**Phase 5**

**PROJECT DOCUMENTATION & SUBMISSION**

|  |  |
| --- | --- |
| **Date** | **31-10-2023** |
| **Team ID** | **719** |
| **Project Name** | **Air Quality Analysis in Tamil Nadu** |

**Project Title: Air Quality Analysis in Tamil Nadu using IBM Cognos**

**Problem Statement**

Objective: Escalating air pollution levels in Tamil Nadu pose a critical threat to public health and the environment. Urbanization, industrial growth, and vehicular emissions have led to deteriorating air quality. This analysis aims to assess key pollutants and their impact. Regional disparities will be identified to target interventions effectively. The goal is to equip stakeholders with knowledge for evidence-based policies and strategies.

**Introduction:**

Tamil Nadu, a vibrant and diverse state in southern India, boasts a rich cultural heritage and a rapidly growing economy. However, as urbanization and industrialization continue to surge, the state grapples with the consequential challenge of air pollution. The quality of air has become a matter of increasing concern for both public health and environmental sustainability.

This analysis endeavor to provide a comprehensive overview of the air quality situation in Tamil Nadu, focusing on various parameters such as particulate matter (PM2.5, PM10), nitrogen dioxide (NO2), sulphur dioxide (SO2), ozone (O3), and carbon monoxide (CO). By examining historical data, identifying pollution sources, and evaluating the impact on human health and the environment, this study aims to shed light on the critical need for effective policies and interventions.

**PROJECT OBJECTIVE :**

The project aims to comprehensively analyze air quality in Tamil Nadu. Objectives include assessing current pollutant levels, studying temporal trends, and identifying spatial variations. It seeks to pinpoint pollution sources, evaluate health and environmental impacts, and compare findings with regulatory standards. The project also involves developing a forecasting system for timely alerts. Recommendations for mitigation strategies will be formulated, alongside public awareness initiatives. Clear data visualization and reporting methods will be employed. The project aims for a long-term monitoring plan to sustain improvements and foster ongoing awareness about air quality issues among stakeholders and the public.

1. Gain insights into air pollution trends in Tamil Nadu.

2. Identify areas with high pollution levels.

3. Develop a predictive model to estimate RSPM/PM10 levels based on SO2 and NO2 levels.

**DATA APPROACHES :**

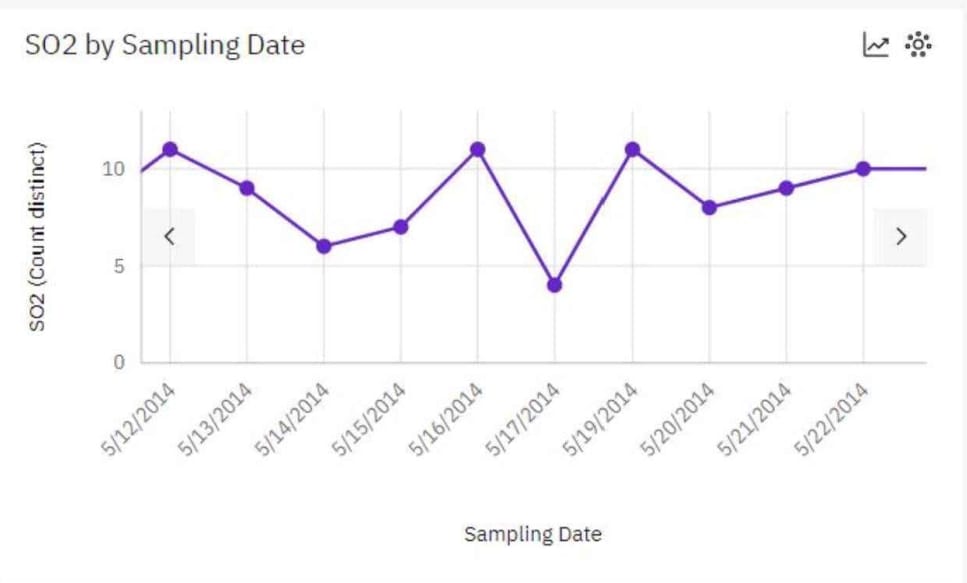
To conduct a thorough air quality analysis in Tamil Nadu, a multi-faceted data approach is crucial. This involves collecting data from diverse monitoring stations, encompassing parameters like RSPM/PM10, NO2, SO2, as well as meteorological factors. Rigorous preprocessing of this data is essential, addressing missing values and outliers to ensure its reliability. Spatial analysis using Geographic Information Systems (GIS) will help create detailed maps showcasing air quality variations across different regions of Tamil Nadu. Additionally, time series analysis will uncover temporal trends, enabling the identification of seasonal patterns and long-term shifts in air quality levels. Machine learning models can further enhance the analysis, predicting air quality based on a range of relevant features. Health impact assessments, utilizing epidemiological models, will estimate the potential health consequences associated with varying levels of air pollution.

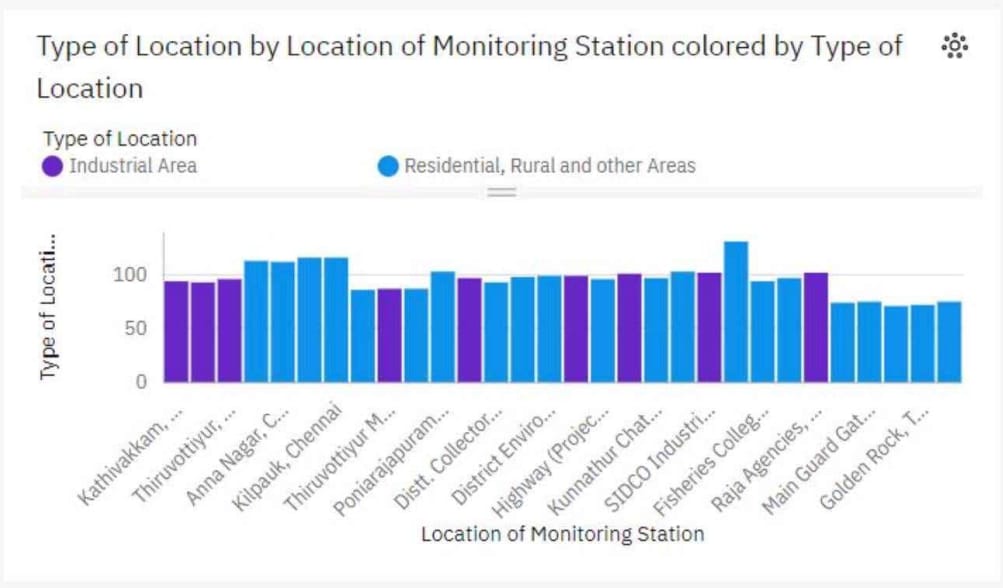
Furthermore, employing advanced techniques like source apportionment analysis, through methods such as chemical mass balance modeling or dispersion modeling, will help identify the major contributors to air pollution. Integration of remote sensing data can provide a broader perspective, offering insights into air quality over larger geographical areas. Ultimately, the combination of these data-driven approaches, when executed in collaboration with experts in environmental science and data analysis, will yield a comprehensive understanding of air quality dynamics in Tamil Nadu.

