Introduction

In recent years, the region of Los Angeles has observed a spike in the level of criminal activity, ranging in terms of severity from petty theft to assaults. First responders and local representatives have struggled to monitor and allocate resources in an efficient manner to limit the number of crimes consistently throughout time. To support first response resource allocation and provide tourists/citizens with a platform to promote safety, the analysis of crimes encompassing a spatial perspective is approached. After preprocessing and exploratory data analysis steps, a time-series visual trend was developed alongside a measure of relative strength respecting significant/anomalous change occurring.

Users and law enforcement are provided with six programs for trend and predictive analysis. Using ArcGIS Pro, a fishnet was developed in order to classify 80x80 pixelated cells on the basis of the most prominent crime type occurrence within the pixelated cell through the dataset's three year history (2020-2023). These platforms when analyzed in conjunction can identify high volume regions of crime, detect anomaly in crime rates, and promote preventative and mitigation efforts.

Data

Pre-Processing

The data was transferred and opened in a Jupyter environment and subsequently null values in the latitude and longitude columns were removed to ensure there exists a geolocation for all crime instances. Data types were also configured to appropriately represent fields as time, string, float, etc., and the final dataset used for analysis was derived from the original dataset through data manipulation functions, such as a dataset detailing 6 crime types. These datasets were leveraged in the various analysis steps of the workflow in order to guarantee consistency and accuracy with output results.

EDA + Data Clock + Hotspot

Exploratory data analysis was conducted on the dataset in order to view relationships between supporting features of crime, prominent regions of crime occurrence, and frequent victims of various crime types. This analysis provides elevated insight to vulnerable groups and regions of crime occurrence which can be responded to by mitigation/preventative measures. Demographic analysis conveyed that there is a disproportion of the population that incurs various crime types (battery, sex, child, misdemeanor, vandalism, homicide) - women are visualized in Figure 2 to be nearly three times as more likely to encounter a sex crime in Los Angeles while men are 1.5 times more likely to incur theft or assault. The Central and 77th Street areas of Los Angeles are observed to contain the highest crime rates, throughout the timespan of the dataset,

In recent months central Los Angeles has consistently recorded the highest level of crimes (Figure 3). Correlation analysis between crime type ('crime_cd') revealed strong linear association levels between the severity of crime and weapon used - .39 and .68 respectively.

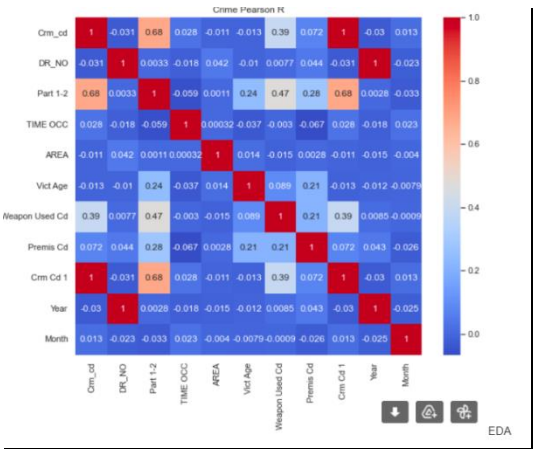

Figure 1: Correlation Matrix (Pearson R)

Finally a Trend analysis function was developed to serve as an indicator respecting the change of crime rates on a week-week basis from a user input LA region. Figure 4,

displays a user input of ‘Central’ passed to the function, a trend is visualized alongside 2 month and 5 month moving averages detailing weekly crime counts per the Central Region. The ROC indicator below measures the crime rate of change week over week, the red upper limit conveys when the rate of change falls above two standard deviations from the mean, detailing a significant change in trend. This can be leveraged in identifying times of interest where crime rates have deviated to an anomaly.

