendency and variability of pollutant concentrations.
- Time Series Analysis:
    - Explore temporal patterns and trends in air quality data. Identify seasonal variations and potential outliers.

4. Feature Engineering:

- Feature engineering involves creating new features or modifying existing ones to improve model performance. In this context, it may involve generating lag features, aggregating data over time intervals, or incorporating weather data ifavailable to capture external factors influencing air quality.

6.Air Quality Index:

- Air quality is a measure of how clean or polluted the air is. Monitoring air quality is important because polluted air can be bad for our health—and the health of the environment. Air quality is measured with the Air Quality Index, or AQI. The AQI works like a thermometer that runs from 0 to 500 degrees. However, instead of showing changes in the temperature, the AQI is a way of showing changes in the amount of pollution in the air.

Air Quality Index Levels of Health Concern	Numerical Value	Meaning
Good	0 to 50	Air quality is considered satisfactory, and air pollution poses little or no risk.
Moderate	51 to 100	Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution.
Unhealthy for Sensitive Groups	101 to 150	Members of sensitive groups may experience health effects. The general public is not likely to be affected.
Unhealthy	151 to 200	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects.
Very Unhealthy	201 to 300	Health alert: everyone may experience more serious health effects.
Hazardous	301 to 500	Health warnings of emergency conditions. The entire population is more likely to be affected.

4.Visualization Techniques:

- To effectively communicate the findings, various visualization techniques were employed:

    - Line Graph:The line graph is used to visualize the density of pollutants in the areas in a specific way.The pollutants are SO2 ,NO2,RSPM/PM10 which is given in the provided data set.

- Pie chart: The pie chart shows the levels of the pollutants in the cities/towns and so on .The main purpose of the pie chart is to show that which city is highly polluted.

- Time Series Plot: The time series plot shows the variance in the pollutant levels according to the sampling dates in a particular monitoring station.

- Heat map: The heat map gives the remarkable insight about the overall pollution made by different types of pollutants.

- Scatter plot: The scatter plot provide a different approach on the basis of the pollution in a city or a town.

5.Machine learning model:

- To Select appropriate machine learning algorithms for the task. Potential modelsinclude:
  - Regression models (e.g., Linear Regression, Random Forest Regression).
  - Time series forecasting models (e.g., ARIMA, LSTM).
  - Ensemble methods for improved accuracy.
- In the above given models Linear Regression is selected
- As Linear Regression is easy to implement and gives effective insights about any given dataset.

- Split the dataset into training and testing sets. Train the selected models on the training data and evaluate their performance using suitable metrics (e.g., Mean Absolute Error, R-squared). Perform hyper parameter tuning to optimize model performance

6.Program:

- First, import the required Python libraries. As given below.

```
# In[1]: import numpy as np
         import pandas as pd
         from sklearn.linear_model import LinearRegression
         from sklearn.model_selection import train_test_split
         from sklearn.metrics import mean_squared_error,r2_score
         from sklearn import datasets
         import pandas as pd
         import plotly.express as px
         import plotly.io as pio
         import plotly.graph_objects as go
```

- To import the dataset into the jupyter notebook, you can use the following Python code given below.

```
# In[2]:data =pd.read_csv(r"C:\Users\Administrator\Downloads\airquality.csv")


# In[3]:print(data.head())
```