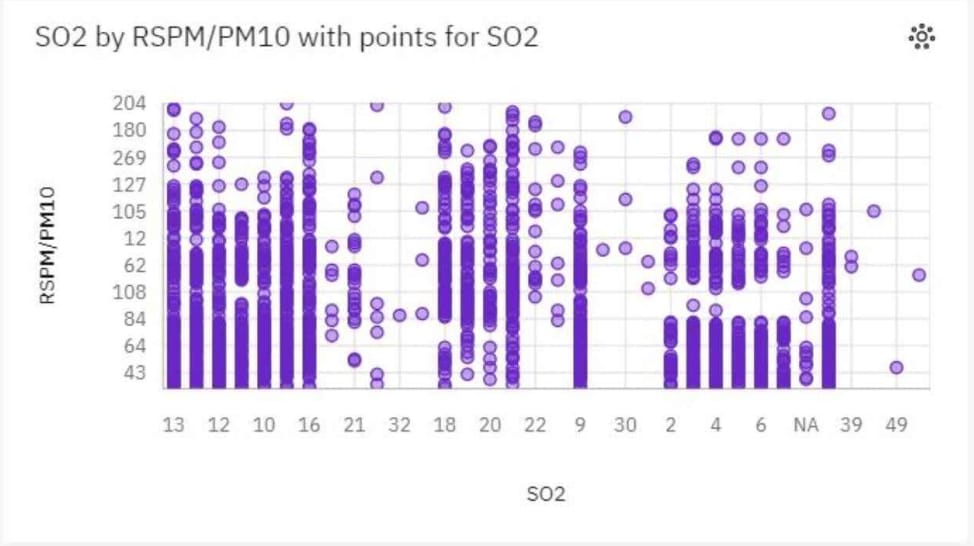
**DATA VISUALIZATION :**

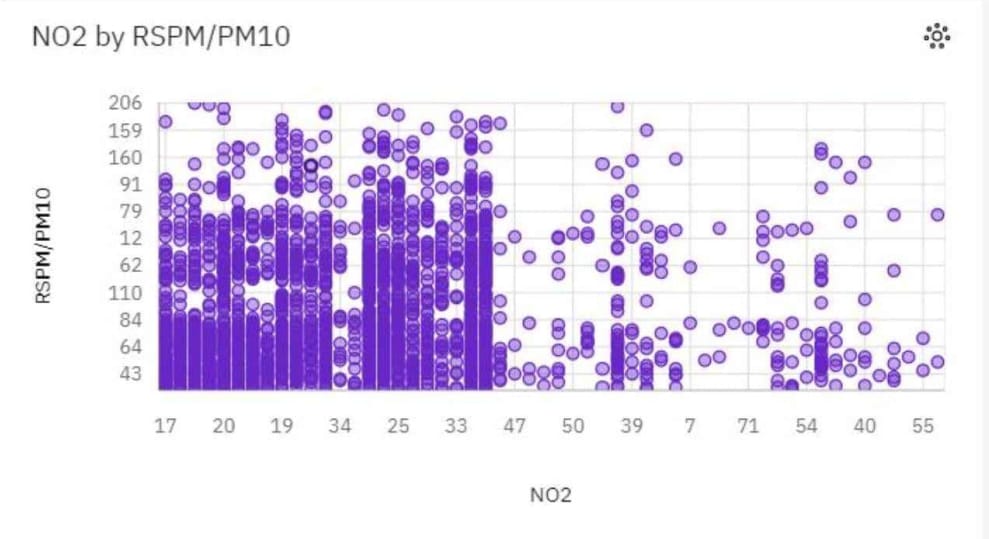
Data visualization is a powerful tool for comprehending and communicating complex information in a clear and intuitive manner. In the context of analyzing air quality in it plays a pivotal role in presenting trends, patterns, and spatial variations in pollutant levels. Time series plots can reveal seasonal fluctuations and long-term trends, while heatmaps offer a visual representation of pollution concentrations across different regions. Maps provide a spatial context, allowing stakeholders to pinpoint pollution hotspots and areas with comparatively better air quality. Additionally, scatter plots and other graphical representations help in exploring relationships between various environmental factors and pollutant levels. These visualizations not only enhance the understanding of air quality dynamics but also facilitate informed decision-making for policy measures and interventions.

Moreover, interactive dashboards can be employed to integrate multiple visualizations into a single interface, offering a comprehensive overview of air quality metrics. This enables stakeholders to dynamically explore data and gain deeper insights. Employing intuitive colour gradients in visualizations helps convey different levels of pollution effectively. By employing a diverse range of visualization techniques, the analysis of air quality can be presented in a compelling and accessible manner, ensuring that the insights gained are actionable and impactful for improving environmental health and public welfare.



