**PROJECT**:  **PUBLIC TRANSPORTATION ANALYSIS**

**Empathize and Understand the Problem**:

Understanding the importance of analyzing public transportation in Tamil Nadu is crucial for addressing the region's specific challenges and concerns. Engage with experts, stakeholders, and potential users to gather insights into the current state of public transportation, identifying pain points and areas for improvement.

**Defining Clear Objectives:**

**Objective 1:** Analyze historical public transportation data to identify usage patterns and service efficiency.

**Objective 2:** Identify key areas with high demand for public transportation and assess the current infrastructure's capacity.

**Objective 3:** Develop a predictive model to estimate peak hours and optimize service frequency.

**Ideation and Analysis Approach:**

**Data Collection:** Identify sources of public transportation data in Tamil Nadu, such as transportation agencies or municipal records.

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

**Data Analysis:** Use statistical analysis and visualization techniques to identify usage patterns and areas of improvement.

**Demand Hotspot Detection:** Develop algorithms or criteria to identify regions with high demand for public transportation.

**Predictive Modelling:** Choose a suitable machine learning algorithm to build a model predicting peak hours and optimizing service frequency.

**Evaluation:** Define metrics to evaluate the model's accuracy and effectiveness.

**Visualization Strategy:**

**Line Charts:** Display historical usage patterns over time to identify trends and fluctuations.

**Heatmaps or Geographic Maps:** Visualize demand hotspots geographically to highlight areas that require increased transportation capacity.

**Scatter Plots or Regression Plots:** Illustrate the relationship between various factors influencing service optimization, such as time of day and passenger count.

.

**Build and Implement:**

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

**Test and Iterate:**

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

**Deliver Insights:** Present the findings and insights in a clear and understandable manner. Utilize the selected visualizations to communicate usage patterns, demand hotspots, and the predictive model's effectiveness.

**Integration:** Integrate the data analysis, visualization, and predictive modeling code into a cohesive pipeline for seamless execution.