AIR QUALITY ASSESSMENT TAMIL NADU

INTRODUCTION:

Data from Tamil Nadu monitoring stations will be analyzed and displayed as part of the effort. Knowing the extent of air pollution and creating a prediction model based on NO2 and SO2 levels to calculate RSPM/PM10 levels. This project's goals are defined, an analytical strategy is created, a design is selected, visualization techniques are applied, and a prediction model is built.using Python along with the required libraries.

DESCRIPTION:

This stage outlines the project's objective, its development phases, and its design thinking approach. And the analysis's objectives, the data collection process, the data visualization using IBM Cognos, and the Python integration. explains how the findings from the study might be used to improve user experience on websites. Describe the

objectives of the study, the data collection process, the usage of IBM Cognos for data visualization, and the incorporation of Python. explains how website owners may utilize the analysis's findings to improve user experience.

OBJECTIVES:

Define goals like examining trends in air quality, locating areas of high pollution, and developing a model to anticipate RSPM/PM10 levels.

➢ RSPM : RSPM stands for "Respirable Suspended Particulate Matter." It is a term used in the context of air quality and environmental monitoring to refer to fine particles or particulate matter that are small enough to be inhaled into the respiratory system, specifically the lungs. These particles can include dust, smoke, pollen, and various pollutants.

➢ PM10 : PM10 stands for "Particulate Matter 10." It is a term used in the field of air quality and environmental monitoring to refer to airborne particles or particulate matter that have a diameter of

10 micrometers (μm) or smaller. PM10 particles are often referred to as fine particles.

PROJECT OBJECTIVES :

➢ Analyze air quality trends : Analyzing historical air quality data to identify patterns and trends.

➢ Identify pollution hotspots : Locating geographical areas with consistently poor air quality.

➢ Build a Predictive Model for RSPM/PM10 Levels: Developing a model to predict RSPM (Respirable Suspended Particulate Matter) and PM10 (Particulate Matter 10 micrometers or smaller) levels in the region.

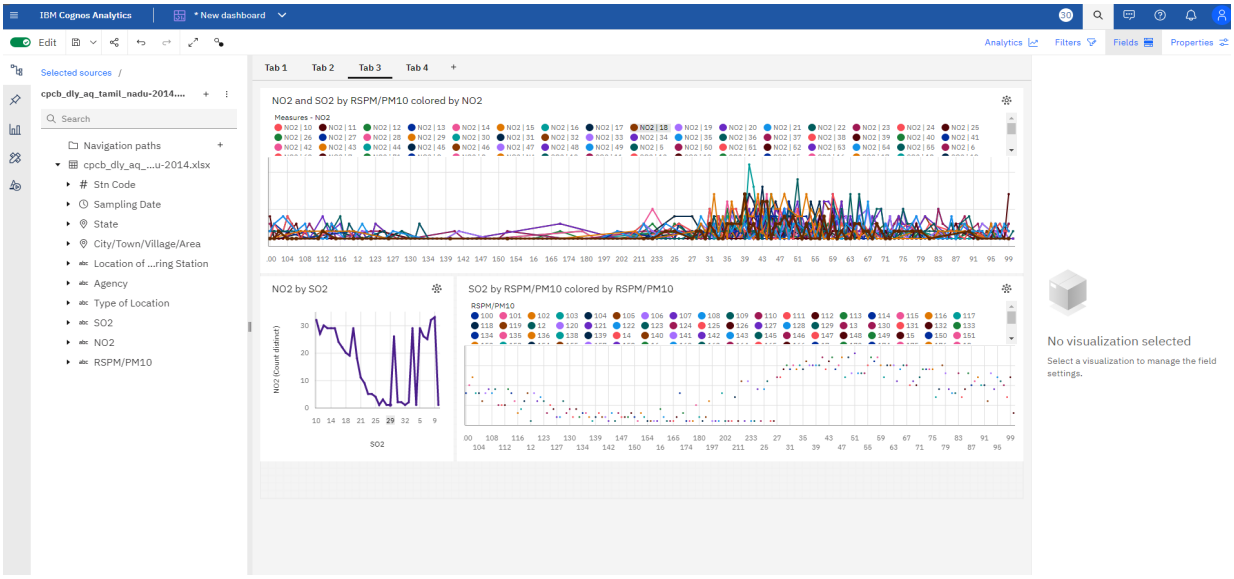
IBM COGNOS:

IBM Cognos is a collection of business intelligence and performance management tools intended to assist companies in deriving insights from their data. It provides a number of tools for scorecarding, reporting, analysis, and event and

metric monitoring. Ad hoc reporting, complex analytics, and interactive dashboards are just a few of the capabilities that IBM Cognos offers to help organizations make better decisions and operate more efficiently. It's widely recognized for its user-friendly interface and strong ability to convert unprocessed data into information that can be applied to strategic planning and decision-making.

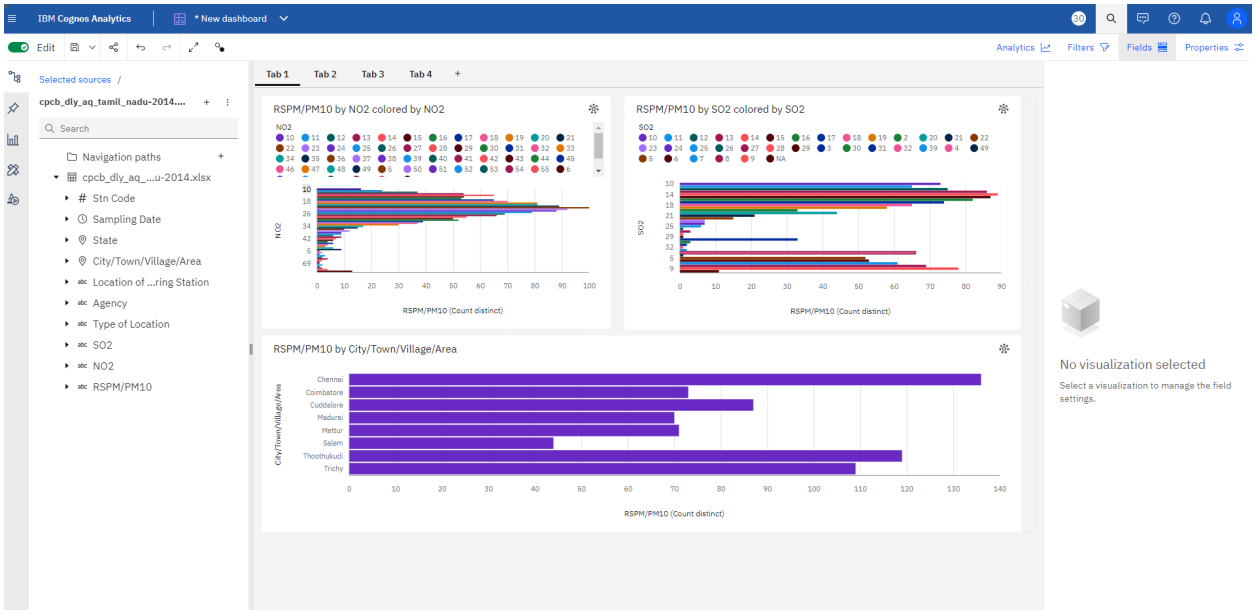
- **LINE CHART:**

  - RSPM/PM10 41 has the highest NO2 due to Stn Code 161.
  - Stn Code 767 has the highest NO2 at 80, out of which SO2 18 contributed the most at 12.
  - Stn Code 767 has the highest SO2 at 108, out of which RSPM/PM10 95 contributed the most at 4.
  - RSPM/PM10 41 SO2 from Stn Code 71 is 4, whereas 44 is only 1.