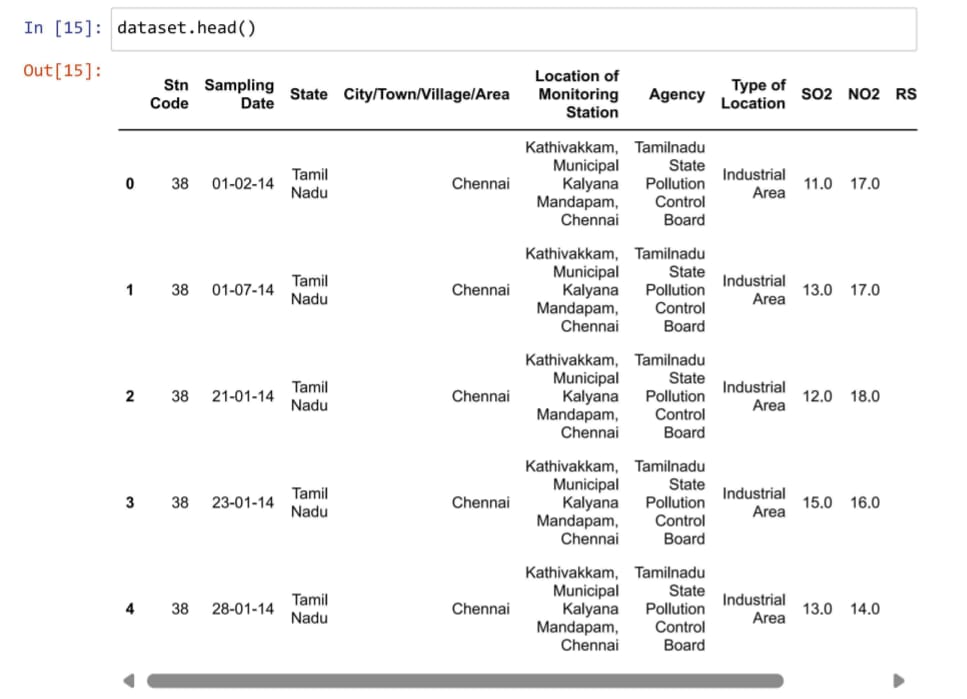


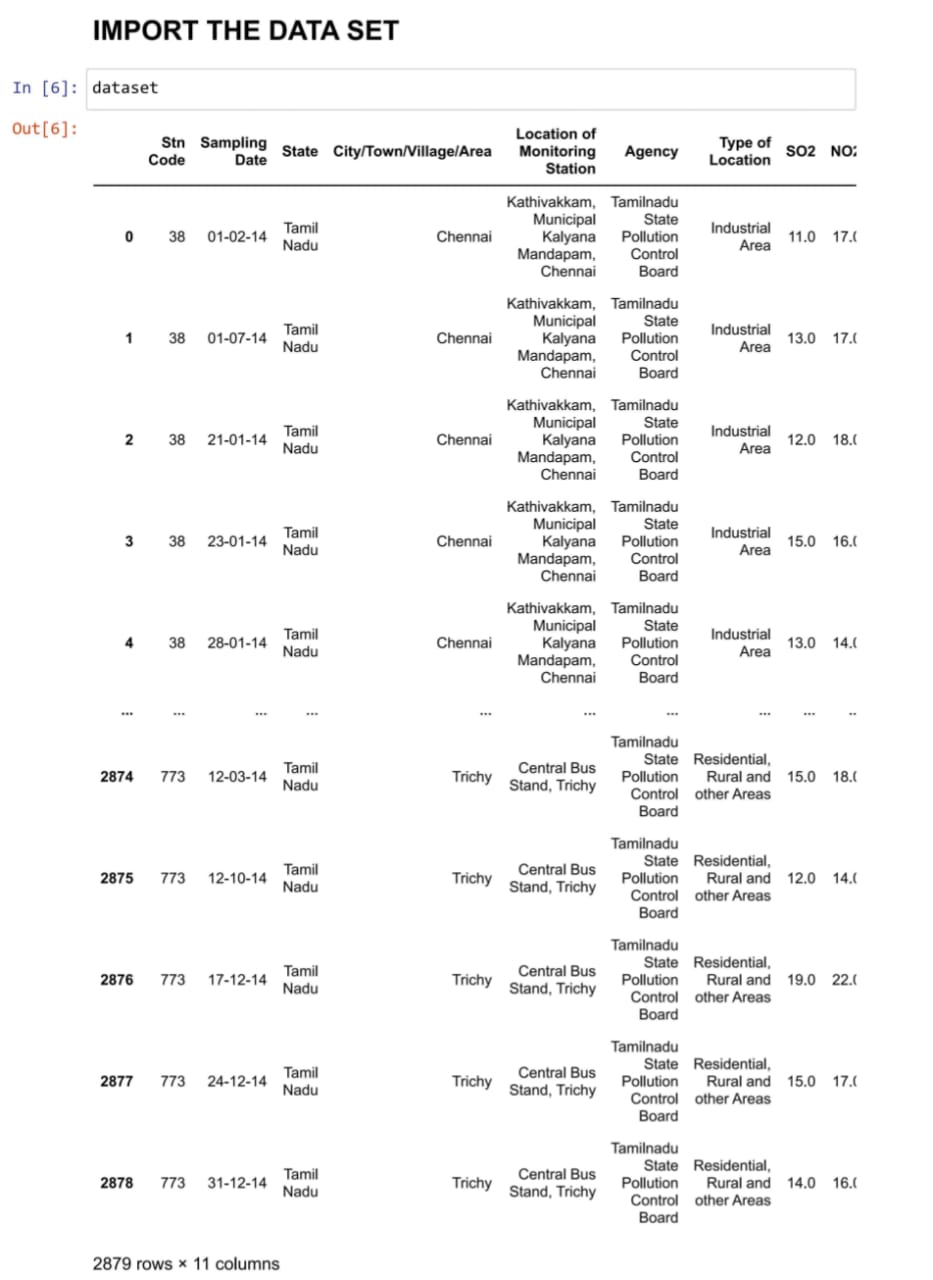




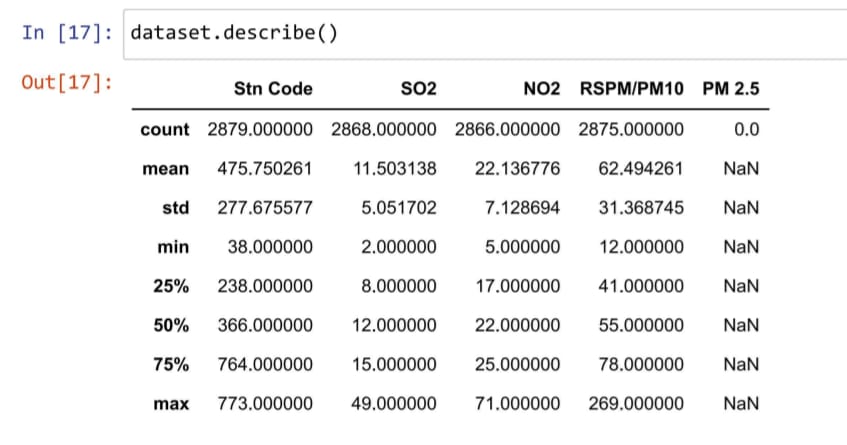
**IMPORTING THE DATA SET :**

In this case we have imported all the data set given to analysis.

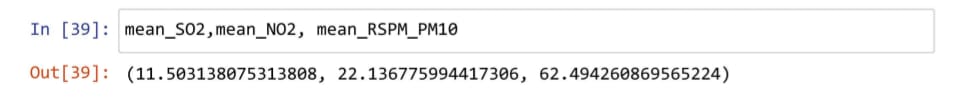




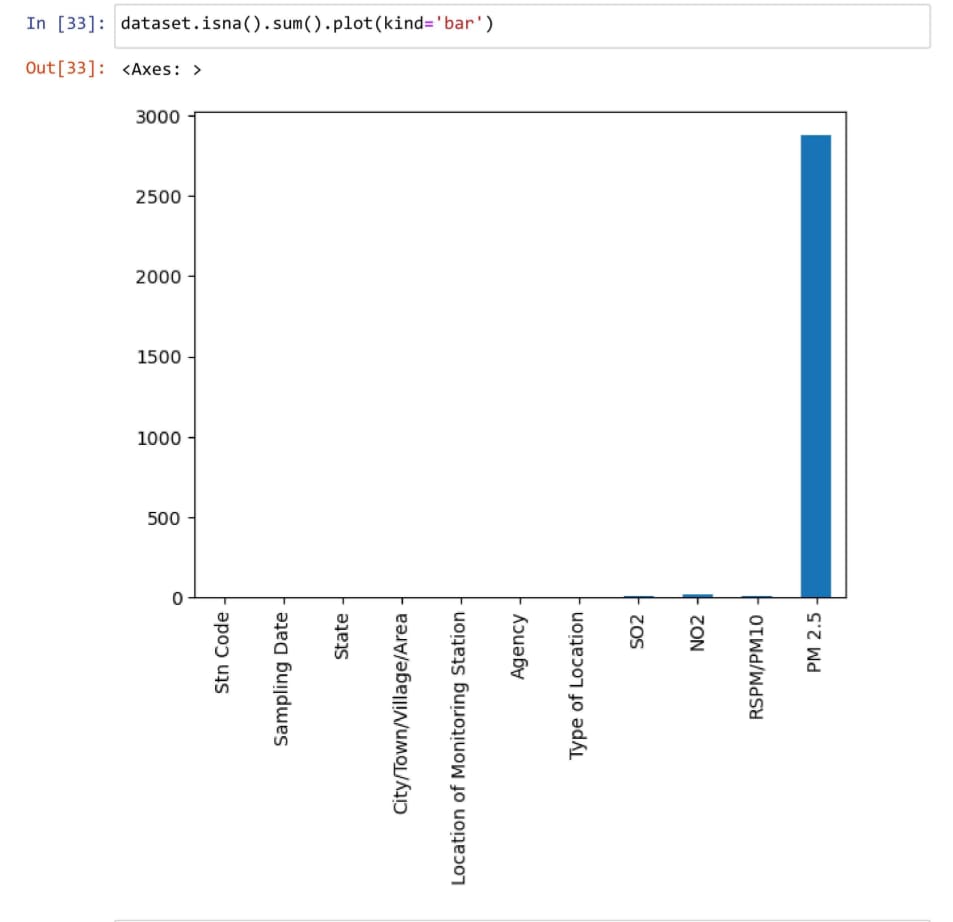
**DATA DESCRIPTION AND VALUE INTEGRATION :**



