**(Crimes in India)**

**Title: Exploring Crime Trends in India: A Data Analysis Approach**

**Submitted by:**

**Disha Sharma**

Abstract:

This capstone project investigates crime trends in India through a comprehensive analysis of available crime datasets. The primary objective is to discern patterns, identify influential factors, and draw meaningful insights to contribute to the understanding of crime dynamics in the country.

Problem Statement:

India, as a diverse and populous nation, faces complex challenges in crime management. This project seeks to address gaps in our understanding of crime patterns, contributing to more effective policy formulation and law enforcement strategies.

Methods:

The analysis leverages advanced statistical and machine learning techniques to process and interpret extensive crime datasets. Exploratory data analysis is employed to identify temporal and spatial trends, while clustering modeling is used to assess the influence of various socio-economic factors on crime rates**.**

**Data Collection:**

In the initial phase of the project, data was collected from various reliable sources to create a comprehensive dataset for analysis. The following data sources were utilized:

Population Data:Source: Census of India

Link: Census of India - Population Data

Literacy Rate Data: Source: National Family Health Survey (NFHS)

Link: NFHS - Literacy Rate Data

Area Data:

**Assumptions:-**

* Assuming that we have yearly data requirements, we utilize projection data provided by the Census of India. Since population counts are available only every decade, we employ projection data for our analysis in this project.
* Assuming a criterion for literacy rates based on the National Family Health Survey, individuals aged 14 to 16 and above are considered literate. We possess population data for individuals aged 14 to 16 and above. According to historical reports, the literacy rate has consistently ranged between 70% and 80% of the total population every decade.
* The Census of India reported an average literacy rate of 73% in 2011, while the National Statistical Commission surveyed a literacy rate of 77.7% in 2017–18. Urban areas exhibited a higher literacy rate at 87.7%, compared to rural areas at 73.5%. Utilizing this data, we consider the literacy rate to be 70% of the total population for our analysis.
* **For visualization use the Power BI file**
* **Phase 1 project is done in Projection Dataset CSV file**

**References**:-

* [NCRB Official Website](http://ncrb.gov.in/)
* [Kaggle Datasets](https://www.kaggle.com/datasets)
* <https://en.wikipedia.org/wiki/Literacy_in_India>
* <https://scikit-learn.org/stable/modules/clustering.html>