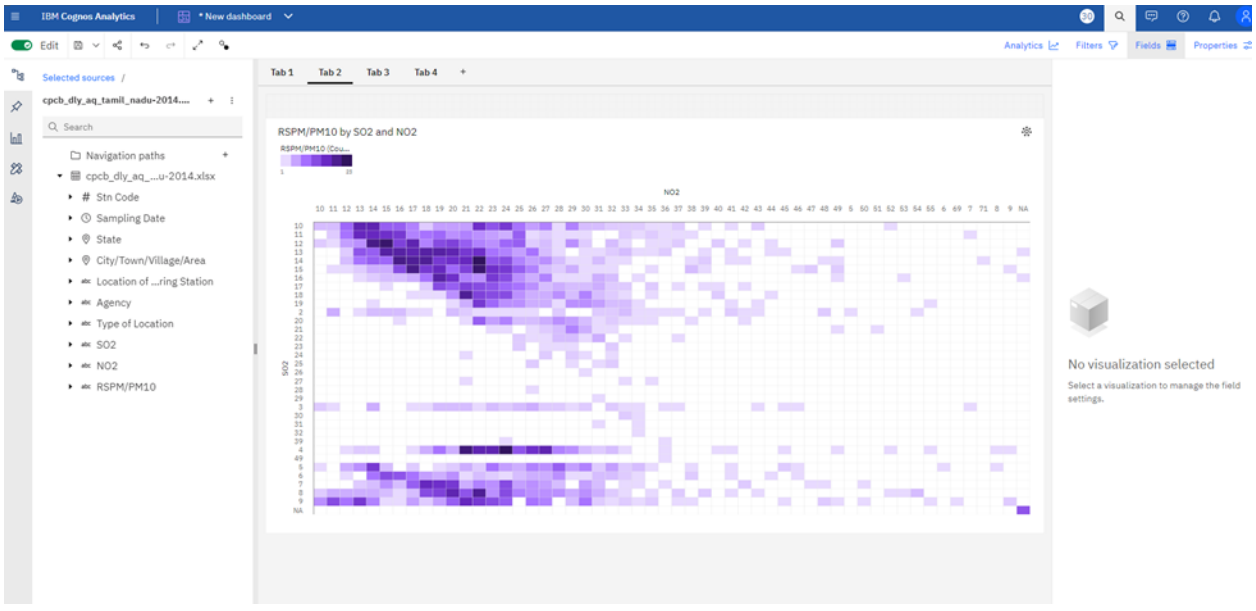


- **BAR CHART:**

  - 13 has a RSPM/PM10 of 20 for Stn Code 366.
  - NO2 22 RSPM/PM10 from Stn Code 239 is 17, whereas 24 is only 7.
  - 4 has a RSPM/PM10 of 45 for Stn Code 375.
  - City/Town/Village/Area Chennai has the highest RSPM/PM10 due to Stn Code 766.

