# Trends & Analysis of Crime in Los Angeles

Insights from Open Data from 2010 to Present

Trent Rivers
Computer Science
University of Colorado Boulder
Boulder, Colorado, USA
trent.rivers@colorado.edu

**Abstract**
This paper undertakes an examination of publicly available datasets from the city of Los Angeles's OpenData initiative. Specifically, two datasets recording crime in the city were studied, recording crimes between the years of 2010 to 2019 and 2020 to Present respectively.

A couple of research questions guided the analysis of this paper, including determining long-term trends in crime rates, seasonality in crime, and any relationship between victim demography and the types of crime that befall them.

According to our analysis here, Los Angeles is facing an increasing general trend in criminality. Crime peaked in 2022, but a change in LAPD's reporting system has obfuscated more precise trend analysis beyond 2023. A significant seasonal trend can be observed in crime rates, with crime falling in the Winter to a minimum in February, and increasing across Spring and Summer to a peak in the months between July and October.

K-Means clustering was employed to determine any demographic groups among the citizens of Los Angeles who are likely to fall victim to specific crimes. No connection could be found between the demographics of victims and the type of crime that befalls them. However, a handful of groups were found to generally be more frequently victimized, including young Hispanic females 17-25, middle-aged Hispanic males 32-50, senior white males 56-70, and elderly white females.

**Introduction**
Criminality is a persistent societal ill, costing thousands of lives and an unquantifiable amount of money every year. Preserving the safety and property of the citizenry is one of the primary roles of government, so addressing crime is perpetually relevant at every level of society, from locally to internationally.

Open Data initiatives have begun to emerge in the last couple of decades to provide public access to government data with the hope that this would foster greater transparency and research that would benefit the community. One city to champion this endeavor has been Los Angeles. As one of the United States's largest and most diverse metropolitan areas, much can be learned through analyzing the large datasets which the city has made public. Two of those datasets will be the subject of this paper: the first recording Crime Data from 2010 to 2019, and the second recording Crime Data from 2020 to Present.

There are a few research questions which guide the analyses conducted in this paper.  be allocated to address the problem. One question to answer is this: are there certain demographic profiles which are more likely to be victimized by certain types of crimes? This question has important ramifications for community education and prevention because understanding who is victimized most would determine who is targeted most.

We will also evaluate the role of time and seasonality: 1) “What is the overall trend for crime in Los Angeles over time?” and 2) “Does the season impact the level of crime on an annual basis?” Understanding larger trends across time is useful because it shows how criminality is changing, and how current resources may need to be reallocated and current methods adjusted to meet those trends.

**Related Work**

Crime is a problem that every government throughout the world continually seeks to alleviate. In 2018, the ubiquity of crime and the necessity of addressing it were captured well by X. Alphonse Inbaraj and A. Seshagiri Rao, who wrote “any research that can help in solving crimes faster will pay for itself” [1]. As a result, there is a substantial body of research which has been published surrounding crime generally and, more relevant to this project, analyzing crime data specifically.

A 2023 paper published in the International Journal of Geo-Information entitled “Systematic Review of Multi-Scale Spatio-Temporal Crime Prediction Methods” recounts the many methods in which researchers have attempted to create prediction models for crime [2]. The most common model architectures are neural networks, ensembled random forests, and random forests. Prediction is one of the most common data mining tasks, and this research presents another avenue for this project to be extended for additional insights to be drawn from the Los Angeles dataset using prediction.

As previously stated, the usefulness of solutions that meaningfully address the problem of crime is self-evident. With ever more research being conducted in the field of data-driven policing, it has become necessary to formally review the research and gauge its effectiveness in impacting crime. Last year, the paper “The Effectiveness of Big Data-Driven Predictive Policing: Systematic Review” was published in the Justice Evaluation Journal. In it, the authors surveyed 161 articles on big data-driven predictive policing [3]. They found that while the range of research is encouraging, there is still more research that needs to be conducted in this area because only 6 articles of the 161 could be said to provide strong evidence of effectiveness. This is encouraging for this project because it demonstrates that the established body of research is incomplete. With more research and novel approaches to data mining informing policing, then perhaps effectiveness could increase.

