CRIME DATA ANALYSIS

IE6400 –Foundations of Data Analytics Engineering

PROJECT 1- FINAL REPORT

Group Number 25

Hari Chandana Kannuru
Likitha Kukunarapu
Sri Likhitha Anuganti
Venkata Anantha Reddy Arikatla
Akilesh Reddy Katreddy

Introduction:

Crime data analysis plays a vital role in understanding and addressing public safety concerns in urban areas. It helps the law enforcement agencies track and respond to criminal activities. It can aid in resource allocation, patrolling, and deployment of officers in high-crime areas. This documentation presents an in-depth analysis of a real-world crime dataset spanning from 2020 to the present year (2023). The dataset provides a comprehensive record of criminal incidents, their types, locations, and associated attributes.

Throughout, the analysis will provide insights of overall crime trends, significant patterns in crime, most common type of crimes. In this project, we have used Python Pandas Library to clean the datasets. We then used the cleaned data for making Visualizations. We have designed graphs in an easy-to-understand manner above crime trends.

Data Sources:

This dataset reflects incidents of crime in the City of Los Angeles dating back to 2020. This data is transcribed from original crime reports.

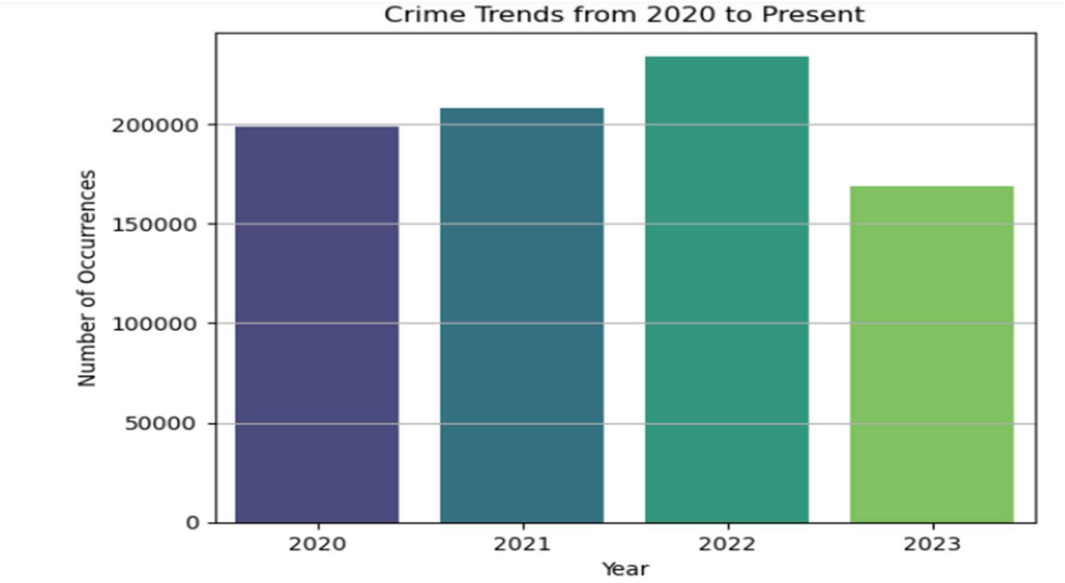
This dataset includes columns such as count of crimes, date of crime reported, time of the incident occurred, area name, area code, area district, description of the crime.

Results and Methods:

The data we have obtained contained large amounts of null values and redundant columns which needed removal to obtain the data for visualizations. For the initial EDA, we chose to use Python to remove null values, drop unnecessary duplicate rows, convert data types, standardize the numerical data, and encode the categorical data. The data was cleaned and processed completely using Python's Pandas library and then using matplotlib we presented it visually.

Exploratory Data Analysis:

