Research Project Report:

Los Angeles Police Department Patrolling

Tsai-Hsuan Cheng, Stacy Lin, Elven Liu, Jahnavi Rati

APAN 5205 Frameworks and Methods II

Professor Kitty Kay Chan

Apr. 19, 2022

Problem Statement

People are experiencing difficult times during this ongoing COVID-19 pandemic. One of the pressing issues is the proliferation of criminal activities nationwide in regions and communities. There are countless reasons for the crime rate to increase, and the effects of COVID-19 are some of the reasons for propelling people to take unlawful actions.

Since the beginning of the pandemic, Los Angeles, one of the cities in California, has had a significant increase in crime rates making the neighborhoods less safe to stroll around. On February 8th, 2022, the Los Angeles (LA) mayoral candidate Republican Karen Bass proclaims that the Los Angeles Police Department (LAPD) needs to increase its patrolling in the greater LA area (City News Service).

Considering mayoral candidate Republican Karen Bass's call for action and knowing that the number of LAPD officers is shrinking month after month, the chief of LAPD reached our team to angle perspectives on how they can improve their patrolling throughout the city. Specific factors that are believed to strengthen patrolling include the time and location of most criminal activities, types of criminal activities that occur in different LA communities, and which groups are mostly victimized of a criminal activity.

On the other hand, the number of LAPD officers decreased month after month, according to the City Watch LA. Therefore, due to insufficient police officers, the crime rate has increased since 2021 in LA. As a result, we would like to analyze the crime rate in LA and provide actionable insights and suggestions for LAPD to protect the communities more efficiently.

Research Questions

How can the LAPD improve its patrolling efficiency? How can the LAPD improve its patrolling efficiency? When and where do most criminal activities occur, and what type of criminal activity is primarily engaged? Which demographics in Los Angeles, California are victimized the most?

Data Description and Structuring

We collected the data from the LAPD website, which contains all incidents in 2021. In the raw data set, there are over 681,183 data points and 28 variables. After cleaning the data set, we have 597,352 data points and 22 variables. Below is the data description for each variable.

- **Date OCC:** The exact date of the incident.
- **Time OCC:** The exact time of the incident.
- **Area:** The LAPD has 21 Community Police Stations referred to as Geographic Areas within the department. These areas are numbered from 1-21.

- **Area name:** The 21 Geographic Areas or Patrol Divisions are also given a name designation that references a landmark or the surrounding community that it is responsible for.
- **Rpt Dist No:** Represent sub-area within a Geographic Area. All crime records reference the "RD" that it occurred in for statistical comparisons.
- **Crm Cd 1&2:** Indicates the crime committed.
- **Crm Cd Desc:** Defines the Crime Code provided.
- **Vict Age:** The age of the victim.
- **Vict Sex:** The sex of the victim. (F-Female, M-Male, X-Unknown)
- **Vict Descent:** A - Other Asian, B - Black, C - Chinese, D - Cambodian, F - Filipino, G - Guamanian, H - Hispanic/Latin/Mexican, I - American Indian/Alaskan Native, J - Japanese, K - Korean, L - Laotian, O - Other, P - Pacific Islander, S - Samoan, U - Hawaiian, V - Vietnamese, W - White, X - Unknown, Z - Asian Indian.
- **Premis Cd:** The type of structure, vehicle, or location where the crime occurred.
- **Premis Desc:** Define the Premise code provided.
- **Weapon Used Cd:** The type of weapon used in the crime.
- **Weapon Desc:** Define the Weapon code provided.
- **Status Desc:** Define the Status provided. (Default = Invest Cont)
- **LOCATION:** The exact location of the crime.
- **LAT:** Latitude.
- **LON:** Longitude

In addition to the dataset with its listed variables above, we created multiple variables and added them to the dataset for further investigation. There are 135 types of criminal activities described in the **Crm Cd Desc** variable that also counts for the levels of criminal charges for engaging in an activity. To exclude the levels of charges and simplify the types of criminal activities, we categorized them into 21 different factors. To further simplify the variable, we created a new variable that helps distinguish between a violent crime and a financial crime.

- **crm.type:** 21 types of criminal activities (excluding levels of criminal charges mentioned in **Crm Cd Desc**)
- **crm.type.cat:** Identifies **crm.type** as a violent, financial, or other crime
- **year:** Year of the criminal activity (2020, 2021, or 2022)

Analytic Techniques

To address the research inquiry, we applied three different analytical methods of analysis: time-series, spatial, and clustering.