**Dataset**

The total dataset for this examination will consist of two individual datasets, downloaded as CSV files from data.lacity.org, and integrated together. The two datasets to be used are: 1) Crime Data from 2010 to 2019 and 2) Crime Data from 2020 to Present. Both can be accessed at the links below.

Data from 2010 to 2019 consists of 2.13M data objects with 28 attributes, and data from 2020 to Present consists of 1M data objects with the same 28 attributes. A full description of all attributes can be found in the Appendix.

Each sample within the dataset represents a “criminal incident”. This is different from a single crime because, during one incident, multiple crimes may be recorded (e.g. driving under the influence of alcohol, resisting arrest, and battery of police could all occur in one record). Each sample has a unique DR_NO, a department of records number,

Data Links:
Crime Data from 2010 to 2019:
https://data.lacity.org/Public-Safety/Crime-Data-from-2010-to-2019/63jg-8b9z/about_data
Crime Data from 2020 to Present:

https://data.lacity.org/Public-Safety/Crime-Data-from-2020-to-Present/2nrs-mtv8/about_data

**Tools**
This analysis was conducted using Python. Several libraries for Python were also used to support this endeavor: Pandas for importing CSV files and managing tabular data, NumPy for efficient handling of vector and matrix math, Matplotlib and Seaborn for visualizations, and Scikit-learn for utilities like One-Hot Encoding and Clustering.

**Techniques Applied**

**Data Cleaning**
Though the dataset comes from the official Los Angeles Police Department OpenData initiative, it is by no means high enough quality to do without data cleaning. The dataset is provided with some disclaimers about its data: "This data is transcribed from original crime reports that are typed on paper and therefore there may be some inaccuracies within the data. Some location fields with missing data are noted as (0°, 0°)". To address these and any other discrepancies, inaccuracies, and aberrations present in the original dataset, the data will now be cleaned. In the interest of reproducibility, the full process has been described here. The total number of samples contained in the original dataset is 3,138,128.

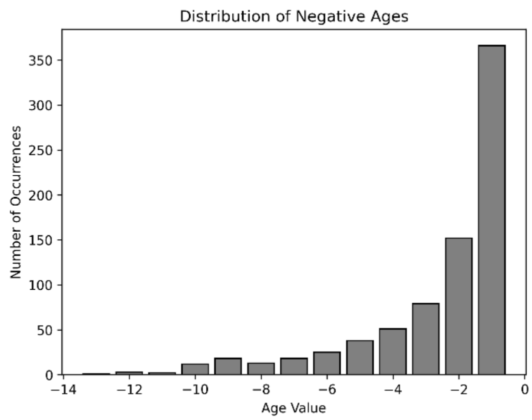
Unexpectedly, when attempting to integrate the two individual datasets which shared all attributes into a single set, a 29th attribute was mysteriously spawned. This was the result of a difference between the two files: one had an attribute named "AREA" while the second had an attribute named "AREA " – the trailing whitespace was causing this error. Upon naming both consistently, the sets could be joined without any additional error.

Upon investigation, 57,809 duplicate records were found to exist within the initial dataset. These duplicates were found using the Division of Records number ("DR_NO") attribute, where each incident should have a unique identifier. It is valid to ask whether the dataset included the same Records number multiple times to represent multiple crimes being committed in a single incident (e.g. assault and resisting arrest by one perpetrator). However, each sample in the dataset contains attributes representing multiple crime codes to accommodate for such an occurrence. Inspecting a handful of these records indicated that they were, in fact, multiples of identical records. Therefore, these duplicate records were removed from the dataset.

There were many problems that needed to be addressed with the Location data included in the initial dataset. As previously described, it has already been reported by the owners of the dataset that some location data is missing. When inspecting the data, no null, whitespace, or NaN values were found; however, 106 samples were found with locations of "UNKNOWN", "UNK", and "00". Due to the small number of records this applied to, they were all removed from the dataset.