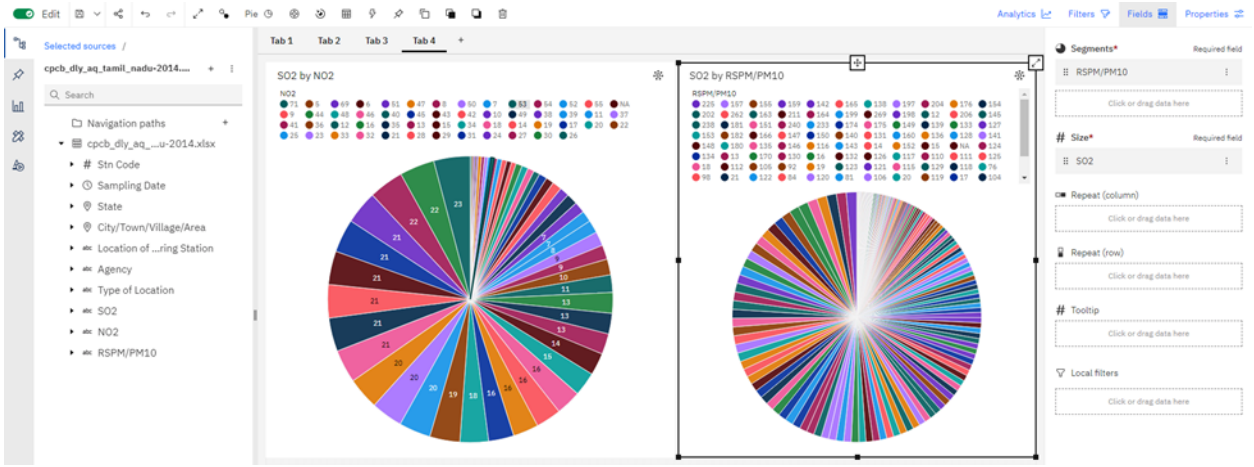

IBM Cognos Analytics
* New dashboard
Edit
Analytics
Filters
Fields
Properties
Selected sources
cpcb_dly_aq_tamil_nadu-2014....
Search
Navigation paths
cpcb_dly_aq_...u-2014.xlsx
Stn Code
Sampling Date
State
City/Town/Village/Area
Location of ...ring Station
Agency
Type of Location
SO2
NO2
RSPM/PM10
Tab 1
Tab 2
Tab 3
Tab 4
RSPM/PM10 by NO2 colored by NO2
RSPM/PM10 by SO2 colored by SO2
RSPM/PM10 by City/Town/Village/Area
RSPM/PM10 (Count distinct)
No visualization selected
Select a visualization to manage the field settings.

- **HEATMAP:**


IBM Cognos Analytics
* New dashboard
Edit
Analytics
Filters
Fields
Properties
Selected sources
cpcb_dly_aq_tamil_nadu-2014....
Search
Navigation paths
cpcb_dly_aq_...u-2014.xlsx
Stn Code
Sampling Date
State
City/Town/Village/Area
Location of ...ring Station
Agency
Type of Location
SO2
NO2
RSPM/PM10
Tab 1
Tab 2
Tab 3
Tab 4
RSPM/PM10 by SO2 and NO2
No visualization selected
Select a visualization to manage the field settings.

- 13 has a RSPM/PM10 of 20 for Stn Code 366.
- 4 has a RSPM/PM10 of 45 for Stn Code 375
- SO2 14 has the highest RSPM/PM10 at 215, out of which NO2 22 contributed the most at 23.

- **PIE CHART:**



- 29 has a SO2 of 6 for Stn Code 161.

- ❖ Stn Code 767 has the highest SO2 at 108, out of which RSPM/PM10 95 contributed the most at 4.
- ❖ RSPM/PM10 41 SO2 from Stn Code 71 is 4, whereas 44 is only 1.
- ❖ 47 (2.2 %) and 41 (2.2 %) are the most frequently occurring categories of RSPM/PM10 with a combined count of 126 items with SO2 values (4.4 % of the total).