 # OP [3]:

```
   Stn Code Sampling Date       State City/Town/Village/Area  \
0        38      01-02-14  Tamil Nadu                 Chennai

   1        38      01-07-14  Tamil Nadu                 Chennai
   2        38      21-01-14  Tamil Nadu                 Chennai
   3        38      23-01-14  Tamil Nadu                 Chennai
   4        38      28-01-14  Tamil Nadu                 Chennai

Location of Monitoring Station  \
  0  Kathivakkam, Municipal Kalyana Mandapam, Chennai
  1  Kathivakkam, Municipal Kalyana Mandapam, Chennai
  2  Kathivakkam, Municipal Kalyana Mandapam, Chennai
  3  Kathivakkam, Municipal Kalyana Mandapam, Chennai
  4  Kathivakkam, Municipal Kalyana Mandapam, Chennai

                                    Agency Type of Location   SO2   NO2  \
  0  Tamilnadu State Pollution Control Board  Industrial Area  11.0  17.0
  1  Tamilnadu State Pollution Control Board  Industrial Area  13.0  17.0
  2  Tamilnadu State Pollution Control Board  Industrial Area  12.0  18.0
  3  Tamilnadu State Pollution Control Board  Industrial Area  15.0  16.0
  4  Tamilnadu State Pollution Control Board  Industrial Area  13.0  14.0
  5

   RSPM/PM10  PM 2.5
  0      55.0     NaN
  1      45.0     NaN
  2      50.0     NaN
  3      46.0     NaN
  4      42.0     NaN
```

- Data Cleansing and Transformation are done by following Python Code.

```
# In[3]: df=data.drop(['PM 2.5'],axis=1)
# In[4]: df.head()

# OP[4]:
```

	Stn Code	Sampling Date	State	City/Town/Village/Area	Location of Monitoring Station	Agency	Type of Location	SO2	NO2	RSPM/PM10
0	38	01-02-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	11.0	17.0	55.0
1	38	01-07-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	17.0	45.0
2	38	21-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.0	18.0	50.0
3	38	23-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapa	Tamilnadu State Pollution	Industrial Area	15.0	16.0	46.0

					m, Chennai	Control Board				
4	38	28-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	14.0	42.0

```
# In[5]: newf=df.dropna()
         newd=newf.drop_duplicates()

# In[6]: newf.head()
        newf.scale()
        newf.get_dummies()

# OP[6]:
```

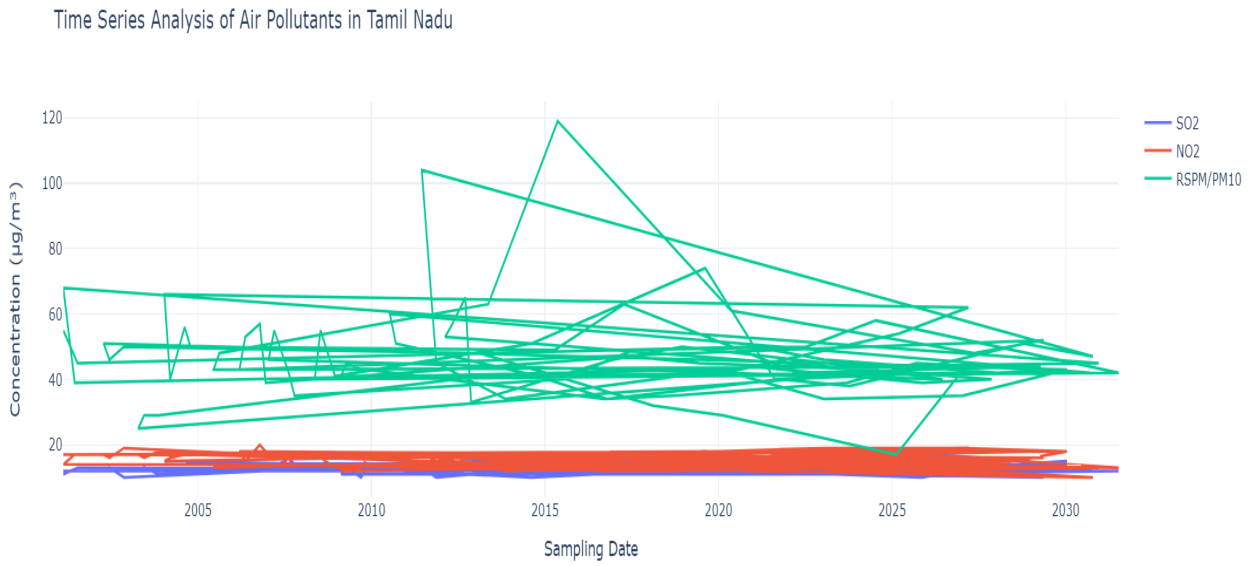
Stn Code	Sampling Date	State	City/Town/Village/Area	Location of Monitoring Station	Agency	Type of Location	SO2	NO2	RSPM/PM10	
0	38	01-02-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	11.0	17.0	55.0
1	38	01-07-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam,	Tamilnadu State Pollution Contro	Industrial Area	13.0	17.0	45.0

