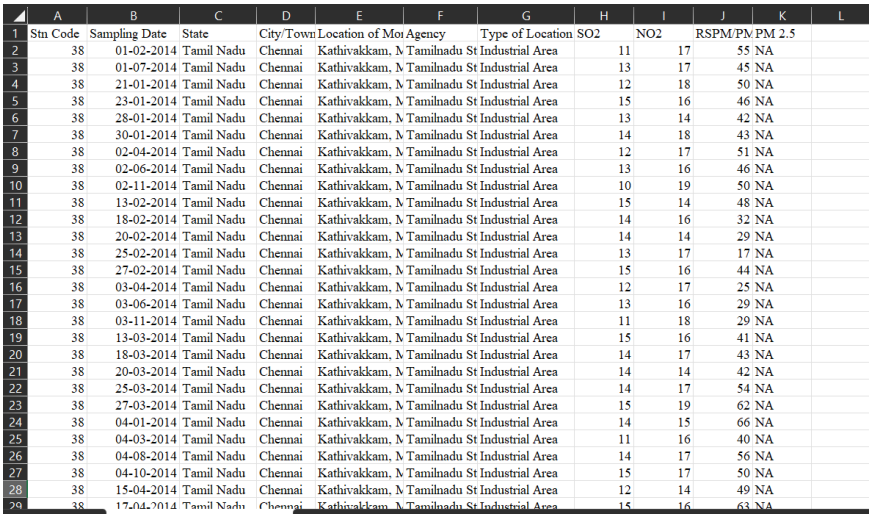
Project Definition and Design Thinking: Air Quality Analysis in Tamil Nadu

Problem Definition:

Air pollution is a significant environmental and public health concern in Tamil Nadu, India. The state experiences a wide range of air quality issues, including high levels of particulate matter (RSPM/PM10) and gaseous pollutants like sulfur dioxide (SO2) and nitrogen dioxide (NO2). Understanding and addressing these air quality challenges are crucial for improving public health and environmental quality.

The project aims to analyze and visualize air quality data from monitoring stations in Tamil Nadu. The objective is to gain insights into air pollution trends, identify areas with high pollution levels, and to visualize RSPM/PM10 levels based on SO2 and NO2 levels. This project involves defining objectives, designing the analysis approach and selecting visualization techniques using IBM Cognos.

The dataset employed for this project is ".csv" file from the link : **https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014**



	A	B	C	D	E	F	G	H	I	J	K	L
1	Stn Code	Sampling Date	State	City/Town	Location of Mo	Agency	Type of Location	SO2	NO2	RSPM/PM	PM 2.5	
2	38	01-02-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	11	17	55	NA	
3	38	01-07-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	13	17	45	NA	
4	38	21-01-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	12	18	50	NA	
5	38	23-01-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	15	16	46	NA	
6	38	28-01-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	13	14	42	NA	
7	38	30-01-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	18	43	NA	
8	38	02-04-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	12	17	51	NA	
9	38	02-06-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	13	16	46	NA	
10	38	02-11-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	10	19	50	NA	
11	38	13-02-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	15	14	48	NA	
12	38	18-02-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	16	32	NA	
13	38	20-02-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	14	29	NA	
14	38	25-02-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	13	17	17	NA	
15	38	27-02-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	15	16	44	NA	
16	38	03-04-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	12	17	25	NA	
17	38	03-06-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	13	16	29	NA	
18	38	03-11-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	11	18	29	NA	
19	38	13-03-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	15	16	41	NA	
20	38	18-03-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	17	43	NA	
21	38	20-03-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	14	42	NA	
22	38	25-03-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	17	54	NA	
23	38	27-03-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	15	19	62	NA	
24	38	04-01-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	15	66	NA	
25	38	04-03-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	11	16	40	NA	
26	38	04-08-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	14	17	56	NA	
27	38	04-10-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	15	17	50	NA	
28	38	15-04-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	12	14	49	NA	
29	38	17-04-2014	Tamil Nadu	Chennai	Kathivakkam, N	Tamilnadu St	Industrial Area	15	16	63	NA	

Design Thinking:

1. Project Objectives

Objective 1: Explore Air Pollution Trends

- Analyze historical air quality data to identify patterns and trends in RSPM/PM10, SO2, and NO2 levels.
- Uncover insights into how air quality has evolved over time in Tamil Nadu.