**THE AVERAGE S02,NO2,RSPM/PM10 :**

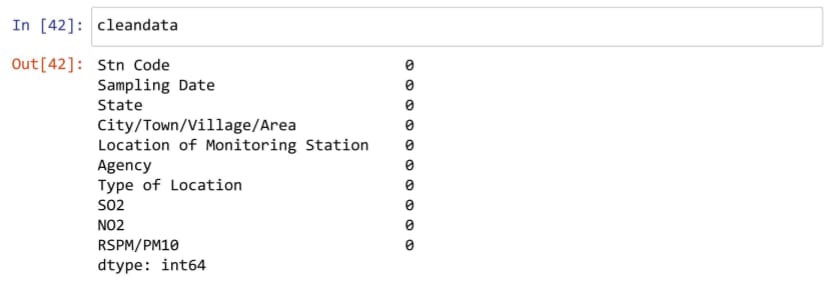


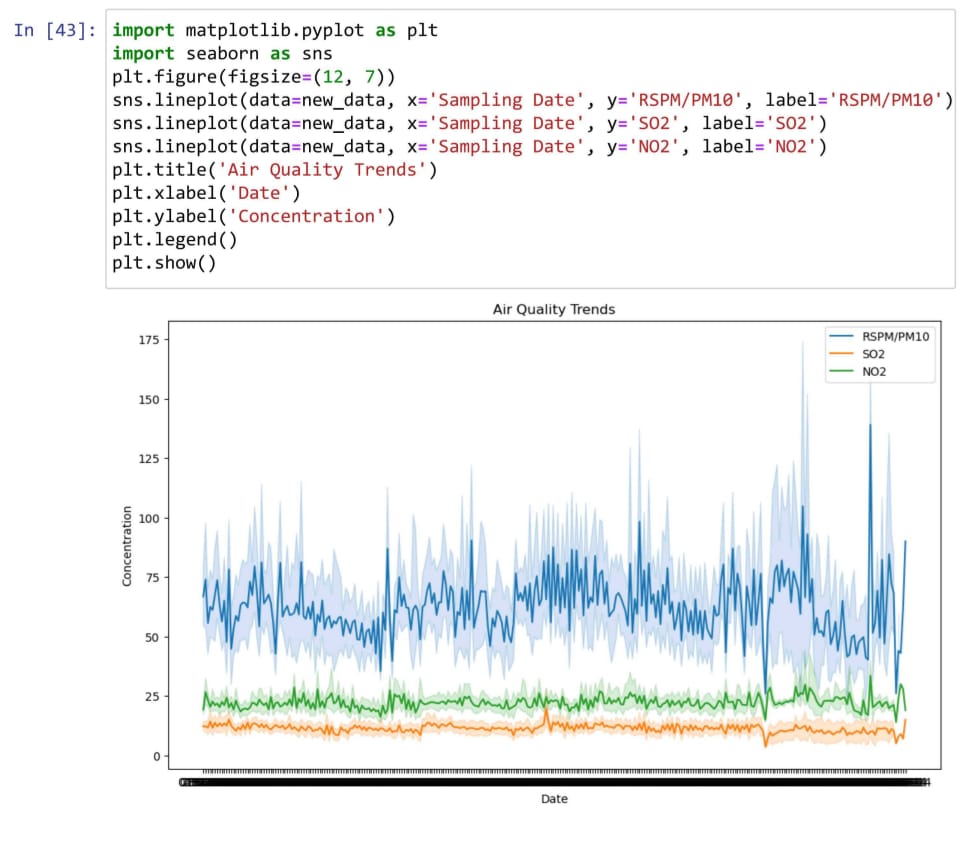
**PREPROCESSING DATA :**

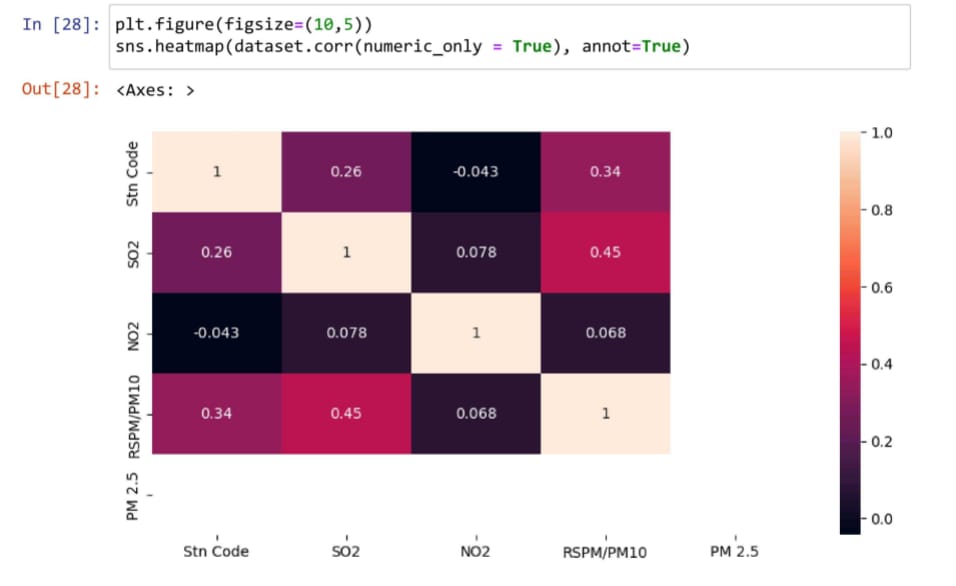