					Chennai	l Board				
2	38	21-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.0	18.0	50.0
3	38	23-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.0	16.0	46.0
4	38	28-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	14.0	42.0

- Constructing a Time series plot for each Air Pollution:

```
# time series plot for each air pollutant
fig = go.Figure()

for pollutant in ['NO2', 'SO2', 'RSPM/PM10']:
    fig.add_trace(go.Scatter(x=data['date'], y=data[pollutant], mode='lines',
                             name=pollutant))

fig.update_layout(title='Time Series Analysis of Air Pollutants in Delhi',
                  xaxis_title='Date', yaxis_title='Concentration (µg/m³)')
fig.show()
```

- Time Series Analysis of Air Pollutants in Tamil Nadu



- Descriptive Analysis on the levels of the pollutants , the average of the pollutants in the monitoring station Kathivakkam

```
sta1['SO2'].mean()    #Average of SO2 in Kathivakkam
```

```
12.925531914893616
```

```
sta1['NO2'].mean()    #Average of NO2 in Kathivakkam
```

```
15.170212765957446
```

```
sta1['RSPM/PM10'].mean()     #Average of RSPM/PM10 in Kathivakkam
```

```
46.851063829787236
```

- The average of the pollutants in the monitoring station Govt.High School,Manali

```
sta2['SO2'].mean()    #Average of SO2 in Govt. High School,Manali
```

```
13.043010752688172
```

```
sta2['NO2'].mean()    #Average of SO2 in Govt. High School,Manali
```

```
15.408602150537634
```

```
sta2['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Govt. High School,Manali
```

```
44.61290322580645
```

- The average of the pollutants in the monitoring station Thiruvotriyur

```
sta3['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Thiruvottiyur
```

```
42.604166666666664
```

```
sta3['NO2'].mean()    #Average of SO2 in Thiruvottiyur
```

```
15.583333333333334
```

```
sta3['SO2'].mean()    #Average of SO2 in Thiruvottiyur
```

```
13.010416666666666
```

- The average of the pollutants in the monitoring station Thiyagaraya Nagar

```
sta4['SO2'].mean()    #Average of SO2 in Thiyagaraya Nagar, Chennai
```

```
18.849557522123895
```

```
sta4['NO2'].mean()    #Average of NO2 in Thiyagaraya Nagar, Chennai
```

```
28.25
```

```
sta4['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Thiyagaraya Nagar, Chennai
```

```
102.32743362831859
```

- The average of the pollutants in the monitoring station Chennai

```
city1['SO2'].mean()    #Average of SO2 in Chennai
```

```
13.014042126379138
```

```
city1['NO2'].mean()    #Average of NO2 in Chennai
```

```
22.088442211055277
```

```
city1['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Chennai
```

```
58.998
```

- The average of the pollutants in the monitoring station Chennai

```
city2['SO2'].mean()    #Average of SO2 in Coimbatore
```

```
4.541095890410959
```

```
city2['NO2'].mean()    #Average of NO2 in Coimbatore
```

```
25.325342465753426
```

```
city2['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Coimbatore
```

```
49.217241379310344
```

- The average of the pollutants in the monitoring station Cuddalore

```
city3['SO2'].mean()    #Average of SO2 in Cuddalore
```

```
8.965986394557824
```

```
city3['NO2'].mean()    #Average of NO2 in Cuddalore
```

```
19.710884353741495
```

```
city3['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Cuddalore
```

```
61.88175675675676
```