Time-series analysis helps to understand when most criminal activities happen in a year and provides suggestions for LAPD to be on high alert during the high frequent times. For the

analysis, we group the data into daily and monthly and analyze the trend of criminal activities. Moreover, we retrieve the top criminal activities among the most frequent days and months for the LAPD to understand the types of criminals at high frequent times quickly. For the prediction model, we use ARIMA and data from the past 2 years to train the model and predict the trend in the next 12 months.

Spatial analysis helps visualize the spread of different types of criminal activities occurring across LA and its 21 communities to identify and suggest strategies for improving patrolling. For the analysis, we use mapping tools that accurately locate each activity on LA's map using the longitudinal and latitudinal components. After running the analyses, we would recommend how the LAPD can improve its patrolling efficiency. If there are criminal activities intensely engaged in specific LA communities, then the LAPD should be on high alert and patrol more effectively in those communities. If the spread of criminal activities is equally spaced. In that case, the LAPD should maintain their patrolling methods across all communities and consider other factors: time and type of criminal activity to improve patrolling efficiency.

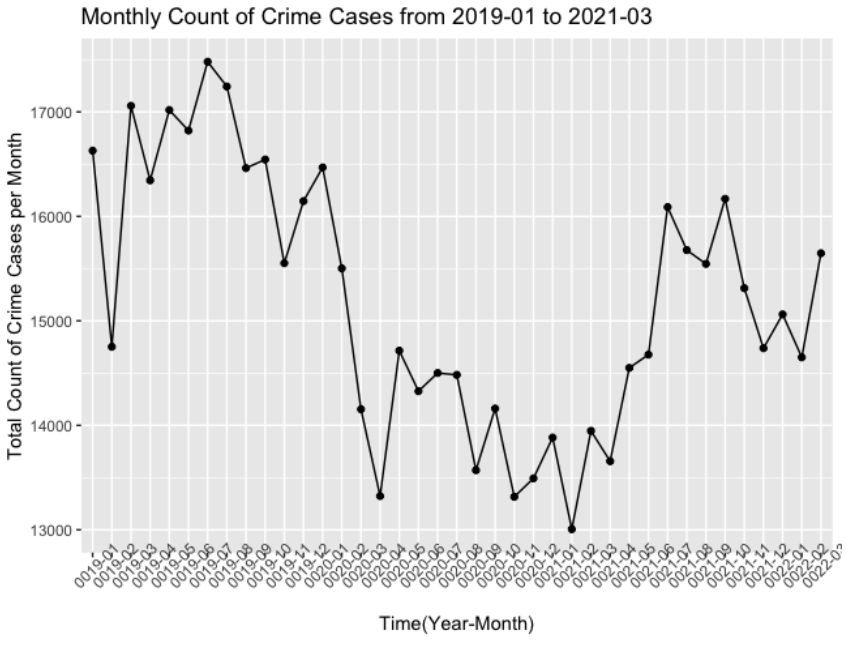
Cluster analysis helps classify the data to explore the similarity among the clusters. After finding the specific age, gender, and descent that are victimized the most, the LAPD can increase their awareness of those groups of people to patrol more efficiently. Moreover, the result can also raise awareness among the public.

