## 

## **Phase-2 Submission Template – Data Analytics**

**Student Name:** Saraswathi.D  
**Register Number:** 731123205046  
**Institution:** Government College Of Engineering,Erode  
**Department:** B.Tech Information Technology  
**Date of Submission:** 9/5/2025  
**GitHub Repository Link:** https://github.com/Saraswathi-D/understanding-crime-trends-through-geograpic-and-temporal-data-analysis.git

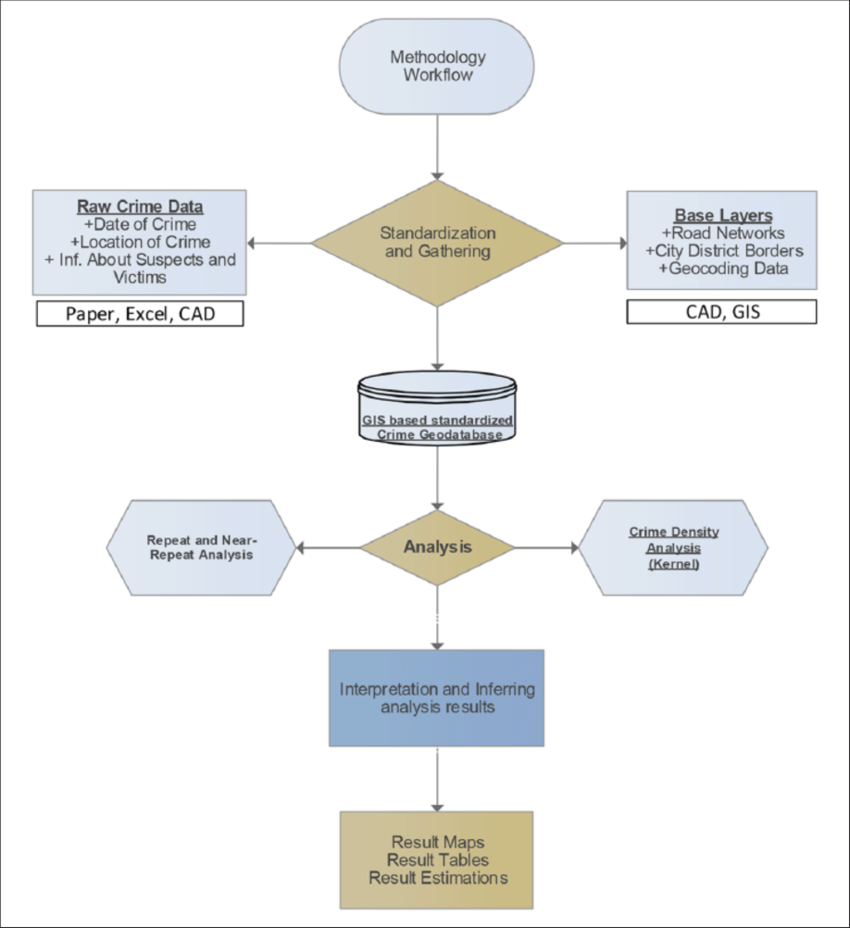
### **1. Problem Statement**

Crime patterns vary across locations and time, making it difficult for authorities to predict and prevent incidents effectively. By analyzing geographic and temporal data, we aim to uncover trends, identify hotspots, and support smarter, data-driven crime prevention strategies.

### **2. Project Objectives**

* Map crime hotspots across different regions.
* Analyze crime patterns over time (daily, monthly, seasonal).
* Identify factors influencing crime trends.
* Predict future crime occurrences using data analysis.Support better resource planning for crime prevention.

### **3. Flowchart of the Project Workflow**



### 

### 

### 

### **4. Data Description**

* **Dataset name and source**

(Kaggle-Crime\_data.csv)

* **Data type**

Structured

* **Number of rows and columns**

6639 rows , 12 columns

* **Static or Dynamic**

Static

* **Key fields and attributes relevant to the problem**

**Key Fields:**

Date or Datetime, Location , latitude and longitude , district or precinct.

**Attributes:**

Category, Disposition, Offender Status ,Victim Fatal Status.

### **5. Data Preprocessing**

### **Handled Missing Values:** Removed records missing key data like location or date; filled others with "Unknown".

### **Removed Duplicates:** Dropped exact duplicates to prevent overcounting.

### **Formatted Data:** Standardized timestamps and cleaned location data (lat/long).

### **Encoded Categories:** Converted crime types and weekdays into numerical codes for analysis.

### **Handled Outliers:** Identified anomalies using statistical methods and reviewed them contextually.

### **Documented Changes:** Logged all transformations for transparency and reproducibility.

### 

### **6. Exploratory Data Analysis (EDA)**

### **Univariate Analysis:** Identified theft, assault, and burglary as most frequent crimes; peak times were 8–11 PM, especially on weekends.

### **Bivariate/Multivariate Analysis:** Heatmaps and bar charts revealed strong correlations between crime type, time, and location; hotspots emerged in specific neighborhoods**.**

### **Key Metrics:** Average daily incidents, crime rates by region, and seasonal spikes (notably in summer and December) were analyzed.

### **Insights:** Crime patterns are both time- and location-dependent, with clear trends useful for prediction and prevention.

### **7. Tools and Technologies Used**

* **Programming Language:** Python
* **Notebook/IDE:** Google Colab
* **Libraries:** pandas, numpy, matplotlib, seaborn, plotly
* **Optional Automation Tools:**

### **8. Team Members and Contributions**

SIVARANJANI.S – Team Leader

SARASWATHI.D – Data Collection

SUVITHA.S – Machine Learning Engineering

RITHIKA.G – Evaluation Analyst