**Mini Project Overview: Traffic Prediction Model**

**Project Title:**

Traffic Prediction Model Using Historical Data

**Objectives:**

1. To develop a predictive model that estimates traffic volume for specific locations and times.
2. To analyze the impact of various factors (time, weather, events) on traffic patterns.
3. To provide insights that can assist in traffic management and urban planning.

**Dataset:**

* **Source:** Open traffic datasets (e.g., city traffic databases, Kaggle datasets, or traffic APIs).
* **Features:**
  + **Traffic Counts:** Historical traffic volume at specific intervals.
  + **Weather Data:** Temperature, precipitation, and conditions (sunny, rainy, etc.).
  + **Time Features:** Hour of the day, day of the week, and holiday indicators.
  + **Event Data:** Scheduled local events (e.g., concerts, sports events).

**Methodology:**

1. **Data Collection:**
   * Gather data from multiple sources, ensuring coverage of the target area over different timescales (e.g., months or years).
2. **Data Preprocessing:**
   * Clean the dataset by handling missing values and outliers.
   * Normalize the data to bring features to a comparable scale.
3. **Exploratory Data Analysis (EDA):**
   * Visualize data to identify patterns and correlations among variables using Matplotlib and Seaborn.
   * Perform correlation analysis to select relevant features.
4. **Model Selection:**
   * Start with simple models like Linear Regression.
   * Progress to more complex models like Random Forest or XGBoost if needed.
5. **Model Training and Evaluation:**
   * Split the dataset into training (80%) and testing (20%) sets.
   * Train the model and evaluate its performance using metrics like MAE, MSE, and R² score.
6. **Deployment:**
   * Create a simple web application using Flask or Streamlit to demonstrate real-time predictions based on user input.

**Expected Outcomes:**

* A functional traffic prediction model that provides accurate estimates of traffic volume.
* A user-friendly interface for stakeholders to input parameters and receive traffic forecasts.
* A report detailing findings, methodologies, and recommendations for traffic management strategies.

**Tools and Technologies:**

* **Programming Language:** Python
* **Libraries:** Pandas, NumPy, Scikit-learn, XGBoost, Matplotlib, Seaborn, Flask/Streamlit for deployment.
* **Development Environment:** Jupyter Notebook or any Python IDE.