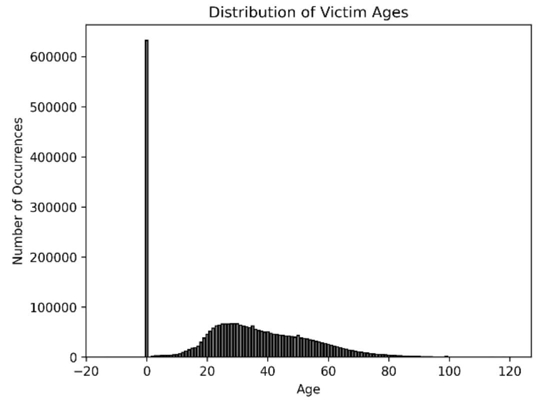
Since these records are derived from handwritten notes on each incident, many of the samples have Location data that's been shortened and does not reflect the exact address of the crime. For example, "6TH" is included instead of "6TH ST" for many samples. Though this does affect the street level information of many samples within the dataset, attempting to use regular expressions to string match all of these and correct them is time consuming with too little benefit to warrant pursuing. Other features like the area name, latitude, longitude, and reporting district would be more useful in examining the geographic impacts of crime. Therefore, this represents a known issue in the dataset which will remain unfixed here.

As stated in the Introduction section above, one of the questions to examine in this paper involves the demography of victims represented in the dataset. The Vict_age attribute, representing the age of the victim, is pertinent to this study. 778 age records were found to have negative values ranging from
-13 to -1 (seen in the distribution below).

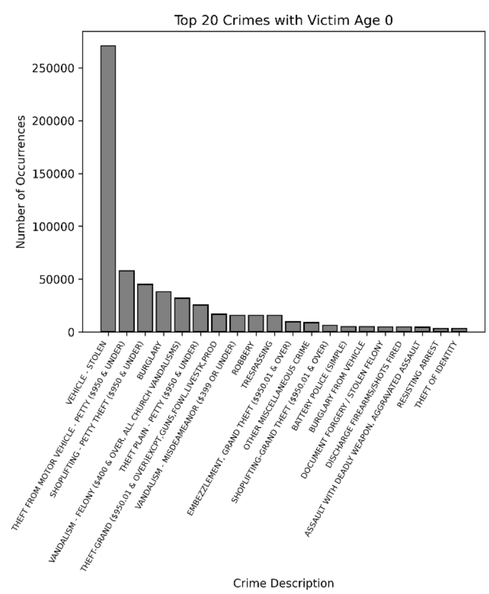



While one might assume these samples represent crimes against minors that errantly had a negative added upon data entry, inspection of an assortment of records revealed this to be incorrect (for example, DR_No 69460 – a Grand Theft, Embezzlement record with a victim of age -1). Due to how few records this represents and that there is no clear value to use to impute them, they have been removed from the set.

A larger problem with the Vict_Age attribute is that a total of 632,108 samples, roughly 20% of all samples, are recorded as having a victim of age 0. This is by far the highest and most common value in all the data.




An instinctual explanation one might propose here is that a victim age of 0 perhaps represents a "victimless crime" (crimes with no individual actively harmed) like violations of liquor laws, receiving stolen property, forgery, or fraud. However, the bar plot seen below disproves that hypothesis, with offenses including robbery, battery of police, and aggravated assault appearing in the top 20 crimes with age data of zero.




As seen in the Distribution of Victim Ages, there is a wide, left-skewed, unimodal distribution that evenly tapers off on both sides. This distribution *could* be randomly sampled and used to impute the missing ages with values that fit the distribution. However, that approach will not be taken here as imputing the ages according to this distribution would predictably distort any subsequent examination of the demographic data.

Another subset of ages with victim age of 0, 337,111 samples were found to have no victim age, no victim sex, and no victim descent data. As discussed above, imputing this demographic data would distort results, so another course of action must be followed. In the interest of not discarding 11% of all data (for samples with no demographic data) or upwards of 20% (for samples with age 0), these data will be preserved. *However,* when questions of demography are examined, these samples will be excluded from consideration, and a subset of the total dataset only including samples with complete age, sex, and descent data will be included.