- The average of the pollutants in the monitoring station Madurai
-

```
city4['SO2'].mean()    #Average of SO2 in Madurai
```

```
3.319727891156463
```

```
city4['NO2'].mean()    #Average of NO2 in Madurai
```

```
5.768707482993197
```

```
city4['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Madurai
```

- The average of the pollutants in the monitoring station Mettur

```
city5['SO2'].mean()    #Average of SO2 in Mettur
```

```
8.429268292682927
```

```
city5['NO2'].mean()    #Average of NO2 in Mettur
```

```
23.185365853658535
```

```
city5['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Mettur
```

```
52.72195121951219
```

- The average of the pollutants in the monitoring station Salem

```
city6['SO2'].mean()    #Average of SO2 in Salem
```

```
8.114503816793894
```

```
city6['NO2'].mean()    #Average of NO2 in Salem
```

```
28.66412213740458
```

```
city6['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Salem
```

```
62.954198473282446
```

- The average of the pollutants in the monitoring station Thoothukudi

```
city7['SO2'].mean()    #Average of SO2 in Thoothukudi
```

```
12.989690721649485
```

```
city7['NO2'].mean()    #Average of NO2 in Thoothukudi
```

```
18.512027491408936
```

```
city7['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Thoothukudi
```

```
83.45890410958904
```

- The average of the pollutants in the monitoring station Trichy

```
city8['SO2'].mean()    #Average of SO2 in Trichy
```

```
15.293956043956044
```

```
city8['NO2'].mean()    #Average of NO2 in Trichy
```

```
18.695054945054945
```

```
city8['RSPM/PM10'].mean()    #Average of RSPM/PM10 in Trichy
```

```
85.05449591280654
```

- Linear regression for the predicting the levels of the pollutants and analysis.

```
# In[7]: x = newf[['NO2','SO2']]


# In[8]: y = newf['RSPM/PM10']


# In[9]:x_train,x_test,y_train,y_test =
         train_test_split(x,y,test_size=0.3,random_state=0)


# In[8]: model = LinearRegression()
```

- Evaluate the Model using Mean Squared Error and R2 score

```
# In[11]: mse = mean_squared_error(y_test,y_pred)


# In[12]: r2 = r2_score(y_test,y_pred)


# In[13]: print("Mean Squared Error:",mse)
           print("R-squared:",r2)

# OP[13]: Mean Squared Error: 908.4528649741137
          R-squared: 0.19877081345863346
```

- Calculating Air Quality Index using RSPM/PM10 and Categorizing Air QualityIndex :

```
# Define AQI breakpoints and corresponding AQI values
aqi_breakpoints = [
    (0, 12.0, 50), (12.1, 35.4, 100), (35.5, 55.4, 150),
    (55.5, 150.4, 200), (150.5, 250.4, 300), (250.5, 350.4, 400),
    (350.5, 500.4, 500)
]

def calculate_aqi(pollutant_name, concentration):
    for low, high, aqi in aqi_breakpoints:
        if low <= concentration <= high:
            return aqi
    return None

def calculate_overall_aqi(row):
    aqi_values = []
    pollutants = ['NO2','SO2','RSPM/PM10']
    for pollutant in pollutants:
        aqi = calculate_aqi(pollutant, row[pollutant])
        if aqi is not None:
            aqi_values.append(aqi)
    return max(aqi_values)

# Calculate AQI for each row
newd['AQI'] = newd.apply(calculate_overall_aqi, axis=1)

# Define AQI categories
aqi_categories = [
    (0, 50, 'Good'), (51, 100, 'Moderate'), (101, 150, 'Unhealthy for Sensitive Groups'),
    (151, 200, 'Unhealthy'), (201, 300, 'Very Unhealthy'), (301, 500, 'Hazardous')
]

def categorize_aqi(aqi_value):
    for low, high, category in aqi_categories:
        if low <= aqi_value <= high:
            return category
    return None
```

