NAAN MUDHALVAN

COURSE : DATA ANALYTICS WITH COGNOS

Phase 1: Problem definition and Design Thinking

# *Topic: Air Quality Analysis in Tamil Nadu*


**Team Members-**

Sangamithra T - au2021103573

Sreenithi A - au2021103051

Shrivarshini Kp - au2021103048

Shri Sakitha D - au2021103047

# TABLE OF CONTENTS

## PROBLEM STATEMENT:

Air pollution poses a significant environmental and public health challenge in Tamil Nadu, India, with adverse effects on the well-being of its residents. Rapid urbanization, industrial growth, vehicular emissions, and seasonal factors have contributed to deteriorating air quality levels across the state. Poor air quality poses severe health risks to the population and has adverse effects on the environment.

To address this critical issue, this project seeks to analyze and visualize air quality data collected from monitoring stations across Tamil Nadu. The primary objective is to gain comprehensive insights into air pollution trends within the region, identify areas that consistently exhibit high pollution levels, and ultimately develop a predictive model to estimate RSPM/PM10 levels. The predictive model will be based on critical pollutants, specifically sulfur dioxide (SO2) and nitrogen dioxide (NO2). This project aims to provide a holistic understanding of air quality issues in Tamil Nadu and to guide policymakers, environmentalists, and stakeholders in implementing measures to safeguard the health and well-being of the state's residents while preserving the environment.

## PROJECT DEFINITION:

The project aims to analyze and visualize air quality data from monitoring stations in TamilNadu. The objective is to gain 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 project involves defining objectives, designing the analysis approach, selecting visualization techniques, and creating a predictive model using Python and relevant libraries.

## PROJECT DESIGN:

### OBJECTIVE:

In the initial design thinking phase, it's imperative to clearly define the project's objectives. For this air quality analysis project, we aim to go beyond just data analysis. We aspire to not only analyze air quality trends but also gain insights into air pollution trends and track the underlying causes.

**i. Analyzing Air Quality Trends**: To gain comprehensive insights into historical air quality data trends, specifically focusing on RSPM/PM10 levels, SO2, and NO2 concentrations.This objective extends beyond surface-level analysis. We aim to delve deep into historical air quality data, using statistical methods and data mining techniques to understand the drivers behind air pollution trends..

**ii. Identifying Pollution Hotspots**: Identifying pollution hotspots involves pinpointing geographic areas within Tamil Nadu that consistently exhibit high air pollution levels, thereby identifying pollution hotspots. This objective requires not only data analysis but also geospatial analysis and potentially even qualitative research to determine the root causes.

**iii. Building a Predictive Model:** Developing a predictive model for RSPM/PM10 levels is a complex task. In the design thinking process, we recognize the need to involve domain experts to develop an accurate predictive model capable of estimating RSPM/PM10 levels using SO2 and NO2 concentrations as key variables.

**ANALYSIS APPROACH:**

Our analysis approach involves the following steps:-

**i. Data Loading**: Collect and integrate air quality data from various monitoring stations in Tamil Nadu, ensuring data quality and consistency.

**ii. Data Preprocessing**: Clean and prepare the data, addressing missing values, outliers, and any inconsistencies.

**iii. Data Analysis**: Utilize statistical methods and data visualization techniques to uncover air quality trends over time and identify significant patterns.

**iv. Hotspot Identification**: Implement spatial analysis and clustering techniques to pinpoint pollution hotspots within Tamil Nadu.

**v. Model Development**: Employ machine learning or statistical modeling techniques to construct a predictive model for RSPM/PM10 levels based on SO2 and NO2 concentrations.

**vi. Model Evaluation**: Assess the model's accuracy and performance using appropriate validation methods.

**VISUALIZATION TECHNIQUES:**

We aim to effectively communicate our findings by employing :-

**i. Time Series Charts**: Utilize line charts and time series plots to depict historical air quality trends, allowing stakeholders to visualize fluctuations over time.

**ii. Heatmaps**: Employ heatmaps to spatially represent pollution hotspots, offering a clear and intuitive view of regions with high pollution levels.

**iii. Geospatial Mapping**: Incorporate geographic information system (GIS) tools to create interactive maps, enhancing the understanding of pollution distribution.

By applying design thinking principles to this project, we ensure a user-centric and goal-driven approach, ultimately facilitating better decision-making for air quality management in Tamil Nadu.

## DATASET:

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

## PROJECT SCOPE:

The insights that can be derived from the Air Quality Analysis Project in Tamil Nadu, along with the project's scope, can include the following:

**i. Air Quality Trends and Variability:** The project can provide insights into the historical trends and seasonal variations in air quality across different regions of Tamil Nadu. This can help identify whether air quality is improving, worsening, or remaining stable over time.

**ii. Identification of Pollution Sources:** By analyzing data and conducting source apportionment studies, the project can pinpoint the major contributors to air pollution in the state. This includes identifying specific industries, transportation sectors, and other sources of emissions.

**iii. Hotspot Mapping:** The project can create pollution hotspot maps that highlight areas with consistently poor air quality. This information can guide targeted interventions and regulatory measures in regions where air quality is particularly problematic.

**iv. Health Impact Assessment:** Through epidemiological studies, the project can estimate the health impacts of air pollution in Tamil Nadu, such as the incidence of respiratory diseases, cardiovascular problems, and premature mortality. These insights can underscore the urgency of addressing air quality issues.

**v. Economic Costs:** By quantifying the economic costs associated with air pollution-related health problems, the project can provide a clear picture of the financial burden on healthcare systems and the overall economy.

**vi. Policy Recommendations:** The project can offer evidence-based policy recommendations and interventions to reduce air pollution. This might include stricter emission standards for industries and vehicles, promoting clean energy sources, and enhancing public transportation.

**vii. Monitoring and Evaluation Framework:** Establishing a framework for ongoing monitoring and evaluation allows for the tracking of progress in air quality improvement over time, ensuring that interventions are effective.

**viii. Environmental Impact:** Assessing the environmental impact of air pollution, such as its effects on ecosystems and natural resources, can also be within the project scope.

Overall, the project's scope encompasses a thorough examination of air quality in Tamil Nadu, from data collection and analysis to policy recommendations and community engagement. Its primary aim is to provide a holistic understanding of the problem and offer actionable insights to mitigate the adverse effects of air pollution on public health and the environment.

T

## SUMMARY:

The project's primary focus is on analyzing and visualizing air quality data collected from monitoring stations across Tamil Nadu, India, with the overarching goal of addressing the critical issue of air pollution in the region. Our key objectives encompass gaining deep insights into air pollution trends, pinpointing geographic areas suffering from consistently high pollution levels, and developing a predictive model to estimate RSPM/PM10 levels based on SO2 and NO2 concentrations. Applying design thinking principles, we intend to not only perform data analysis but also understand the root causes of pollution, involve stakeholders, and ensure that our findings are actionable. The analysis approach will involve rigorous data preprocessing, comprehensive spatial and temporal analysis, and the development of an accurate predictive model. In terms of visualization, we aim to create interactive and informative representations of air quality trends, pollution hotspots, and model outputs. Ultimately, this project seeks to empower policymakers, environmental agencies, and the public with the knowledge and tools needed to make informed decisions, leading to improved air quality and enhanced public health in Tamil Nadu.