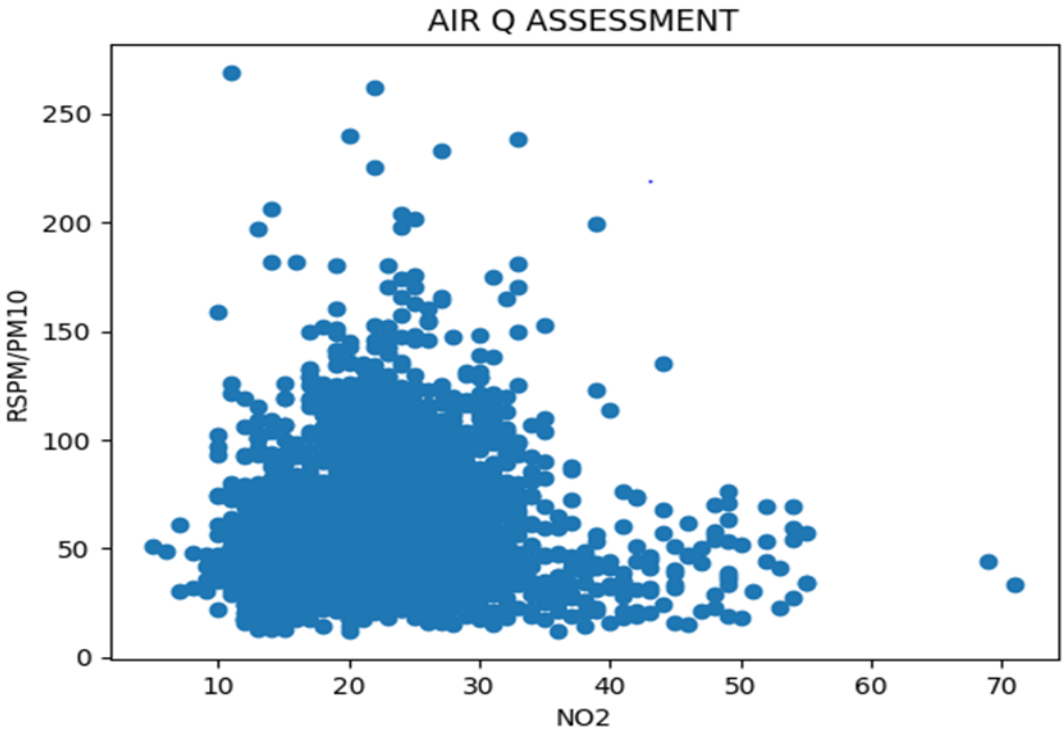
PYTHON:

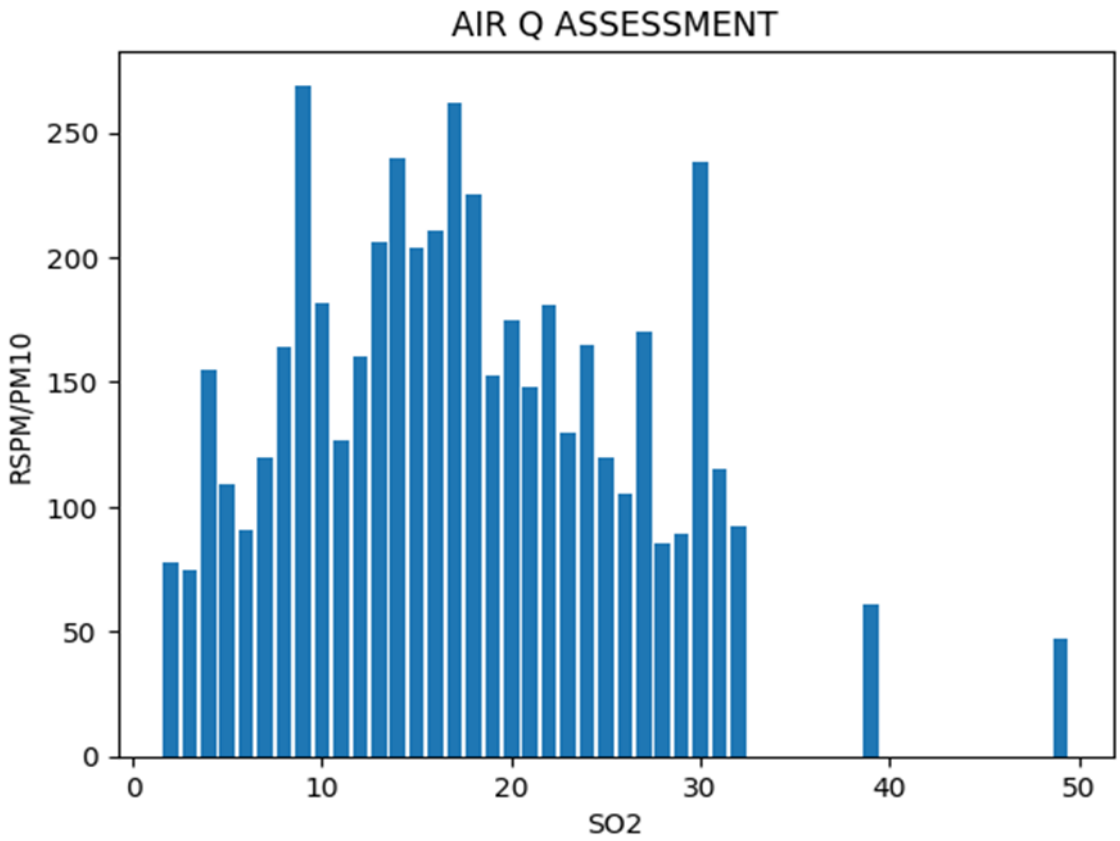
Simple, adaptable, and ideal for beginners. Its syntax encourages readability and simplifies the code. It drives automation, data science, and web development with its extensive standard library. Because it is collaborative and community-driven, Python is a popular choice among developers globally.

