**AIR QUALITY ANALYSIS IN TAMILNADU**

**Introduction:**

The aim of this document is to outline the problem statement and design thinking process for the project that focuses on analyzing and visualizing air quality data from monitoring stations in Tamil Nadu. The primary objective of this project is to gain valuable 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 document will provide a comprehensive overview of the project's objectives, the proposed analysis approach, and the selection of visualization techniques.

**Project Objectives:**

The project's main objectives can be summarized as follows:

**Analyzing Air Quality Trends:**

- Understand the historical trends of air quality parameters in Tamil Nadu.

- Identify seasonal variations and long-term trends in air pollution.

**Identifying Pollution Hotspots:**

- Locate regions with consistently high levels of air pollution.

- Determine the key contributing factors in these pollution hotspots.

**Building a Predictive Model:**

- Develop a predictive model that estimates RSPM/PM10 levels.

- Utilize SO2 and NO2 levels as input variables for the model.

- Assess the model's accuracy and reliability for forecasting air quality.

**Analysis Approach**

To achieve the project objectives, we propose the following analysis approach:

**Data Collection**:

- Gather historical air quality data from monitoring stations across Tamil Nadu.

- Ensure data consistency and reliability.

**Data Preprocessing:**

- Clean and preprocess the collected data to handle missing values and outliers.

- Standardize data to ensure uniformity across different stations.

**Exploratory Data Analysis (EDA):**

- Conduct EDA to uncover insights, trends, and patterns in the air quality data.

- Perform statistical analysis to identify correlations and dependencies among variables.

**Visualization Selection:**

- Select appropriate visualization techniques to effectively represent air quality trends and pollution levels.

- Consider using line charts, heat maps, and geographical maps for visualization.

**Predictive Model Development:**

- Utilize Python and relevant libraries to build a predictive model.

- Split the dataset into training and testing sets for model evaluation.

- Train the model using SO2 and NO2 levels as independent variables and RSPM/PM10 levels as the target variable.

- Evaluate the model's performance using appropriate metrics.

**Conclusion:**

This problem statement document clearly explains the project's objectives, design thinking, and the proposed approach to analyzing and visualizing air quality data in Tamil Nadu. By following the outlined plan, we aim to gain insights into air pollution trends, identify pollution hotspots, and develop a predictive model that can contribute to better environmental management and public health in the region. The successful execution of this project will rely on thorough data analysis, effective visualization, and the development of a reliable predictive model.