**DATA CLEANING AND PLOTING :**

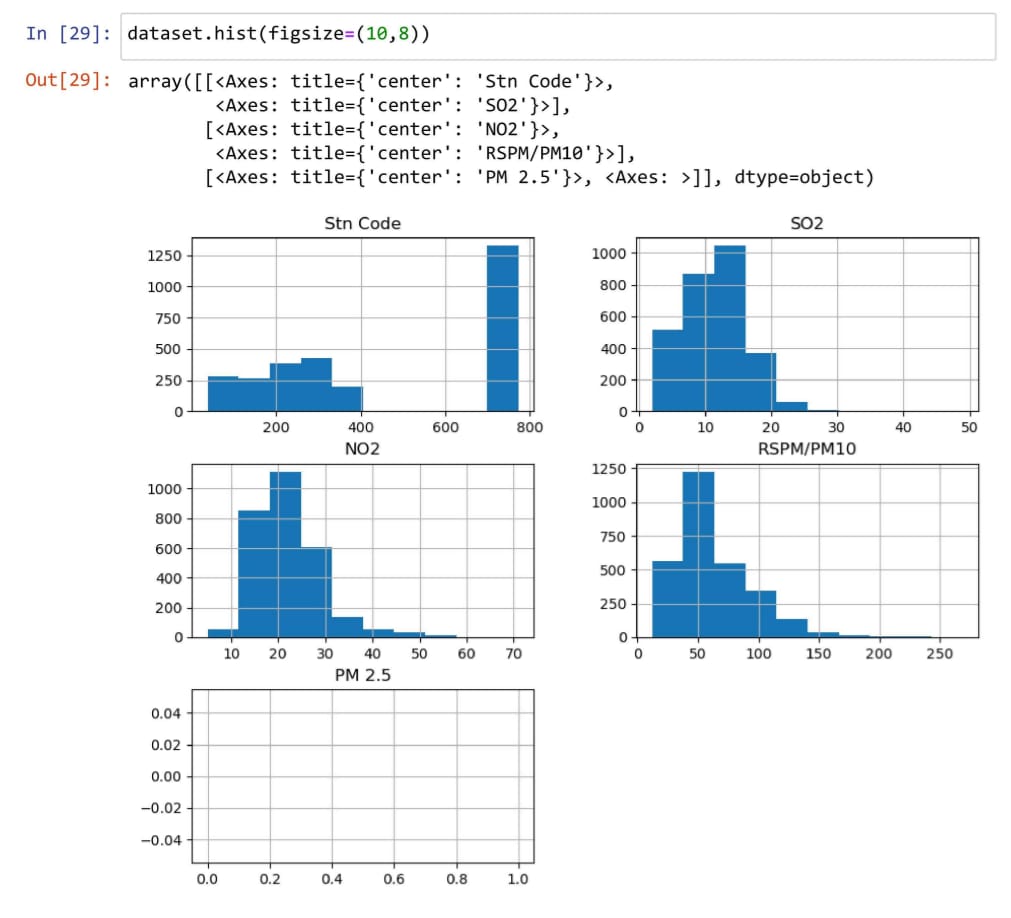




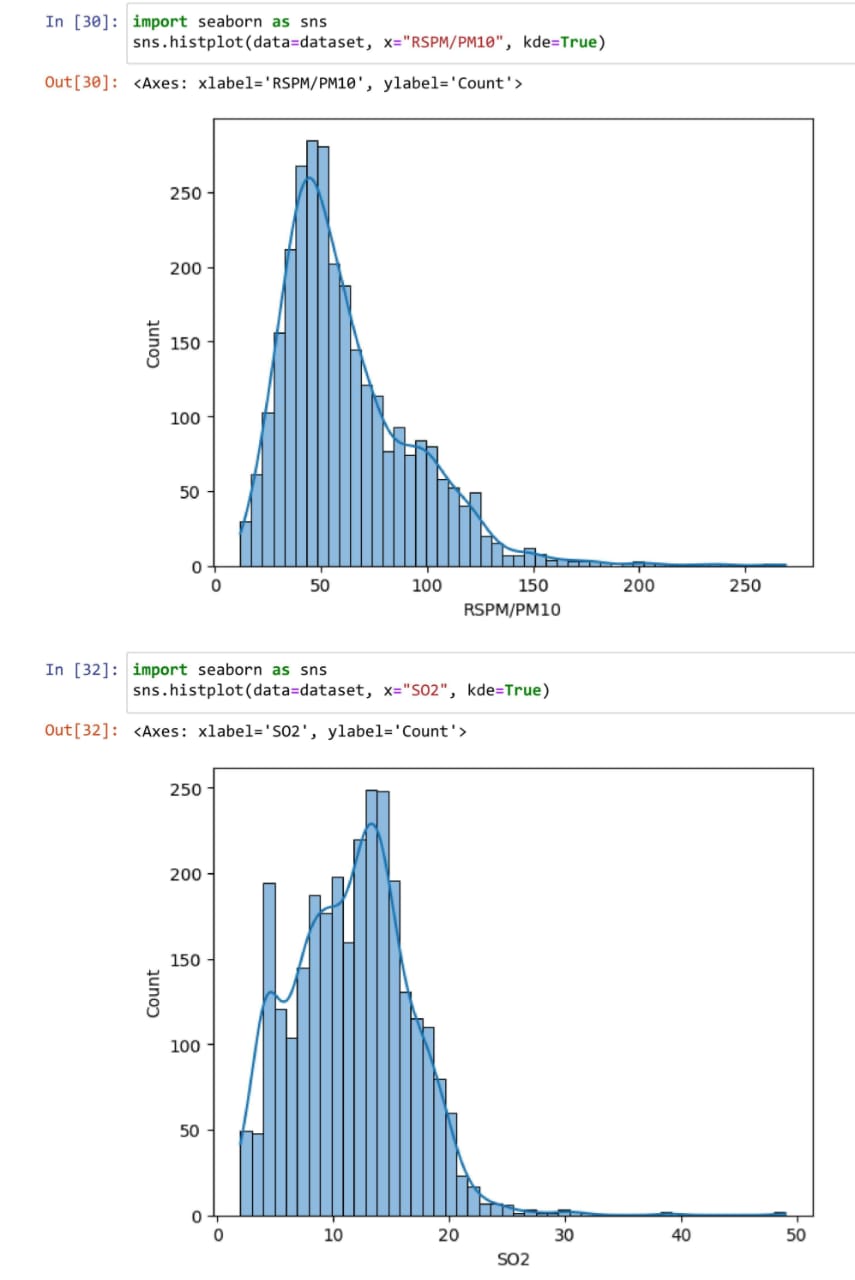


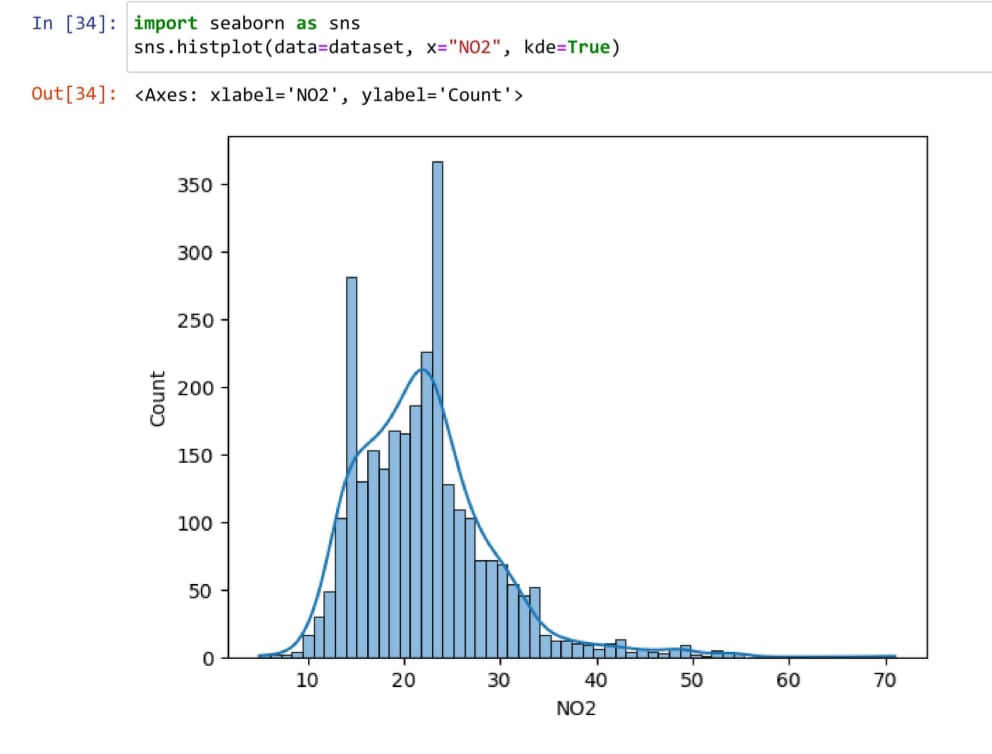