- Categories the Air Quality Index Values as ‘Good’,’Moderate’,’Unhealthy for Sensitive Groups’ and ‘Unhealthy’ .

```
# Categorize AQI
newd['AQI Category'] = newd['AQI'].apply(categorize_aqi)
print(newd.tail())
```

AQI Categories:

	Stn Code	Sampling Date	State	City/Town/Village/Area	Location of Monitoring Station	Agency	Type of Location	SO2	NO2	RSPM/PM10	AQI	AQI_Category
0	38	01-02-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	11.0	17.0	55.0	150	Unhealthy for Sensitive Groups
1	38	01-07-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	17.0	45.0	150	Unhealthy for Sensitive Groups
2	38	21-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.0	18.0	50.0	150	Unhealthy for Sensitive Groups
3	38	23-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.0	16.0	46.0	150	Unhealthy for Sensitive Groups
4	38	28-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	14.0	42.0	150	Unhealthy for Sensitive Groups

- Identify pollution trends and areas with high pollution levels by highlight function:

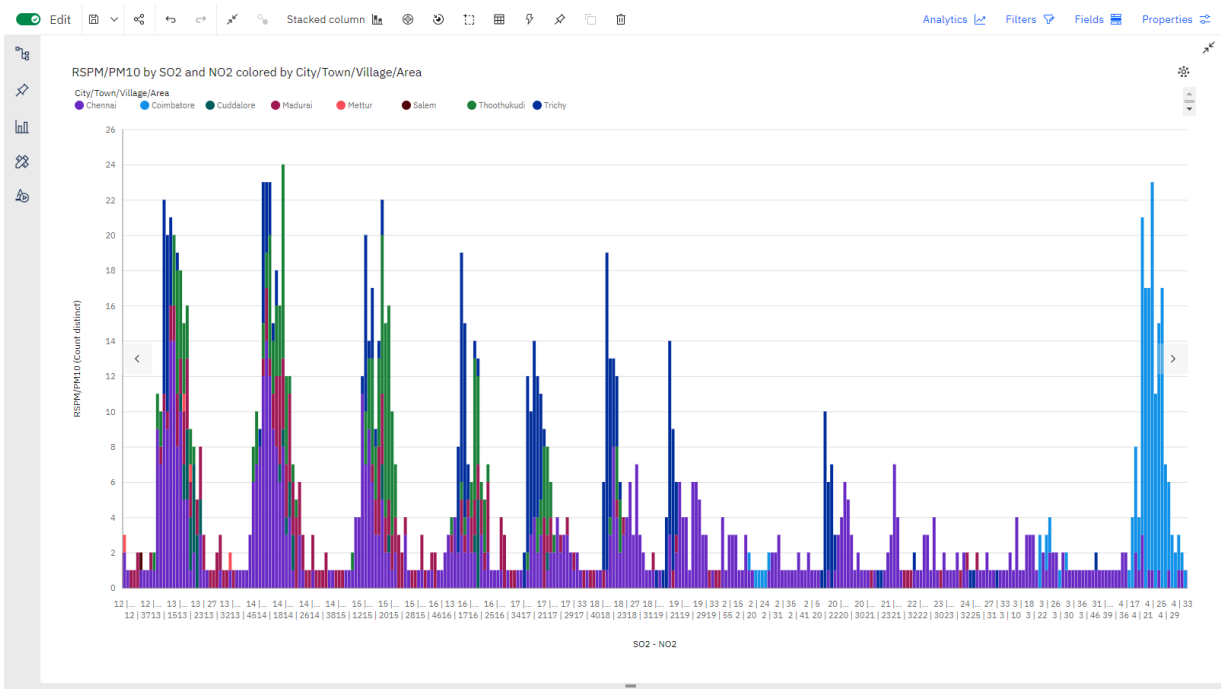
```
def highlight_unhealthy(val):
    color = 'red' if val == 'Unhealthy' else ''
    return f'background-color: {color}'

# Apply the highlight function to the 'AQI Category' column
styled_df1= newd.style.apply({'AQI Category': highlight_unhealthy}, subset=pd.IndexSlice[newd[newd['AQI Category'] == 'Unhealthy'].index, 'AQI Category'])

# Display the styled DataFrame
styled_df1
```