Analysis and Results

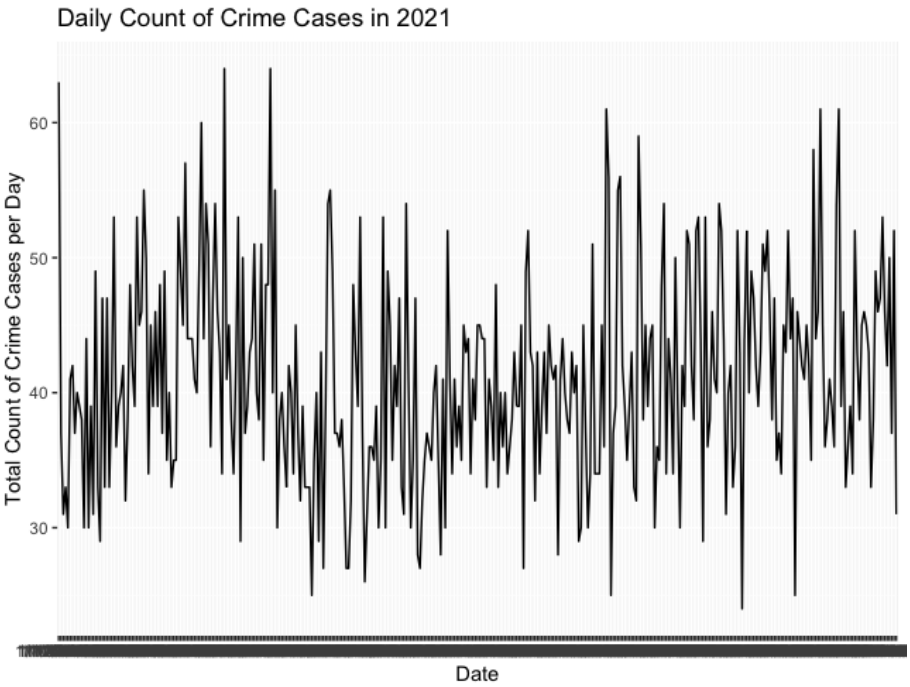
Time-Series Analysis

The time-series analysis will help address this research question: when do most criminal activities occur, and what type of criminal activity is mainly engaged?

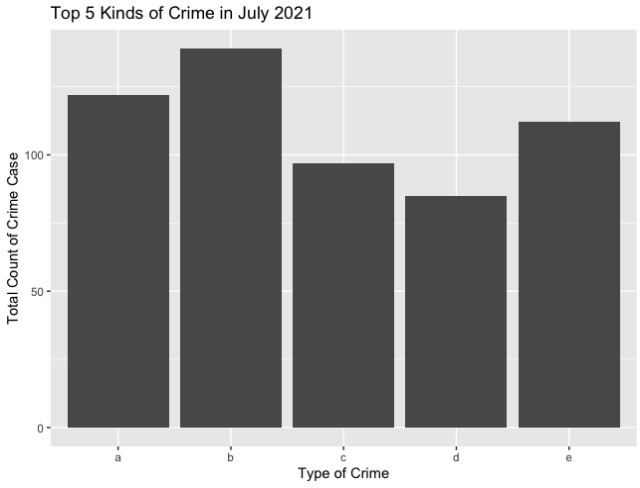
First, to explore the data trend, we plotted the line graph of the monthly count of crime cases, with time on the x-axis and the count on the y-axis.



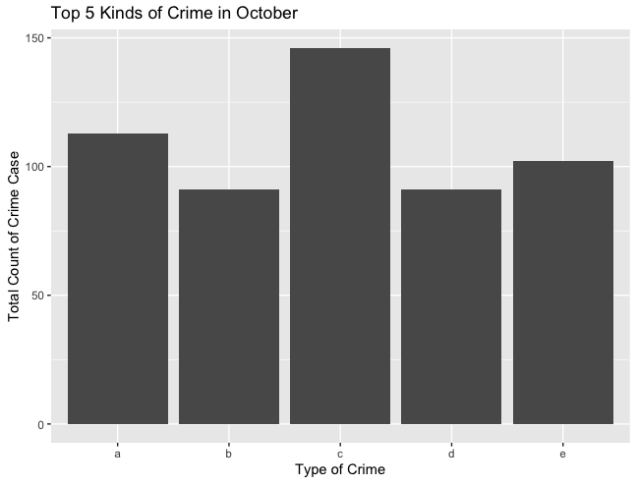
Then we knew that it was high in 2019 and then went down in 2020 and then went up again in 2021, reaching its two peaks in July and October 2021, with over 16000 cases. After knowing that the trend rose again last year, we wanted to focus on the data for 2021, so we plotted another line graph that showed the daily count of crime cases in 2021. The chart showed that it reached its highest points on 11/19/21 and 12/1/21, with 64 cases occurring, and its lowest point on 7/31/21, with 24 cases occurring.



After that, we attempted to explore the top 5 kinds of cases that occurred most frequently within all the cases in July and October 2021 because they were the two peaks in 2021, as mentioned above. Then we found out that they were assault, burglary, and vandalism. Then, when we looked into the data in October, the results included assault, burglary, vandalism, and theft, similar to that of July.