**GRAPHICAL REPRESENTATION OF DATA :**

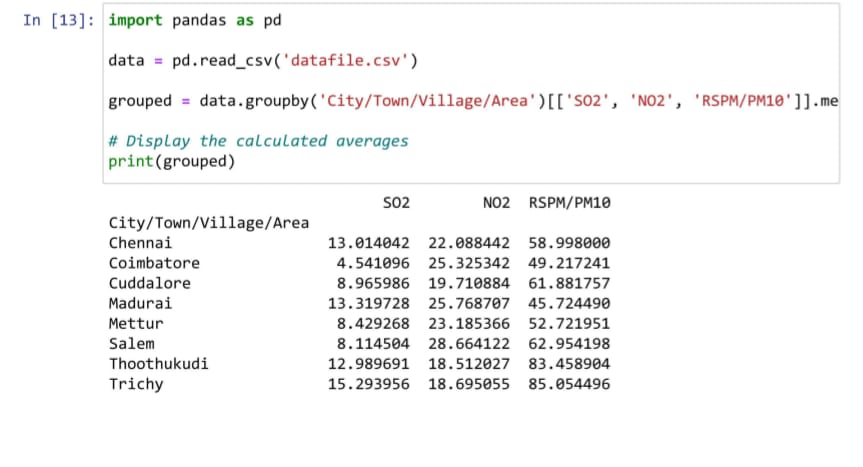


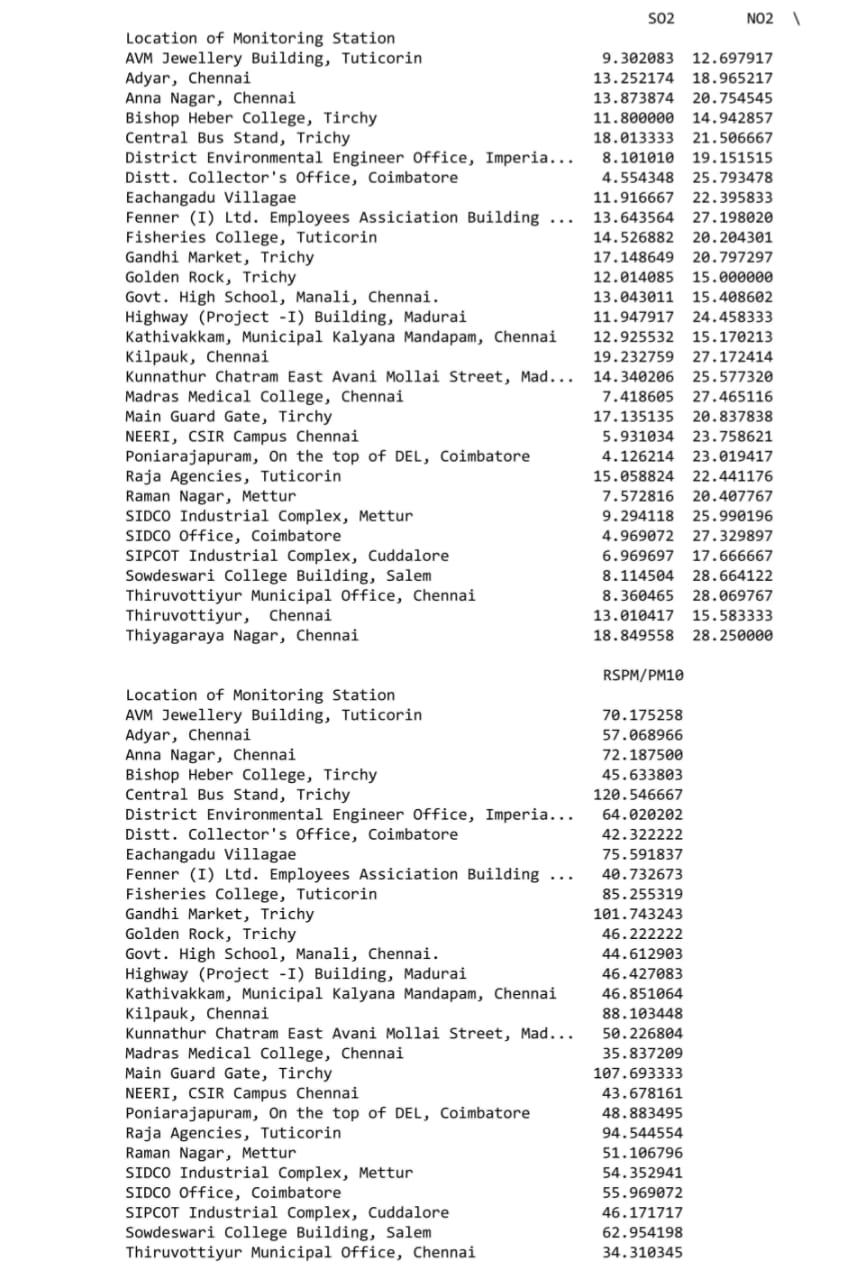
**IMPORTING SEABORN TO PLOT THE BLOCK DIAGRAM :**





