Team - 1

Crime Data Analysis of a Los angeles of year 2023

Public safety remains a pressing concern in major metropolitan areas, and the City of Los Angeles has made strides in promoting transparency by releasing a comprehensive crime dataset spanning from 2020 to the present day. This dataset provides a wealth of information to analyze recent crime patterns and trends, enabling data-driven approaches to enhance community safety and security.

While crime rates can be influenced by many factors, including location characteristics, socioeconomic conditions, and temporal variables, this dataset offers an opportunity to uncover the specific attributes and factors driving crime occurrences in Los Angeles. By leveraging this valuable data resource, law enforcement agencies, policymakers, and community stakeholders can gain insights to develop targeted interventions and allocate resources more effectively.

The dataset encompasses a wide range of features, such as incident details, crime types, location information, and case statuses, providing a robust foundation for in-depth analysis. We will employ statistical analysis techniques to extract meaningful insights from this data, conduct exploratory data analysis (EDA), and utilize data visualization tools.

SMART Questions:

- How accurately can we predict the likelihood of crime being solved in Los Angeles based on the available data features in 2023?
- What are the key factors influencing crime rates across various neighborhoods or communities, and how have these factors evolved over the recent years?
- Can we identify emerging spatial and temporal patterns or hotspots for crime categories to inform proactive and targeted interventions?

Predicting the likelihood of crime types (Question 1) could be tackled using supervised learning models like decision trees, random forests, logistic regression.
Identifying key influencing factors (Question 2) could involve techniques like feature selection, dimensionality reduction, ensemble methods like gradient boosting and examining feature importance from tree-based models.
Detecting spatial and temporal patterns/hotspots (Question 3) may require specialized spatial analysis techniques, but clustering algorithms like k-means could potentially help identify hotspot areas.

DATASET SOURCE :https://catalog.data.gov/dataset/crime-data-from-2020-to-present

GITHUB REPOSITORY : https://github.com/varshith2233/DATS6103_Team1