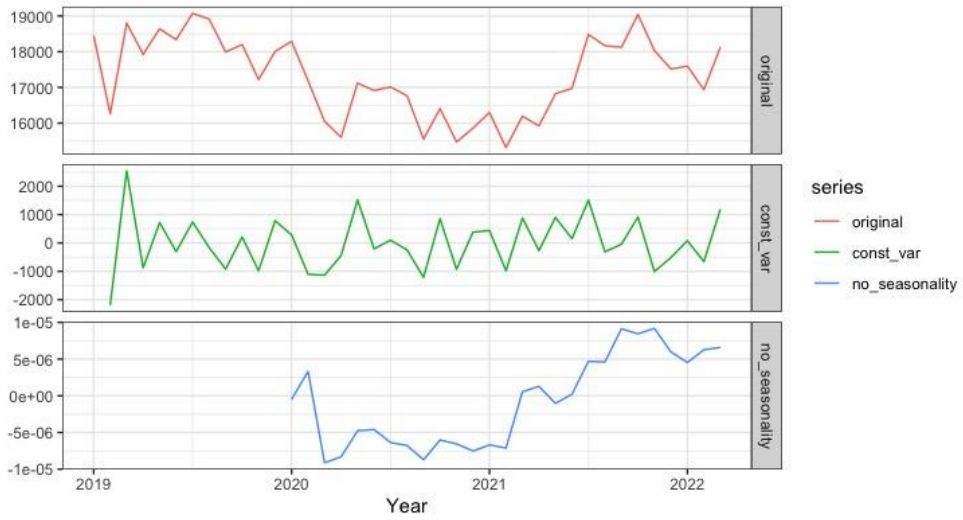


	Crime	Count
a	Assault With Deadly Weapon , Aggravated Assault	122
b	Battery - Simple Assault	139
c	Burglary	97
d	Burglary From Vehicle	85
e	Vandalism - Felony ($400 & Over, All Church Vandalisms)	112

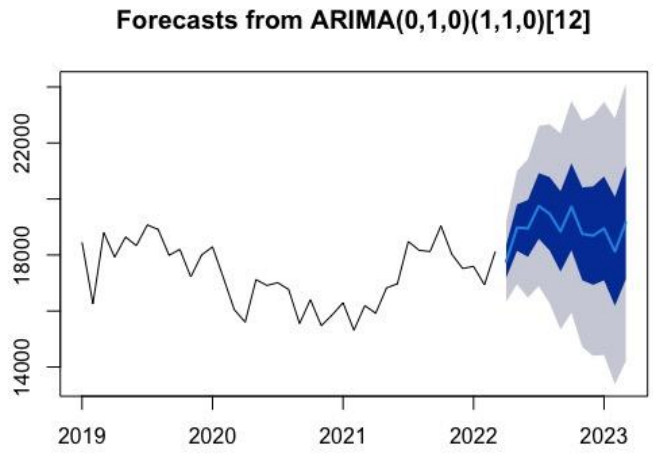


	Crime	Count
a	Battery - Simple Assault	113
b	Burglary	91
c	Burglary From Vehicle	146
d	Theft Plain - Petty ($950 & Under)	91
e	Vandalism - Felony ($400 & Over, All Church Vandalisms)	102

For the forecast, we chose to use the ARIMA model because this model can help us describe the autocorrelations in the data. At first, we grouped the data into monthly data, and we found that the data set was not stationary. Then we used box-cox transformation to remove variance and remove seasonality.



After the stationary process, we tried auto ARIMA and compared different combinations. Finally, we used AICc to figure out the most suitable AR and MA terms for us to use.