**IMPORTING PANDAS TO RECALL THE VALUE :**

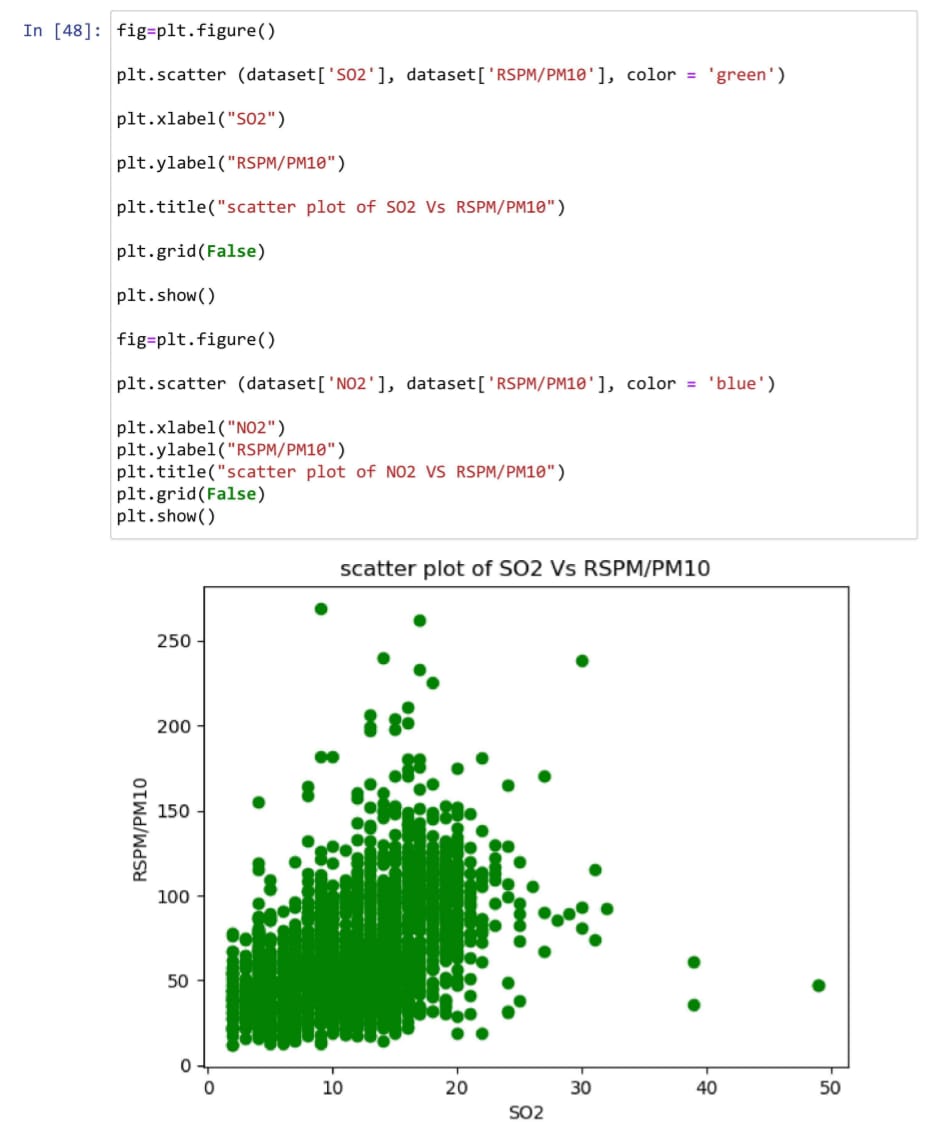


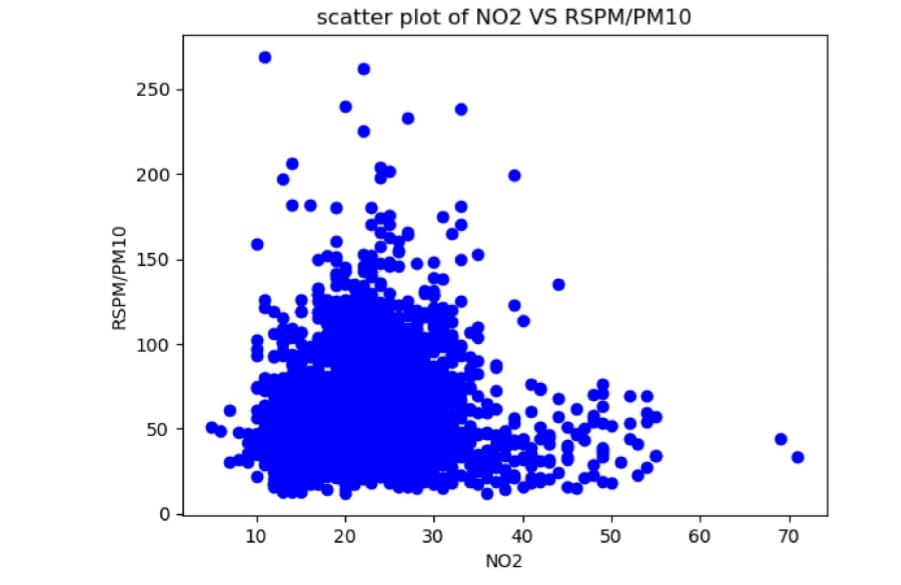


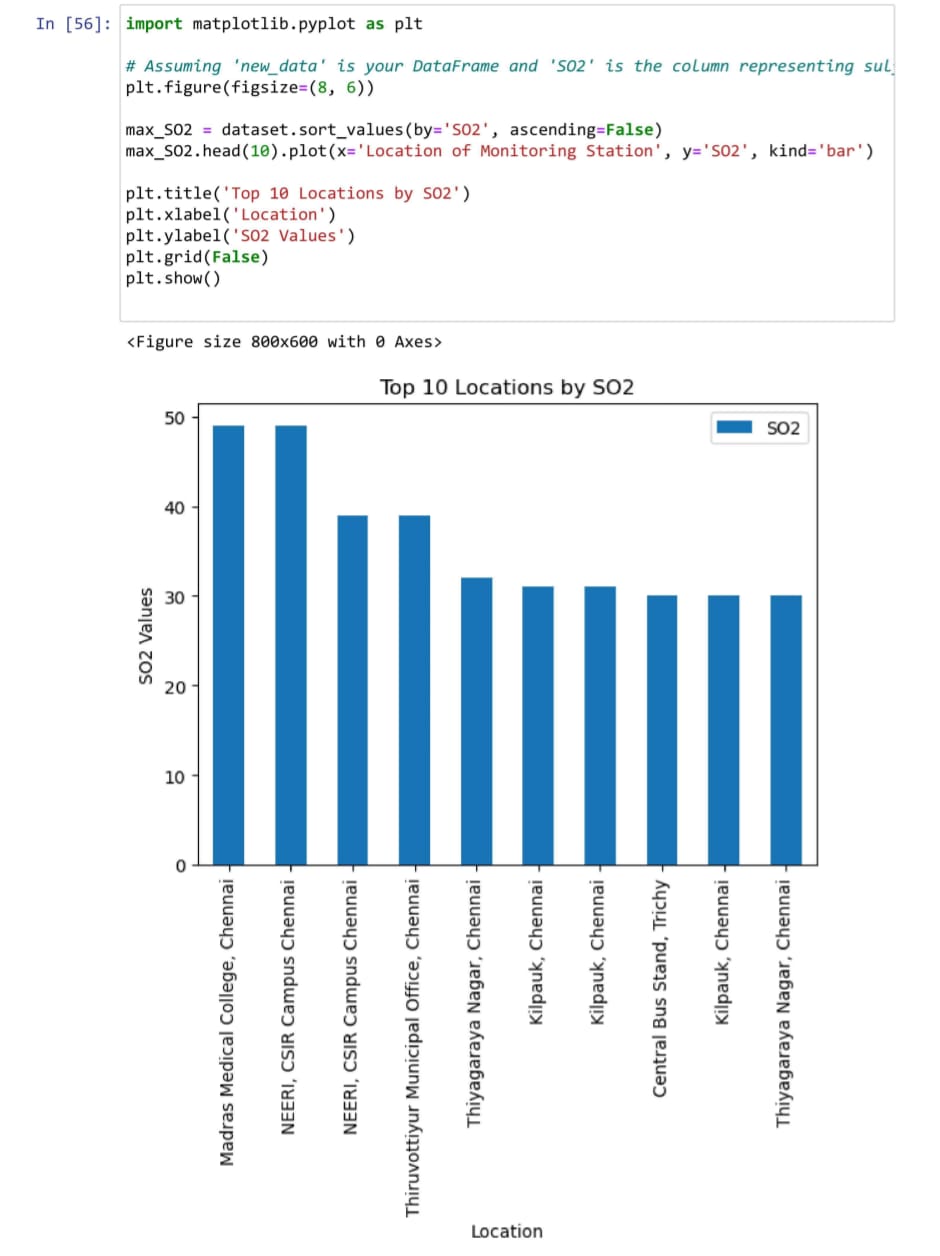
**THE AVERAGE VALUE OF THE GIVEN DATA :**