Areas with High Pollution Levels:

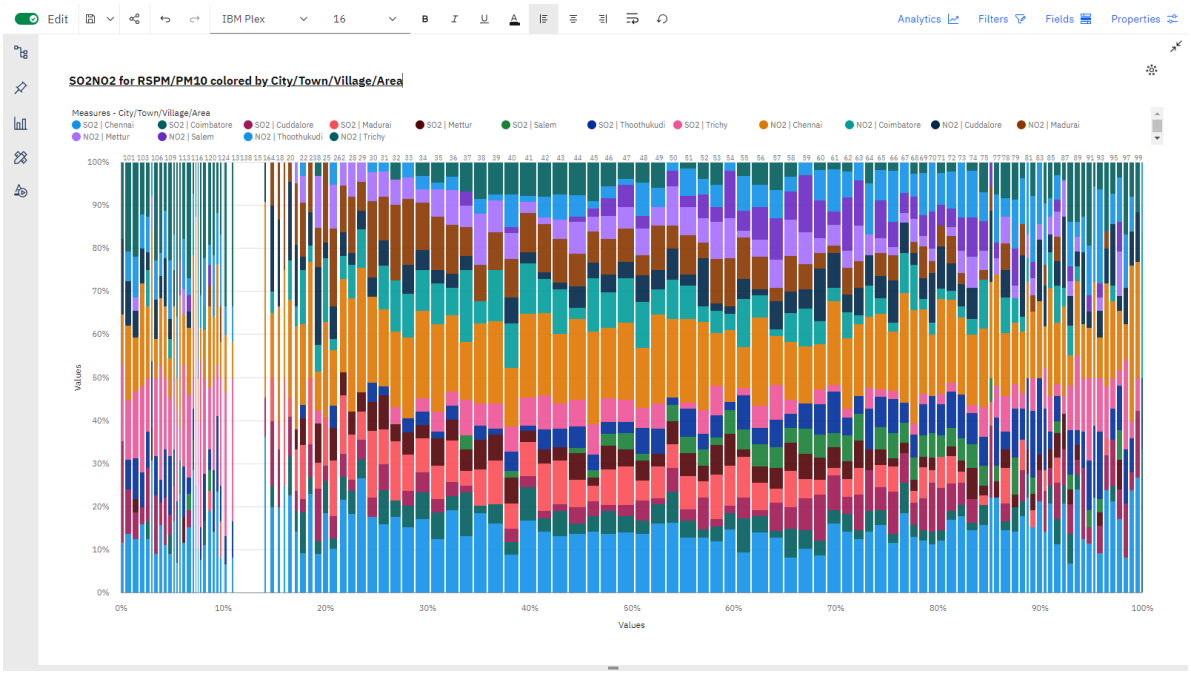
20	38	25-03-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	14.000000	17.000000	54.000000	150	Unhealthy for Sensitive Groups
21	38	27-03-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.000000	19.000000	62.000000	200	Unhealthy
22	38	04-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	14.000000	15.000000	66.000000	200	Unhealthy
23	38	04-03-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	11.000000	16.000000	40.000000	150	Unhealthy for Sensitive Groups
24	38	04-08-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	14.000000	17.000000	56.000000	200	Unhealthy
25	38	04-10-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.000000	17.000000	50.000000	150	Unhealthy for Sensitive Groups
26	38	15-04-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.000000	14.000000	49.000000	150	Unhealthy for Sensitive Groups
27	38	17-04-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.000000	16.000000	63.000000	200	Unhealthy
28	38	22-04-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.000000	18.000000	42.000000	150	Unhealthy for Sensitive Groups
29	38	29-04-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.000000	18.000000	44.000000	150	Unhealthy for Sensitive Groups
30	38	05-06-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.000000	13.000000	43.000000	150	Unhealthy for Sensitive Groups
31	38	05-08-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	14.000000	14.000000	48.000000	150	Unhealthy for Sensitive Groups
32	38	13-05-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	14.000000	13.000000	63.000000	200	Unhealthy
33	38	15-05-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.000000	15.000000	119.000000	200	Unhealthy
34	38	20-05-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.000000	18.000000	61.000000	200	Unhealthy

