**PHASE-2**

**Student Name** : R.vasanth

**Register Number :** 212923243051

**Institution :** St.Joseph college of Engineering

**Department :** Artificial Intelligence and Data Science

**Date of Submission:** 8-5-2025

**Github Repository link:** https://github.com/AndrewBenaiah/Andrew-Benaiah-NM.git

* **Project Statement**

Building on foundational insights from Phase 1, Phase 2 aims to deepen the analytical rigor of crime trend analysis by integrating advanced data processing, interactive visualizations, and predictive modeling. This phase emphasizes detailed pattern recognition in spatial and temporal dimensions using modern analytical tools to support proactive policing and urban planning.

* **Project Objective**
* Perform thorough exploratory and statistical analysis to extract actionable patterns.
* Apply geospatial clustering and temporal decomposition techniques.
* Build an intuitive project workflow integrating all stages of crime data analytics.
* Utilize advanced tools for preprocessing, modeling, and visualization to ensure scalability and usability.
* **Flowchart of the Project Workflow**

Data Acquisition

↓

Data Preprocessing

↓

Exploratory Data Analysis (EDA)

↓

┌───────────────┐

↓ ↓

Geographic Temporal

Analysis Analysis

↓ ↓

└─── Combined───┘

↓

Modeling & Prediction

↓

Visualization & Report

* **Data Description**
* Crime Incident Records: Type, date/time, location (latitude/longitude), outcome.
* Temporal Features: Timestamps, day of week, month, year, holidays.
* Geographic Features: Region, district, police jurisdiction.
* External Factors (optional): Weather, demographic, and socioeconomic data.
* **Data Preprocessing**
* Cleaning: Handling missing, inconsistent, and duplicate entries.
* Standardization: Formatting date-time fields, standardizing crime categories.
* Geocoding: Converting addresses to coordinates (if needed).
* Feature Engineering: Extracting time-based features (hour, day, season), aggregating location data into zones or grids.
* Data Integration: Merging with external datasets like weather or census data.
* **Exploratory Data Analysis**
* Univariate & Bivariate Analysis: Frequency of crime types, correlations.
* Temporal Trends: Time series plots to detect peak hours, days, or months.
* Geospatial Visuals: Heatmaps, choropleths to identify hotspots.
* Clustering: K-means or DBSCAN for detecting spatial crime clusters.
* Hypothesis Testing: To validate statistical significance of observed patterns.
* **Tools and Technologies Used**
* Languages: Python (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn)
* Geospatial Analysis: Folium, Geopandas, QGIS, PostGIS
* Visualization: Plotly, Tableau, Power BI
* Databases: PostgreSQL + PostGIS
* ML/Stats: Scikit-learn, Statsmodels
* **Team Members and Roles**
* **S. Andrew Benaiah -** Maintaine detail documentation of methodologies , finding , and updates.Prepare the final project report and any presentation material
* **M.Dhuvaaragesh –** Builds dashboards and interactive visual Reports to present insights clearly. Uses to tools like Tableau, Power BI , or custom python libraries.
* **Vasanth -** Performs exploratory data analysis to Uncover trends and patterns.