Another attribute that similarly has many missing or erroneous values is the victim's sex. In the official documentation on the dataset, the letter X is meant to represent a case when the victim's sex is unknown. Along with some mis-transcribed letters, these unknown values represent 152,176 samples among the set. The decision that was made regarding age will also be applied in this case: the samples will be withheld during examination of demographic questions but will be preserved for other analyses like time-series analysis and geographic relationships.

3,133 records were found to have either longitude or latitude values of 0. This represents a small segment of the overall dataset; however, unlike previous data, where there is no simple method to impute values, latitude and longitude are more easily replaced. With 21 distinct values across the dataset, the Area Name attribute can be used to compute an average latitude and longitude to impute the missing records. Since no samples are missing an Area Name, that approach was used here.

**Preprocessing & Data Transformation**

The two features most relevant to time series analysis, DATE OCC and Date Rptd, are recorded as simple strings in the dataset. To make them more useful, they were first converted to Pandas datetime objects. This is a non-destructive process and only parses the string information to a different format. Using the newly formatted DATE OCC attribute, two new features were added: YEAR OCC and MONTH OCC.

As previously mentioned, a subset of the data needs to be used when considering demography due to the presence of erroneous values. In Pandas, a second DataFrame excluding these erroneous values has been created, and will be used as the subset for analyzing demographic questions. Excluding the samples with missing demographic information brings the subset to 2,430,525 samples.

Since the demographic data will be used with clustering using K-Means, and K-Means considers every feature to determine distance between samples, some further preprocessing is necessary. A subset of features will be selected and transformed to make the data usable for clustering. These include all the victim demographic features (Vict Age, Vict Sex, and Vict Descent), the crime code features (Crm Cd 1, Crm Cd 2, Crm Cd 3, and Crm Cd 4), and the Part 1-2 feature.

The specific research question examining demography and victimization we are analyzing in this paper is about different groups being subject to specific crimes. The initial dataset records each sample as a separate *criminal incident* (not at the granularity of individual crimes), with up to 4 crime codes recorded per sample. For each crime code present on a sample initially, the demographic subset has been altered so that each crime code represents its own sample. This brings the number of samples to 2,620,397.

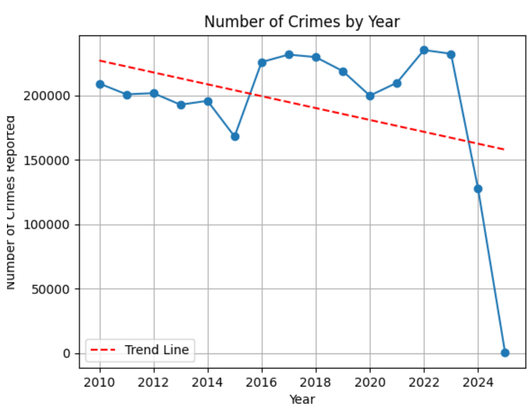
Only the Vict Age and Part 1-2 features are numeric in the original dataset. To further make this subset usable with clustering, one-hot-encoding has been used to make categorical features into Boolean feature columns. This results in a sparse matrix with 195 features.

Another smaller subset will be withheld for use with a Naïve Bayes classifier using the Vict Age, Vict Sex, Vict Descent, and Part 1-2 features. This classifier will use the Part 1-2 feature as the target since this would indicate if the victim experienced a serious crime. The age and descent were one-hot encoded.

## Analysis

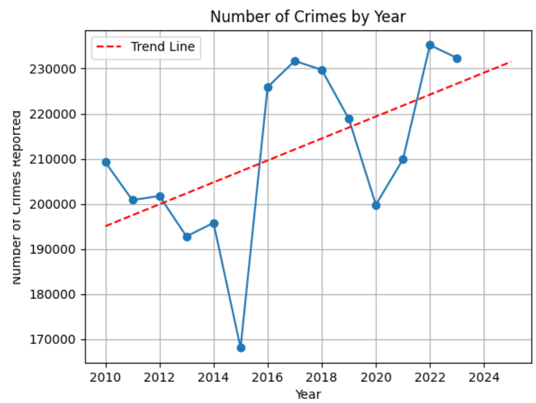
### Time-Series Analysis

To begin our analysis, let's first survey some of the overall trends in the data. The plot below shows the trend across the whole span of years under examination.



