**AIR QUALITY ASSESSMENT TAMIL NADU**

**INTRODUCTION :**

**Data from Tamil Nadu monitoring stations will be analyzed and visualized as part of the project. The goal is to learn more about the extent of air pollution and create a forecast model to calculate RSPM/PM10 levels based on SO2 and NO2 levels. In this project, the goals are set, the analysis strategy is designed, visualization approaches are chosen, and a predictive model is built using Python and pertinent libraries.**

**OBJECTIVES :**

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

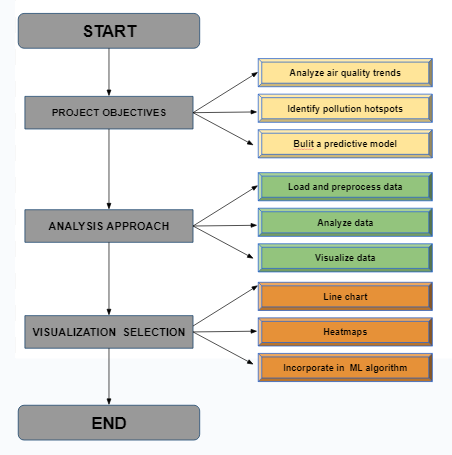
* ****RSPM :****

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.

**FLOW** **CHART :**

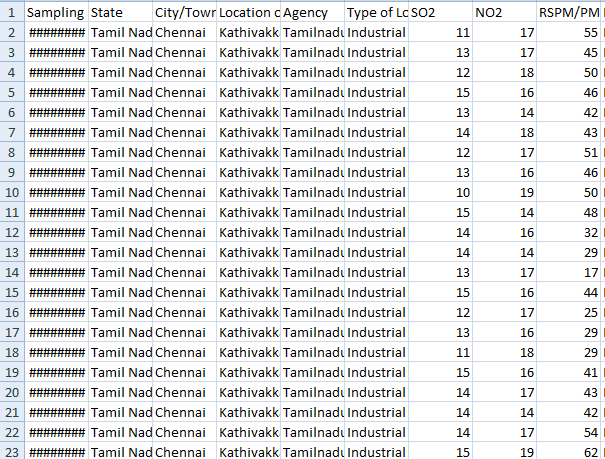


****PROJECT**** ****OBJECTIVES :****

* + ****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.

****ANALYSIS APPROACH :****

* ****Load and preprocess**** ****data :** This data is extracted from**[**https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014**](https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014)



* + **Analyze** **data :** Conduct various analyses, which may include statistical tests, time-series analysis, and spatial analysis, to derive meaningful insights.
  + **Visualize** **data :**: Create visual representations of the data to make it more understandable. Visualization is a powerful tool for conveying insights.

****VISUALIZATION**** ****SELECTION :****

* **Line** **chart :** Using line charts to illustrate trends and changes over time.
* **Heatmap :** Employing heatmaps to depict the concentration of pollutants across different regions.
* **Incorporate in ML** **algorithm :** Optionally, consider integrating machine learning algorithms to enhance the accuracy of your predictive model . We have used the K-Nearest Neighbour algorithm.
  + - **KNN** **algorithm :** The k-Nearest Neighbors (KNN) algorithm is a machine learning method for classification and regression. It makes predictions based on the similarity of new data points to the k nearest neighbors in the training dataset**.**
    - **Modules :** to implement KNN algorithm in python we use certain modules and libraries such as Numpy, pandas, matplotlib, seaborn, sklearn .