**Project: Air Quality Assessment of TamilNadu**

**Empathize and Understand the Problem:**

Understanding the problem and the context. Why is analyzing air quality important in Tamil Nadu? What are the specific challenges and concerns regarding air pollution in the region? Gather insights from experts, stakeholders, and potential users of your analysis.

**Defining Clear Objectives:**

Objective 1: Analyze historical air quality data to identify trends and patterns.

Objective 2: Identify regions or monitoring stations with consistently high levels of air pollution.

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

**Ideation and Analysis Approach:**

**Data Collection:** Identify sources of air quality data in Tamil Nadu, such as government agencies or research institutions.

**Data Pre-processing**: Clean and pre-process the data, handling missing values, outliers, and data quality issues.

**Data Analysis**: Use statistical analysis and visualization techniques to identify trends and patterns in the data.

**Pollution Hotspot Detection:** Develop algorithms or criteria to identify areas with consistently high pollution levels.

**Predictive Modelling**: Choose an appropriate machine learning algorithm to build the predictive model for RSPM/PM10 levels.

**Evaluation**: Define metrics to evaluate the model's performance.

**Prototype and Visualization Selection:**

Matplotlib, Seaborn, Plotly, for visualization.

Time series line charts to show air quality trends over time.

Heatmaps or geographical maps to identify pollution hotspots.

Scatter plots or regression plots to visualize the relationship between SO2, NO2, and RSPM/PM10 levels.

**Build and Implement:**

Develop the full data analysis and visualization pipeline based on the refined approach.

**Test and Iterate**:

Continuously test analysis and visualization as progress, making adjustments and refinements based on feedback and new insights.

**Deliver Insights:**

Presenting the findings and insights in a clear and understandable manner. Use the selected visualizations to communicate trends, hotspots, and the predictive model's performance.