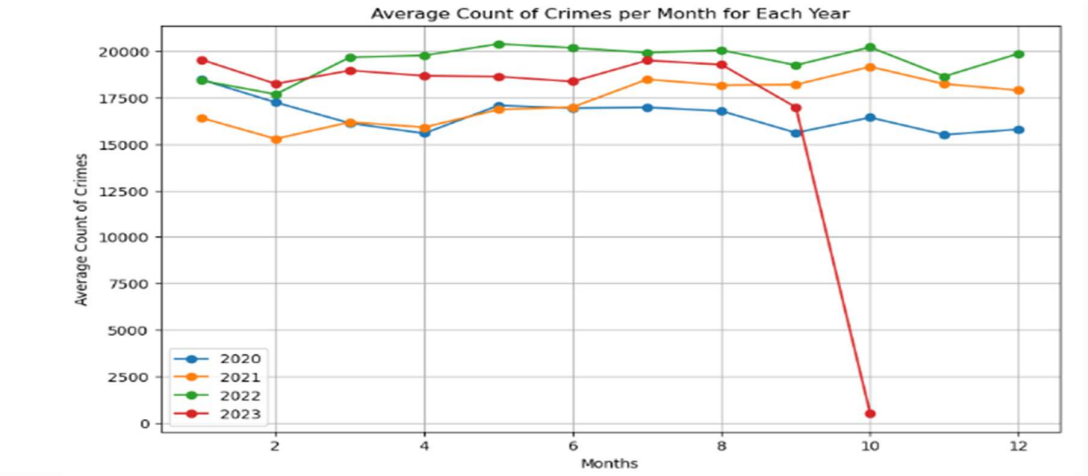
The below graph shows general trend of decreasing crime rates from 2022 to 2023. The highest number of reported crimes occurred in 2022, with 234,263 cases. Subsequently, in 2023, the number decreased to 182,161, marking a decrease of approximately 22.2%. This suggests that there may be some improvement in overall safety within the community over this period. Also, it's interesting to note that 2020 experienced a relatively high number of reported crimes, with 198,625 cases. This could be attributed to various factors, including the COVID-19 pandemic, which may have influenced crime rates as people's routines and socioeconomic conditions were impacted. The decreasing trend in crime rates in 2023 suggests that positive measures may have been taken to address certain issues.



Analysis of Crimes over the years:

The data reveals consistent seasonal patterns in crime rates. The crime counts tend to be higher during specific months, often in the spring and summer. While there are seasonal patterns, there are variations from year to year. Out of all the years , the year 2022 shows higher crime counts in the spring and early summer compared to previous years.

Also, the significant drop in crime counts in October 2023 suggests a notable change compared to previous years. Further analysis could help determine the cause of this decrease.



Most Common Crime Type:

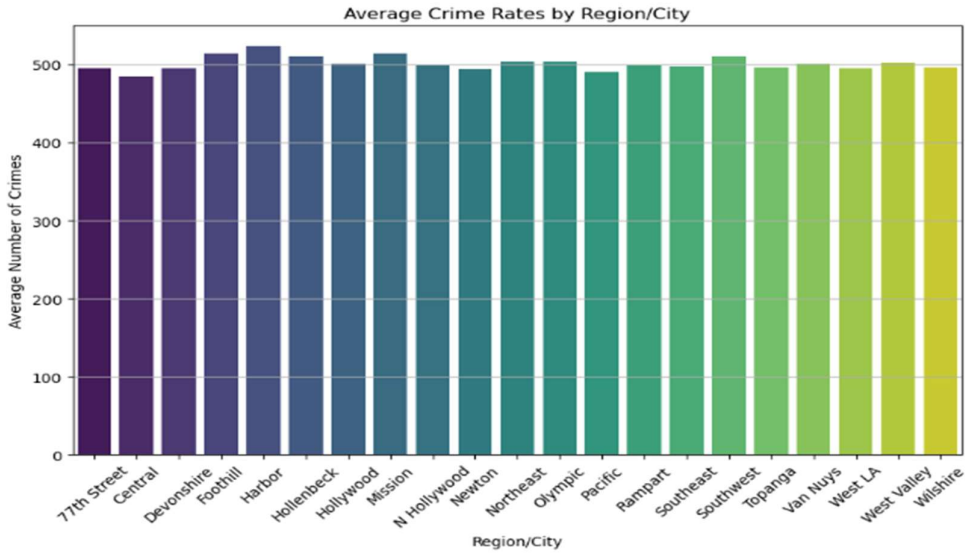
The fact that 'VEHICLE - STOLEN' is the most common crime with a high occurrence count of 86,750 indicates that auto theft is a significant issue. This could be a matter of concern for law enforcement and communities, as it poses a threat to property and public safety.

Implementation of strategies such as public education on safeguarding vehicles, improving surveillance in high-risk areas, and implementing technology solutions to prevent theft will help the situation.

Analysis of Crime rates by region:

The output highlights significant variations in average crime rates among different regions or cities. Regions like 'Harbor' and 'Mission' stand out with notably higher average crime rates, suggesting that they may require more resources and attention to address crime-related issues effectively.

On the other hand, regions like 'Central' and 'Pacific' have lower average crime rates, indicating a relatively safer environment.