- Data Visualization done by IBM Cognos:

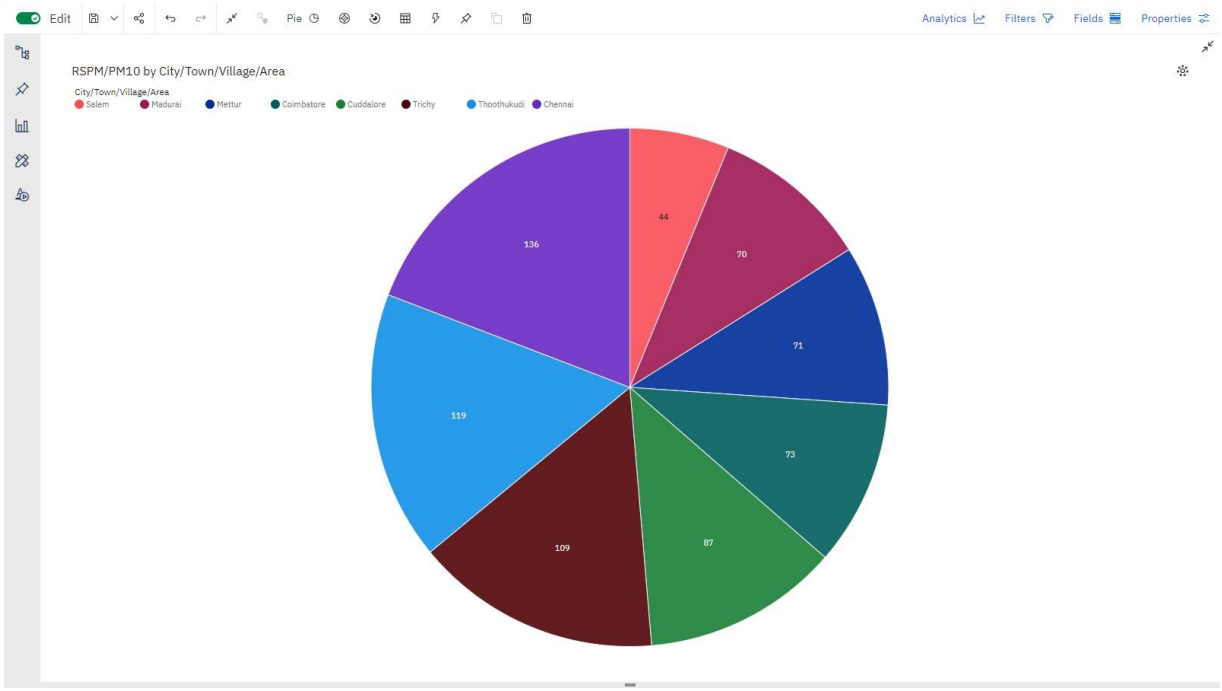
The Below Data Visualization is to visualize the NO2 and SO2 using IBM Cognos.



- The Visualization of SO2 and NO2 for RSPM/PM10 values for City/Town/Village/Area



- The Pie chart for RSPM/PM10 values for City/Town/Village/Area



- The Scatter plot for the pollutant levels in City/Town/Village/Area.

9.Conclusion:

The proposed approach aims to enhance the accuracy of predictive models for ambient air quality in Tamil Nadu through the incorporation of machine learning algorithms. The successof this project will lead to better air quality predictions, enabling more effective pollution control measures and safeguarding public health and the environment.