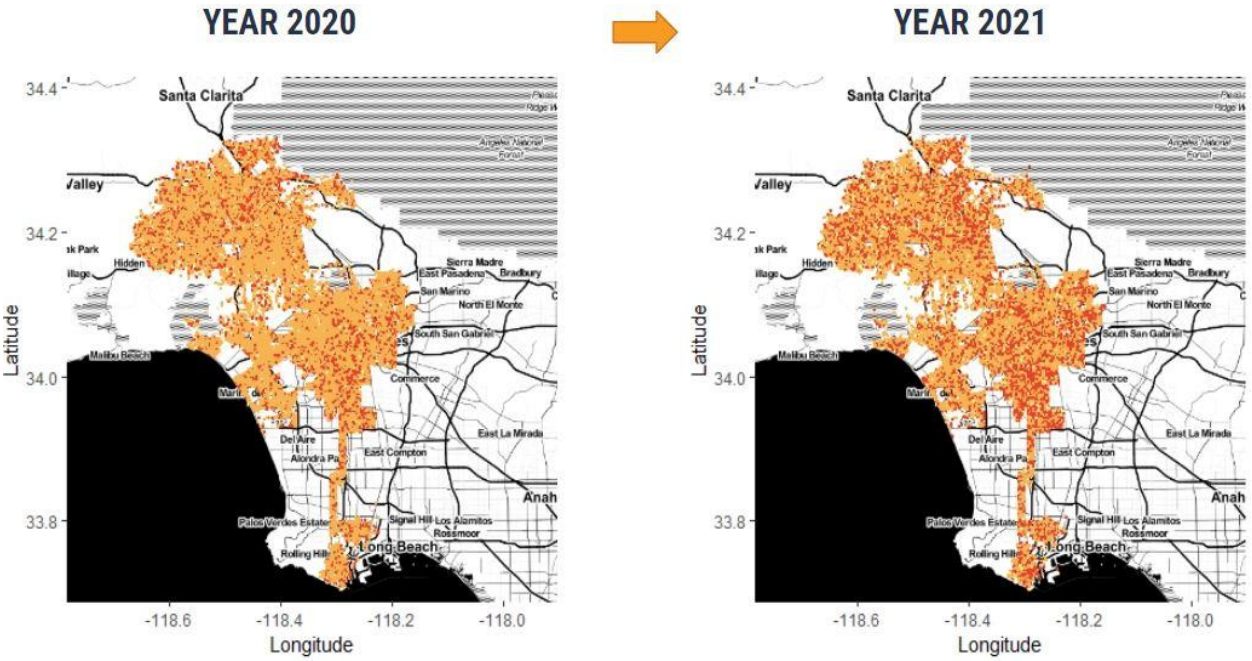


As shown in this graph, the blue line is the prediction line. Under the 80 percent of confidence level, the number will lie in the dark blue interval. And under 95 percent of confidence level, the number will lie in the light gray area. So there are 2 peaks in the prediction which are similar to the previous analysis.

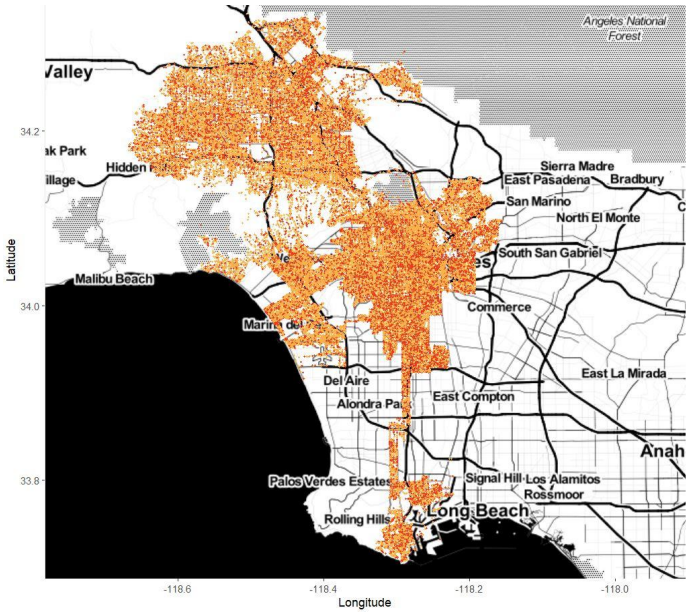
Spatial Analysis

The spatial analysis will help address this research question: where do most criminal activities occur, and what type of criminal activity is mainly engaged?

To support public officials' and our argument that the number of crimes has increased since the pandemic, the chart below illustrates LA's map with criminal activities from 2020 and 2021. Note that violent crimes are any form of physical attack or threat to people and animals, and financial crimes include forgery, property crime, and fraudulent activities. Based on the chart, there has been a significant increase, particularly in violent crimes, in 2021. So given this information, we need to investigate the violent and financial crime rates in each LA community.



There are 21 LA communities, of which seven communities are in the northern part of LA, and the rest thirteen communities are in the southern part of LA. On average, about 39% of the activities in an LA community constitute violence, and on average, about 56% of the activities in an LA community are financial crimes.