Objective 2: Identify High Pollution Areas

- Determine geographic regions or specific monitoring stations within Tamil Nadu that consistently exhibit high levels of air pollution.

- Facilitate informed decision-making by highlighting areas requiring pollution mitigation efforts.

Objective 3: Visualize RSPM/PM10 Based on SO2 and NO2

- Develop a predictive model using IBM Cognos that estimates RSPM/PM10 levels based on SO2 and NO2 concentrations.
- Create visual representations of these predictions to enhance understanding and awareness of air quality dynamics.

2. Analysis Approach

Step 1: Data Collection and Integration

- Import historical air quality data from monitoring stations in Tamil Nadu into IBM Cognos.
- Ensure data quality and consistency through data cleaning and validation using IBM Cognos data preparation tools.

Step 2: Data Preprocessing

- Handle missing data, outliers, and inconsistencies using IBM Cognos data cleaning capabilities.
- Aggregate data by time periods (e.g., daily, monthly) for trend analysis using IBM Cognos data aggregation functions.
- Prepare geospatial data for hotspot identification, leveraging IBM Cognos' geospatial analytics capabilities.

Step 3: Data Analysis

- Utilize IBM Cognos' built-in statistical techniques to identify trends, patterns, and correlations in air quality parameters.
- Visualize findings through line charts, time series plots, and correlation matrices using IBM Cognos' visualization tools.

Step 4: Identifying Pollution Hotspots

- Employ geospatial tools within IBM Cognos (e.g., spatial analytics) to map pollution levels across Tamil Nadu.
- Generate heatmaps and geographic visualizations to visualize pollution hotspots, leveraging IBM Cognos' geospatial mapping features.

Step 5: Reporting and Visualization

- Create informative reports and interactive dashboards within IBM Cognos to present the project's findings.

- Communicate insights regarding air quality trends, pollution hotspots, and the predictive model's accuracy using IBM Cognos' reporting and dashboarding capabilities.
- Enable users to explore and interact with the visualizations and reports for data-driven decision-making.

3. Visualization Selection:

The following visualization techniques (e.g., line charts, heatmaps) helps to effectively represent air quality trends and pollution levels.

Time Series Visualizations:

- Line charts and time series plots capture trends in air quality parameters over time, aiding in trend identification.

Histograms and Box Plots:

- These visuals present the distribution and variability of pollutant concentrations, allowing for insights into data spread.

Geospatial Heatmaps and Maps:

- Leveraging geospatial capabilities in IBM Cognos, heatmaps and geographic maps reveal spatial patterns of air quality data, particularly useful for identifying pollution hotspots.

Regression and Scatter Plots:

- These visualizations help demonstrate the relationships between pollutants (RSPM/PM10, SO2, NO2), enhancing comprehension of interdependencies.

Interactive Dashboards:

- The interactive dashboards integrate all visualizations and analysis components, enabling users to explore and analyze the data interactively. This promotes data-driven decision-making and enhances the accessibility of insights.

Here are some questions that can be answered through visualization based on the air quality analysis project in Tamil Nadu:

1. How has the air quality in Tamil Nadu changed over the past decade?
2. Which cities or regions in Tamil Nadu consistently experience the highest levels of air pollution?
3. Are there seasonal variations in air quality in different regions of Tamil Nadu?
4. What is the distribution of SO2, NO2, and RSPM/PM10 concentrations in the state?
5. Are there any correlations between the levels of different pollutants (e.g., SO2 and NO2)?

6. How do pollutant levels in urban areas compare to those in rural areas?
7. Has there been a significant improvement or deterioration in air quality after the implementation of specific pollution control measures in certain regions?
8. What is the impact of industrial emissions on air quality in specific industrial zones within Tamil Nadu?
9. Can we predict RSPM/PM10 levels based on SO2 and NO2 concentrations accurately?
10. How does air quality vary between coastal areas and inland regions in Tamil Nadu?