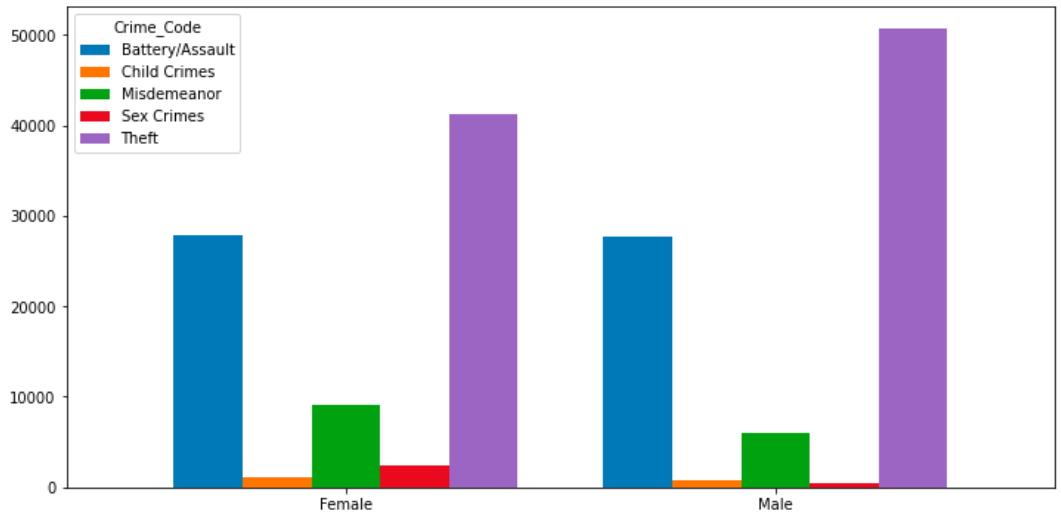

Figure 2: Bar chart representing distribution of victims by gender (Demographic Analysis)

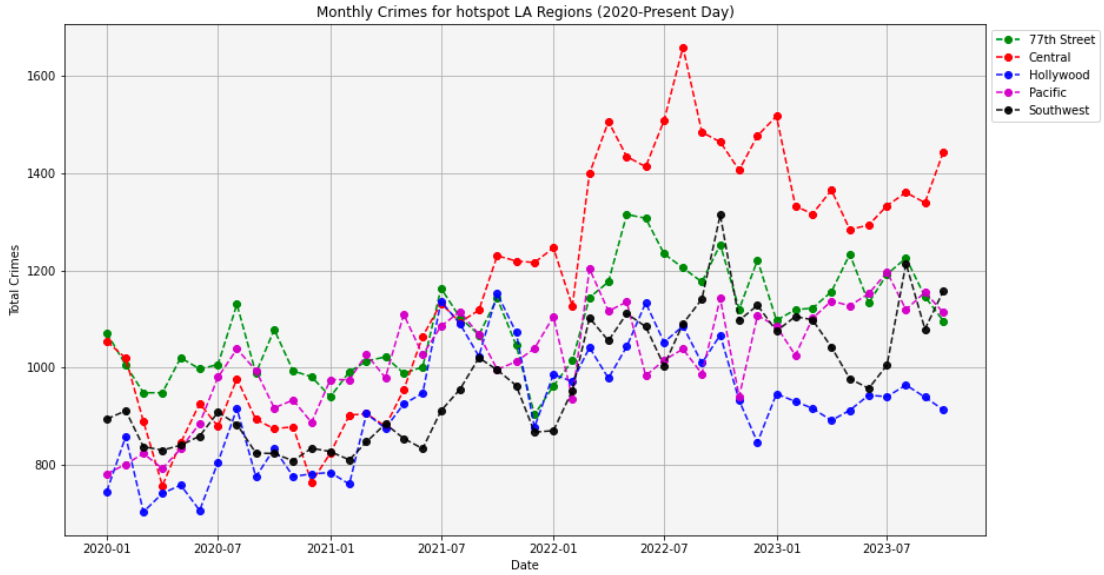

Figure 3: Monthly crimes per region (Regional Analysis)