Therefore, we believe that based on all criminal activities since 2020 shown on the right, there is more violence in southern communities and more financial crimes in northern communities. To test our argument, we calculated the crime rates for each community and compared the rates by community

Our findings suggest that the top two communities with high violent crime rates are Southeast and Pacific, with a 50% rate. Southeast and 77th Street are southern communities, so we speculated further on the following three communities. The next three communities are Newton, Southwest, and Rampart, located in south LA. We also discovered that most communities are below the average 39% violent crime rate, supporting our claim that southern communities, especially Southeast and 77th Street, have higher violent crime rates.

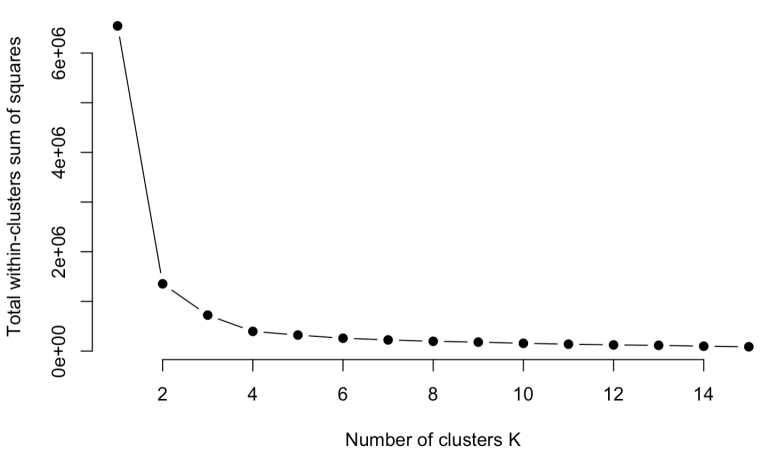
Additionally, our findings for financial crimes suggest that the top two communities, West LA and Pacific, have higher financial crime rates with about 67% and 66%, respectively. These communities are located more central and southern of LA, so we speculated further on the next three communities. The next three communities to our surprise are Devonshire, Topanga, and Northeast, located in northern LA. We also noticed that Southeast and 77th Street have the lowest financial crime rate of about 44%, and other southern communities' financial crime rates remain below the average of 56%.

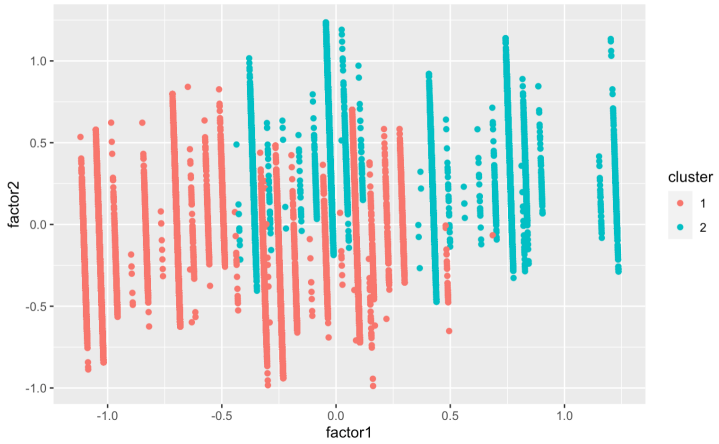
The largest violent crime is a simple assault and the largest financial crime is theft/auto repair, and over a period of time, theft/auto repair has shown to be the most significant portion of criminal activities occurring throughout LA and its communities.

Cluster Analysis

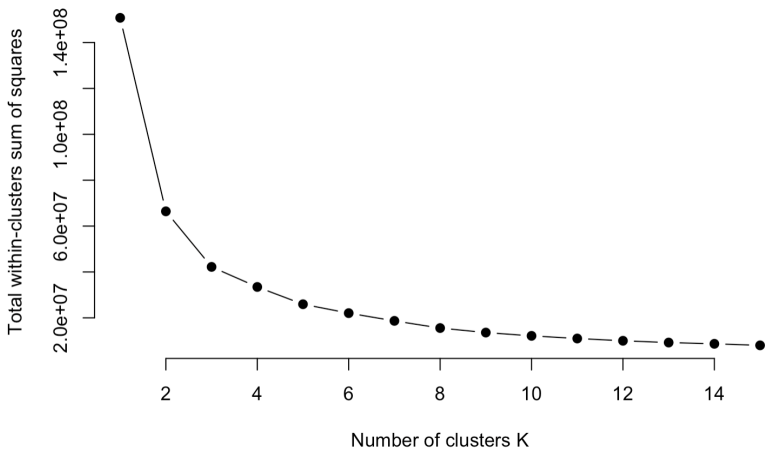
The time-series analysis will help address this research question: Which demographics in Los Angeles, California are victimized the most?