VISUALIZATION USING PYTHON:

Python visualization is the process of transforming data into graphical representations that

users may utilize to see trends, patterns, and insights. There are several libraries available for data visualization in Python; two well-liked options are Matplotlib and Seaborn. A variety of charts, including line, bar, scatter, and histogram plots, can be made with the help of these libraries.





PERFORMANCE METRICS:

Accuracy is a statistic indicating the accuracy of measurements or predictions, calculated by dividing the total number of measurements by the number of accurate predictions, with low errors indicating high accuracy. Tamil Nadu assesses air quality using important performance criteria, such as CO, NO2, SO2, PM2.5, and ozone. By classifying contaminants, the Air Quality Index enables government agencies to take appropriate measures to safeguard public health and lower pollution levels.

Metric	SO2	NO2	RSPM/PM10
MAE	13.67	16.75	47.71
MSE	240.87	346.42	2972.43
RMSE	15.52	18.60	54.49
R-squared	0.34	0.14	-1.02

Let's delve into the metrics and what they convey about the performance of a regression model using the provided data,

Mean Absolute Error (MAE):

- SO2: The average absolute difference between the predicted and actual SO2 values is approximately 13.67. This represents the typical magnitude of error in the model's predictions for SO2.
- NO2: For NO2, the MAE is around 16.75, indicating a slightly larger average absolute error compared to SO2.
- RSPM/PM10: The model's predictions for RSPM/PM10 have an average absolute error of about 47.71.

Mean Squared Error (MSE):

- SO2: The MSE for SO2 is 240.87, reflecting the average squared difference between predicted and actual SO2 values. A lower MSE indicates better model performance.
- NO2: MSE is 346.42 for NO2, suggesting a larger spread of squared errors compared to SO2.
- RSPM/PM10: The model has a higher MSE of 2972.43 for RSPM/PM10, signifying more variability in the squared errors.

Root Mean Squared Error (RMSE):

- SO2: The RMSE of 15.52 is the square root of the MSE for SO2, providing an interpretable scale similar to the original target variable.

- NO2: RMSE is 18.60 for NO2, indicating the average magnitude of error in the model's predictions for NO2.
- RSPM/PM10: The RMSE of 54.49 for RSPM/PM10 is relatively high, suggesting larger errors in predicting this variable.

R-squared:

- SO2: The R-squared value of 0.34 indicates that the model explains about 34% of the variability in SO2 values. A higher R-squared is desirable.
- NO2: With an R-squared of 0.14, the model's explanatory power is limited for NO2.
- RSPM/PM10: The negative R-squared (-1.02) suggests that the model doesn't fit well to the RSPM/PM10 data, and its predictive power is worse than a simple mean.

CONCLUSION:

In conclusion, The air quality assessment project in Tamil Nadu aims to analyze air quality trends, identify pollution hotspots, and develop a predictive model for RSPM/PM10 levels. The project uses IBM Cognos for data visualization and Python for analysis, providing insights into pollutants like NO2, SO2, and RSPM/PM10. Visualizations like line charts, bar charts, heatmaps, and pie charts help understand the distribution of air quality indicators. Performance metrics like MAE, MSE, RMSE, and R-squared values highlight the model's accuracy and explanatory power. This project contributes to better understanding of air quality and environmental management.