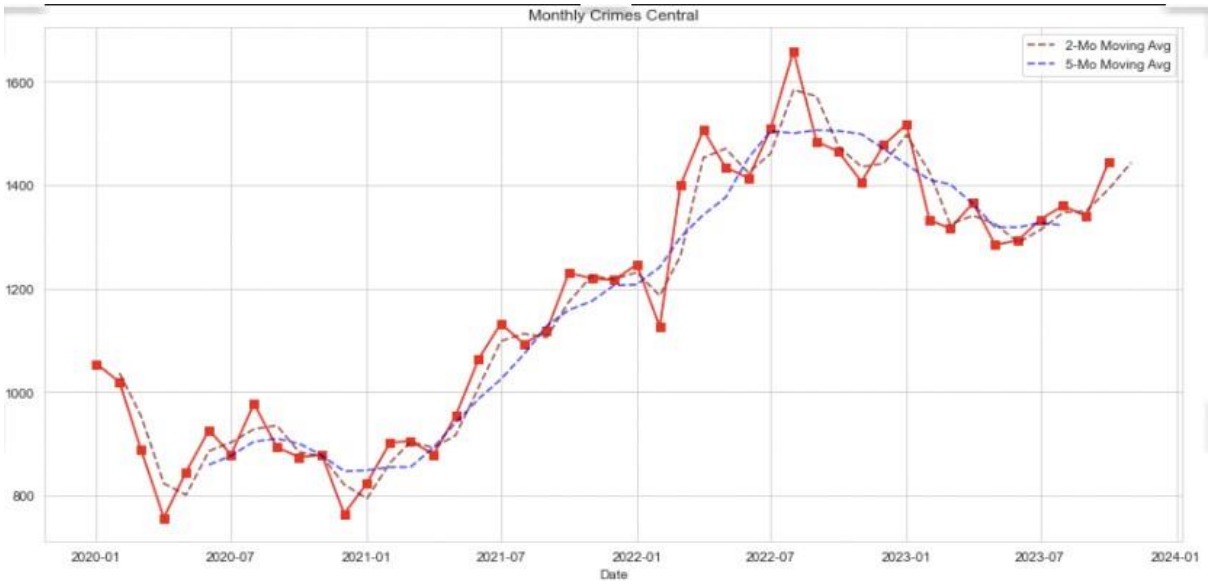

Figure 4: Anomaly detection and technical analysis of monthly crimes - Central region

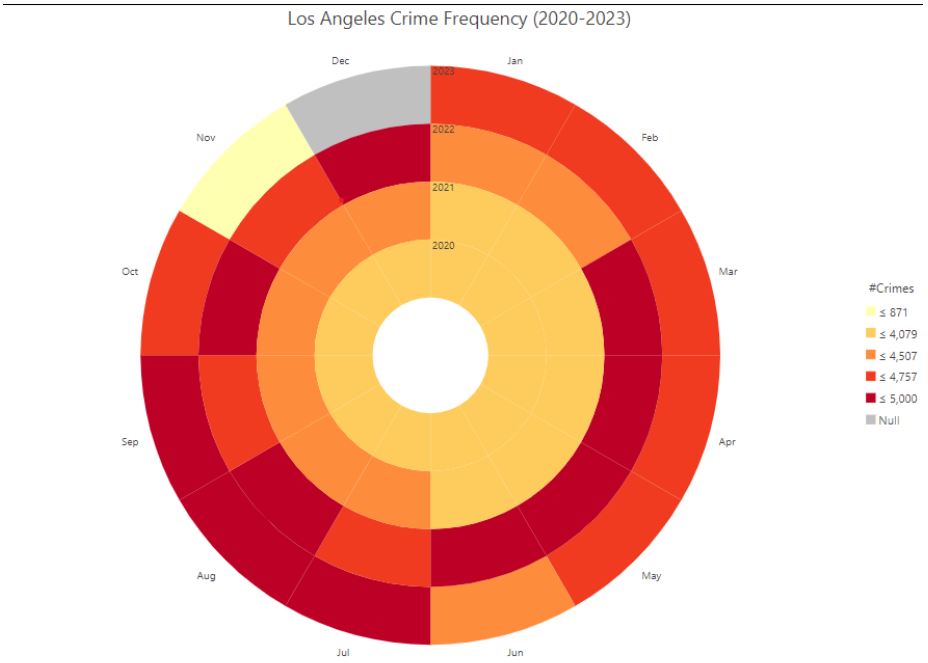

Figure 5: Data clock displaying crime frequency (2020-2023)