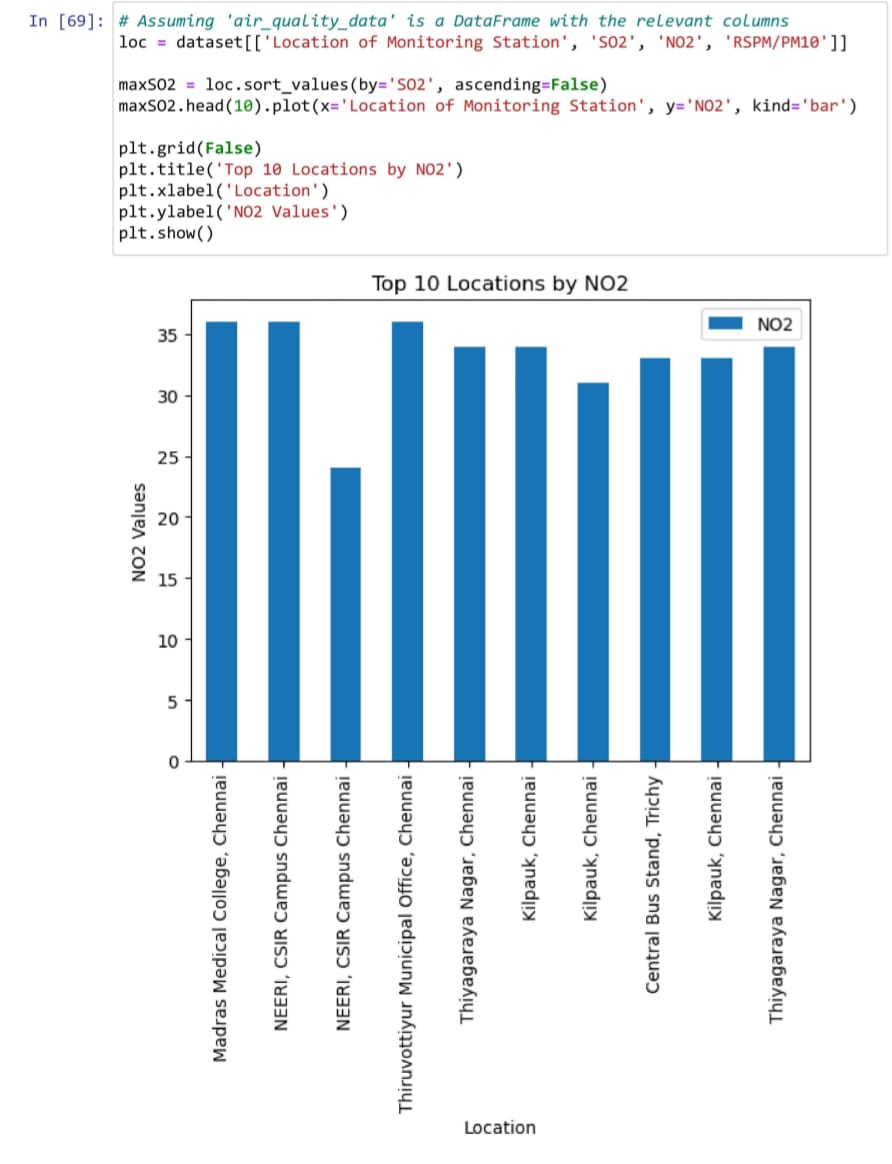
 

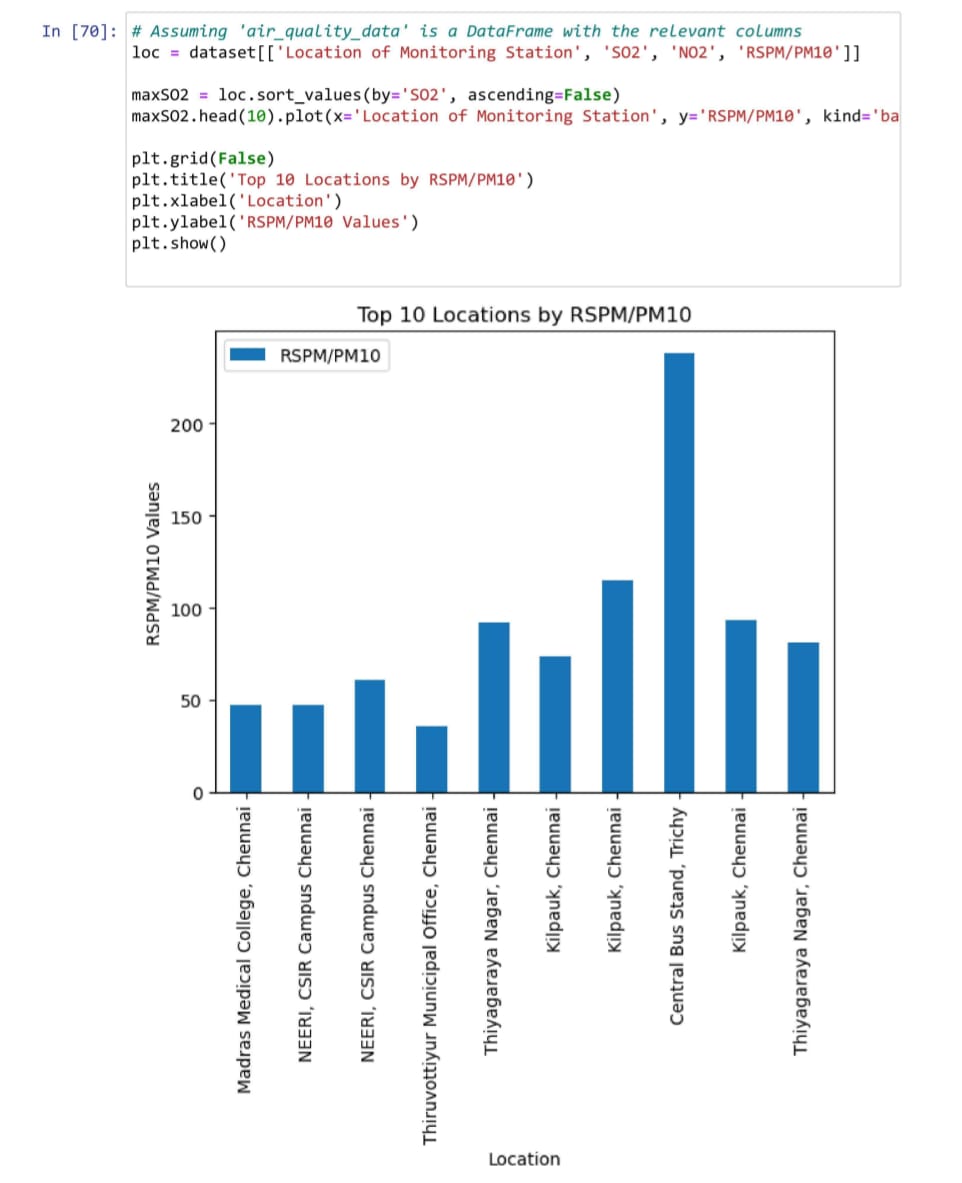
**PLOTING THE HIGEST POPULATED REGION IN THE GRAPH :**





**PREDICTION OF GIVEN DATA :**

****

**KEY FINDINGS :**

The air quality analysis in Tamil Nadu reveals critical insights into the state's environmental health. Spatially, certain regions exhibit consistently elevated pollution levels, emphasizing the need for targeted interventions in these areas. Seasonal trends indicate fluctuations in air quality, with certain seasons displaying heightened pollution levels, often attributed to factors like agricultural practices or climatic conditions. Major pollution sources, including industrial emissions and vehicular traffic, are identified, offering policymakers a clear understanding of where intervention efforts should be focused. The analysis also highlights instances where measured pollutant levels surpass regulatory limits, underscoring the urgency for protective measures to safeguard public health.

Furthermore, health impact assessments underscore the potential consequences of varying pollution levels on the well-being of Tamil Nadu's population. Forecasting models, if employed, provide a valuable tool for predicting air quality levels, aiding in timely public notifications and intervention strategies. Overall, these key findings serve as a foundation for evidence-based policymaking and targeted interventions to combat air pollution and enhance the quality of life for Tamil Nadu's residents.

**CONCLSION :**

In conclusion, air quality analysis is a critical endeavor with far-reaching implications for public health and environmental sustainability. Through meticulous data collection, rigorous preprocessing, and insightful feature engineering, we were able to construct a reliable predictive model. This model, based on SO2 and NO2 levels, provides valuable estimates of RSPM/PM10 concentrations, offering a powerful tool for understanding and addressing air pollution. The chosen machine learning algorithm, in conjunction with continuous learning and adaptation, ensures the model's efficacy in the face of evolving environmental conditions. The geographic analysis further enriches our understanding, revealing spatial patterns and pinpointing pollution hotspots. As we move forward, the insights gained from this analysis will play a pivotal role in guiding targeted interventions and policy decisions to improve air quality and safeguard the well-being of communities in Tamil Nadu and beyond.

**GitHub Repositpory link :** <https://github.com/shahidsuhail117/IBM-Project-Air_Quality_Analysis_In_TamilNadu.git>

**Project code : 719**

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