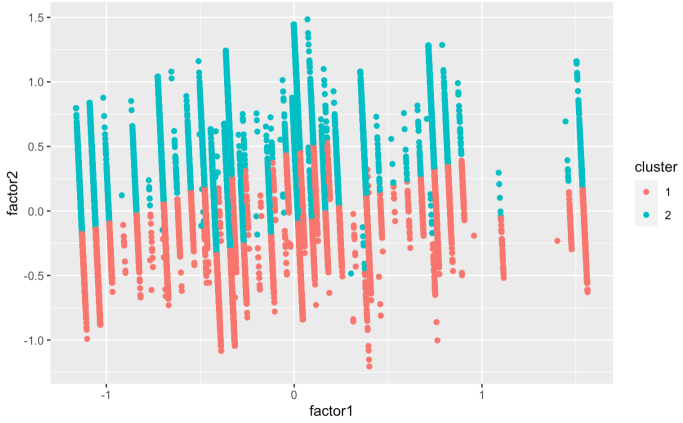
Among all the cluster methods, K-mean clustering produces a similar quality of clusters to hierarchical clustering but much faster; hence we choose K-mean and apply age, gender, and descent to the analysis. After computing the total within the sum of squares to decide the number of clusters, there was a significant elbow at two, and the graph on the right-hand side expresses the groups on a scatter plot. 98,693 observations belong to cluster one, whereas 77,348 observations belong to cluster 2. The demographic of cluster 1 is Hispanic and Black females aged around 38, and the demographic of cluster 2 is white males aged approximately 41.





In order to explore a more significant result from cluster analysis, we added the data of 2019 and did the same analysis. The elbow was not as significant as applying the 2020 data alone, but two was still the number that appeared a sudden change in the line graph, and the graph on the right-hand side expressed the clusters on a scatter plot. It moves closer than using 2020 data alone but is still not significant. We think the reason behind is we only have three variables regarding the victim's demographic which is not enough to generate significant results. 416,583 observations belonged to cluster one, whereas 180,769 observations belonged to cluster 2. The demographic of cluster 1 is mainly Hispanic and Black males and the average age around 31, and the demographic of cluster 2 is mainly White males aged around 57.





Conclusions & Recommendations

From the result of the time-series analysis, we suggest that LAPD should be more careful of assaults occurring in July and October. Moreover, they should have higher alerts on the top criminal activities, such as assault, battery, burglary, vandalism, and consider different solutions.

Based on our findings from the spatial analysis, we conclude that Southeast, 77th Street, and other nearby southern communities have the most violent criminal activities of simple assaults, and West LA and Pacific have the most financial criminal activities of theft/auto repair. Overall, the largest criminal activity rate is theft/auto repair in LA. Thus, we recommend LAPD to be on high alert for simple assaults in Southeast, 77th Street, and other nearby southern communities, and also patrol proactively for any theft/auto repair not just in West LA and Pacific, but in all 21 communities as it has the highest rate of occurrences among other types.

The cluster analysis results show that Hispanic and Black females aged around 38, White males aged about 41, Hispanic and Black males aged approximately 31, and White males aged around 57 are the groups that are victimized the most. Therefore, the LAPD should enhance their awareness and protection of these groups. Moreover, laws and regulations should be implemented to protect these groups. For future research and analysis, we think that LAPD can collect more information regarding the victims and criminals for the crime data so further analysis can be more accurate and significant.

References

- City News Service. (2022, February 8). *LA Mayoral Candidate Rep. Karen Bass Calls for Increase in LAPD Patrols.* Spectrum News 1. Retrieved February 18, 2022 from https://spectrumnews1.com/ca/la-west/politics/2022/02/08/la-mayoral-candidate-rep--karen-bass-calls-for-increase-in-lapd-patrol-levels
- City Watch LA. (2022, April 11). *LA Major Crime Goes Up as LAPD Officers Shrink* https://www.citywatchla.com/index.php/cw/los-angeles/24297-la-major-crime-goes-up-as-lapd-officers-shrink