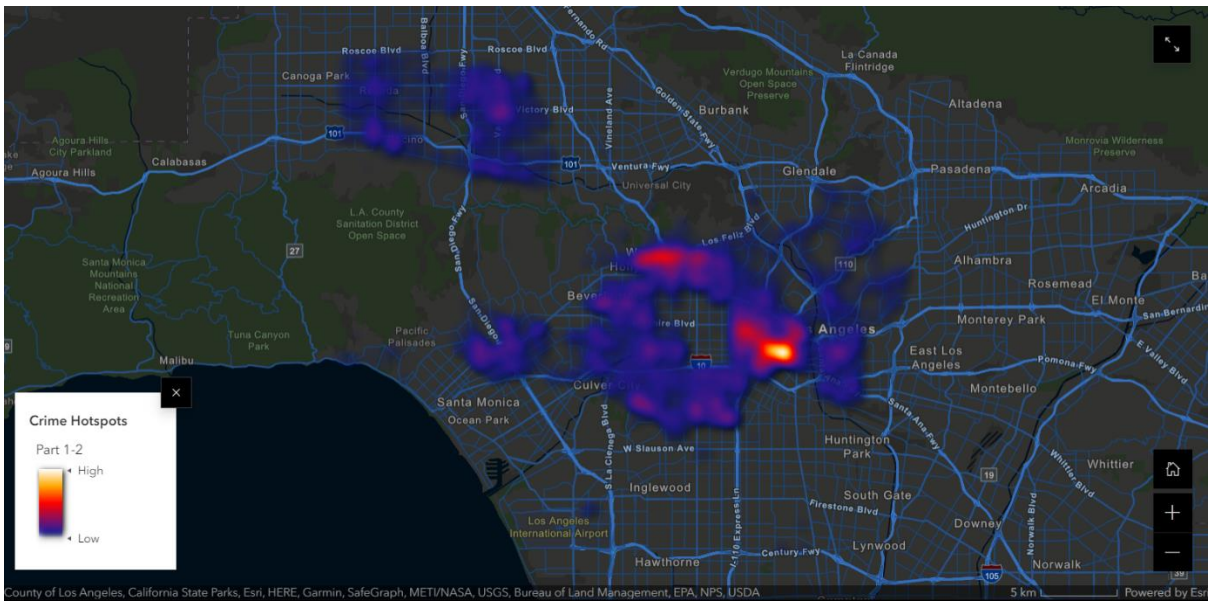

Figure 6: Hotspot analysis displaying severe crime hotspots

Fishnet Grid:

The innovative use of a fishnet grid, where each pixel uniquely represents a distinct category of crime, has proven to be a transformative approach in crime analysis. In this spatial representation, colors are employed to code various crime categories, providing a

visually intuitive and comprehensive overview of crime patterns. For instance, a blue pixel signifies instances of theft, the most prevalent category. This gridded fishnet allows for a granular examination of crime distribution across geographic regions, aiding law enforcement and policymakers in identifying hotspots, allocating resources efficiently and formulating targeted strategies for crime prevention. The grid's pixel-based categorization not only simplifies the visualization of complex crime data but also enhances the interpretation and communication of patterns, fostering a more informed and proactive approach to addressing urban security challenges.

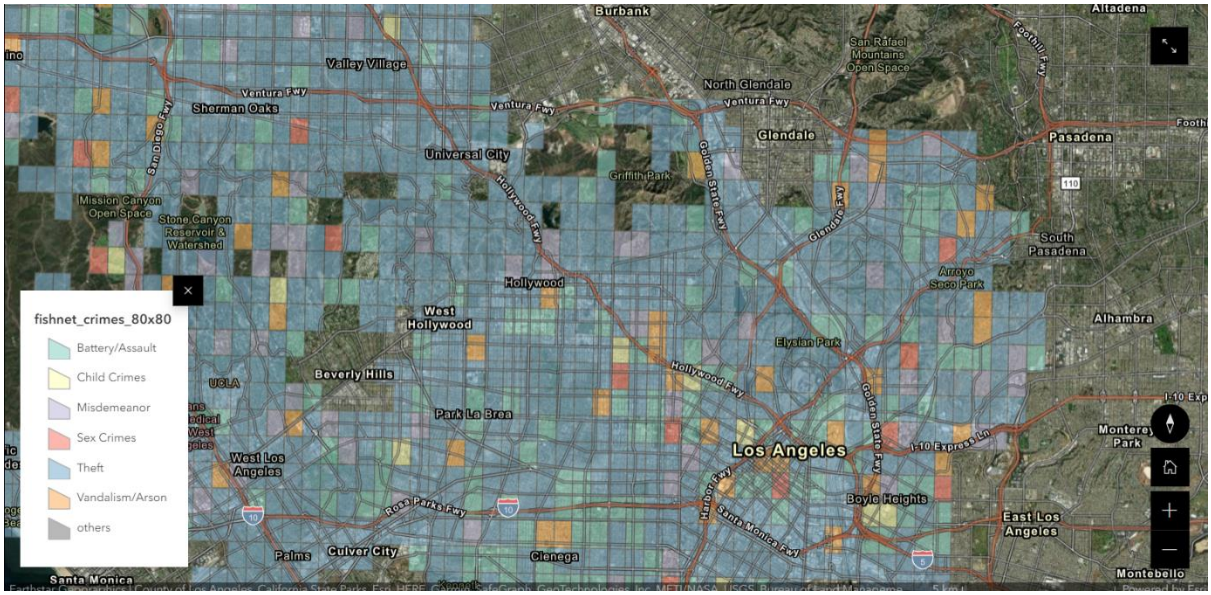

Figure 7: 80x80 pixel fishnet grid displaying color-coded crime types

Future Work

Given more time, potential directions to improve the project could include enhanced predictive modelling, utilizing more advanced machine learning models or ensemble methods to improve the accuracy of crime predictions. Additionally, real-time analysis can be leveraged to implement real-time data streaming and analysis for up-to-date insights and quicker response to changing crime patterns.

Conclusion

In conclusion, this approach to crime analysis has proven to be a powerful and insightful tool for understanding and addressing urban security challenges. The project successfully identified patterns and trends associated with criminal activities across different regions of Los Angeles. The project's outcomes underscore the potential of data-driven approaches in optimizing law enforcement strategies and resource allocation. As we navigate the complex landscape of urban safety, the insights gained from this machine learning based analysis contribute greatly to the ongoing efforts to create safer and more resilient communities.