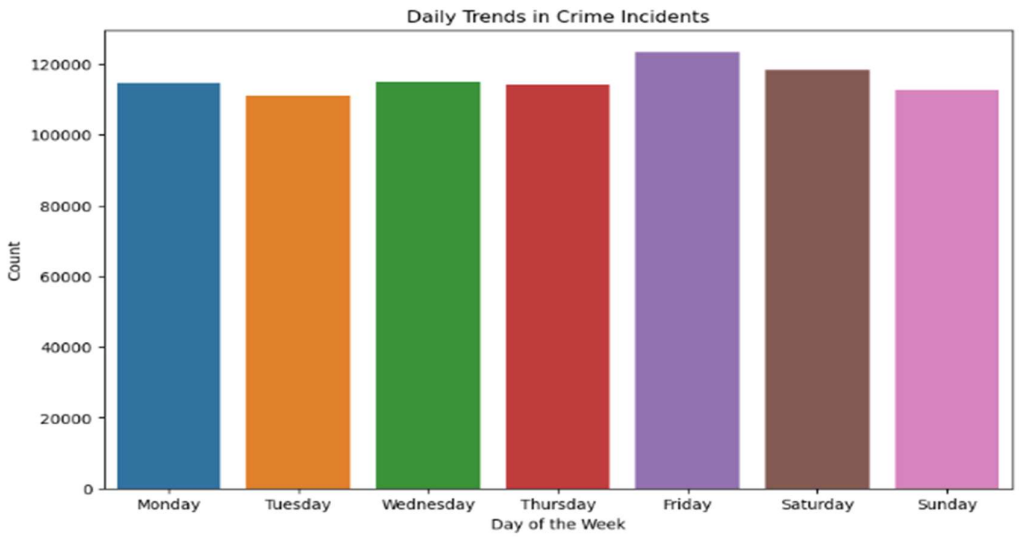


Analysis of number of crimes by day of the Week :

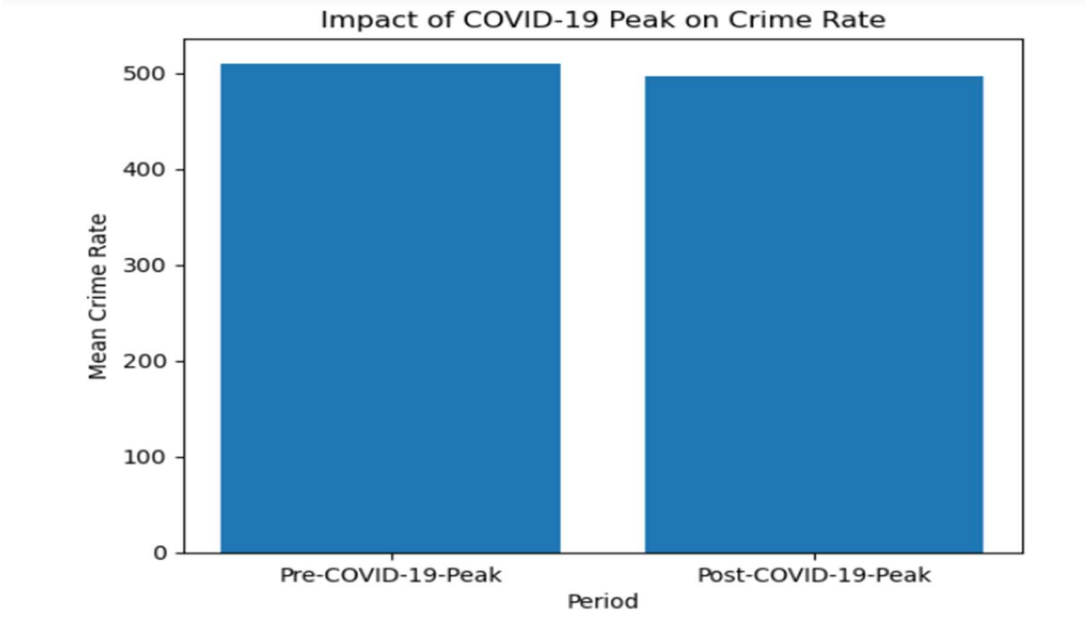
Overall, on the weekdays from Monday to Friday, generally have higher crime incident counts compared to weekends (Saturday and Sunday). This suggests that criminal activity tends to be more frequent during the workweek.

Friday stands out as the day with the highest count of crime incidents. This could be due to various factors, such as increased social activities and gatherings on Friday evenings, which may lead to higher incidents of certain types of crimes.

Law enforcement should consider the option of increasing patrols or focusing on crime prevention efforts on Fridays and other weekdays when crime rates are higher. Also public individuals should be more vigilant and take precautions.



The Impact of Major Events:

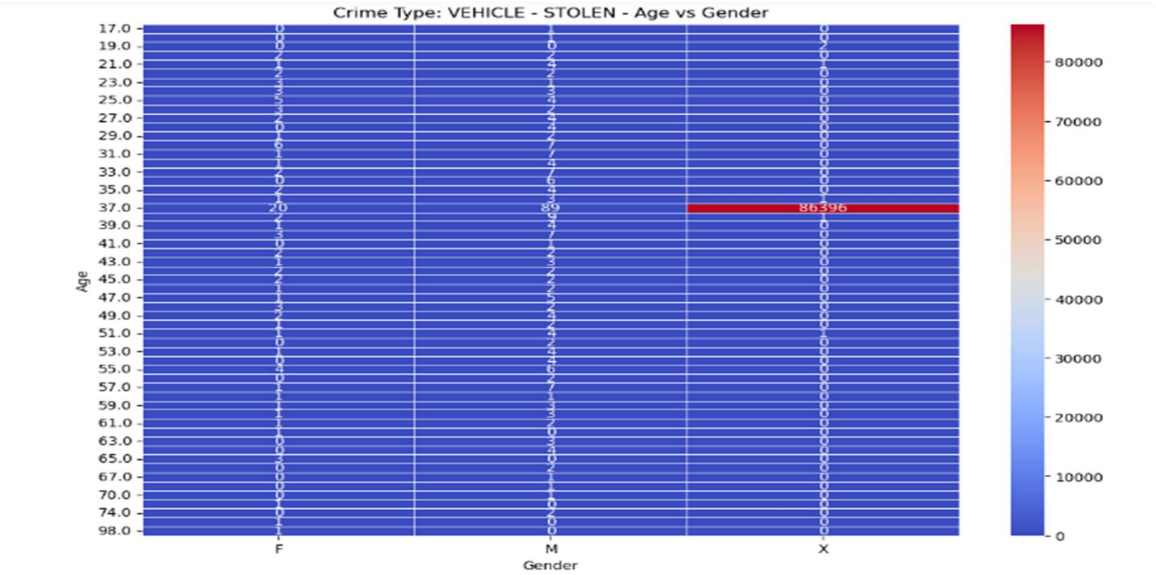


Before the COVID-19 peak, the mean crime rate was approximately 510.16. This represents the average number of reported crimes per unit of time before the pandemic's peak.

After the COVID-19 peak, the mean crime rate decreased to approximately 497.39. This indicates that, on average, fewer crimes were reported per unit of time after the peak of the pandemic.

The percentage change in crime rates from before the COVID-19 peak to after is calculated to be approximately -2.50%. This negative percentage indicates that there was a slight reduction in reported crimes after the COVID-19 peak compared to the period before it. This decrease might be due to various factors associated with the pandemic, such as lockdowns, social distancing measures, and reduced economic activities.

Analysis of Victim age and gender:



The table indicates that, for age 37 (and in some cases for 20s and 30s), there is a significant number of 'VEHICLE - STOLEN' crimes, particularly among males ('M'). This suggests that individuals in their late 30s, especially males, may be more vulnerable to vehicle theft.

Also, there are cases where the victim's gender is unspecified ('X'). It's important to note that the dataset may not have complete information for all cases.

Correlation Analysis:

	Vict Age	Count
Vict Age	1.000000	-0.019732
Count	-0.019732	1.000000

The correlation matrix shows that the value is approximately -0.019732, indicating a very weak negative correlation between victims age and the count of stolen vehicles crimes. This means that there is a very slight tendency for the count of stolen vehicle crimes to decrease slightly as the victim's age increases, but the correlation is very close to zero. In practical terms, age is not a strong predictor of the number of stolen vehicle crimes, as the correlation is weak and not statistically significant.

In summary, the correlation matrix shows that there is no significant linear relationship between the victim's age and the count of 'VEHICLE - STOLEN' crimes.

Final Insights:

Analyzing the city’s crime dataset from 2020 to 2023 has provided us with a comprehensive understanding of the city's crime landscape. The insights gained from this analysis have far-reaching implications for public safety, law enforcement strategies, and community well-

being. Over the examined years, we observed a general trend of decreasing crime rates. The significant drop in crime counts in 2023, compared to 2022, indicates that positive measures may have been taken to address various issues affecting public safety.

Moreover, the identification of stolen vehicles as the most common crime with a high occurrence count highlights the need for targeted actions in addressing auto theft.The variations in average crime rates among different regions underscore the importance of tailored law enforcement strategies. Some areas may require more resources and focused efforts, while others may serve as models for creating safer environments.

In conclusion, the insights from this crime data analysis emphasize the importance of data-driven decision-making for public safety and community well-being. Adapting to evolving crime patterns, addressing specific crime types, and focusing resources where they are needed most are vital steps toward a safer and more secure city. This analysis serves as a valuable tool for law enforcement agencies, city officials, and policymakers in their continuous efforts to protect and enhance the quality of life in the city.