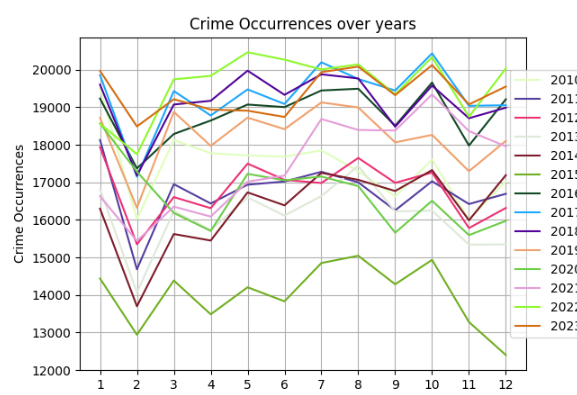

We can see that crime was in a slow upward trend from 2010 to 2022, where it peaked, before dropping precipitously in 2024, and finally plummeting close to 0 in 2025. While this would be encouraging news under normal circumstances, information on the official website for the data may provide an explanation. "LAPD will adopt a new Records Management System for reporting crimes and arrests… during this transition, users will temporarily see only incidents reported in the retiring system." [4]

Due to this technology change, a more accurate trend might be established by evaluating the data through the year 2023, the last complete year recorded in the dataset.




The trendline seen in the above plot likely better reflects the reality of crime in Los Angeles because the incomplete data from other years was excluded. This change makes the jump from a local minima in 2020 to a peak in 2022 more pronounced, with only a slight decrease in 2023. It also underscores the importance of the work undertaken in this study to better understand factors surrounding crime in the city.

We can analyze the time data at a more granular level, however. To determine any seasonal trends, we can use line plots with the number of crimes per month over each year:




A few key insights can be drawn from the graph above. The first is that crime typically reaches its lowest level in the month of February each year after falling off over the Winter months. Next, we can observe that the crime rate climbs across the Spring and Summer months before reaching its peak from August to October. This seasonality does have repercussions we will discuss later in the Applications section.

**Clustering**

As previously mentioned, clustering using the K-Means algorithm will be conducted on the demographic subset of the data to determine if specific groups are victimized more frequently for particular crimes. Since this data has no true labels, an unsupervised evaluation criterion must be used to evaluate how well the clusters are separated. Here, the Davies-Bouldin Index was selected as the evaluation criteria (a description of this metric can be found in the appendix).

Scikit-learn's GridSearchCV utility was used to find the best number of clusters by Davies-Bouldin Index. 5-fold cross-validation with k clusters in the range k = [10, 15, 20, 25, 30, 40, 50] was conducted. You can see the results of the grid search for the best number of clusters below (for Davies-Bouldin Index, lower is better):


The best performing clustering used 10 clusters, and had a Davies-Bouldin Index of 0.7087. In the table below, you can see the demographic component of the clusterings:

| CLUSTER | cluster size | most frequent gender | Gender Frequency | most frequent descent | Descent Frequency | Median Age |
|---|---|---|---|---|---|---|
| 0 | 326422 | M | 168878 | H | 138659 | 38.0 |
| 1 | 528896 | F | 285472 | H | 242589 | 25.0 |
| 2 | 195189 | M | 109269 | W | 64865 | 56.0 |
| 3 | 45597 | F | 24518 | W | 22730 | 82.0 |
| 4 | 262066 | F | 149850 | H | 157729 | 17.0 |
| 5 | 249471 | M | 137421 | H | 88129 | 50.0 |
| 6 | 468362 | M | 235613 | H | 194194 | 32.0 |
| 7 | 98157 | M | 52521 | W | 41997 | 70.0 |
| 8 | 150257 | M | 84599 | W | 55260 | 63.0 |
| 9 | 278455 | M | 146928 | H | 114786 | 44.0 |

There are discrepancies between the most represented demographic in each and the total size of the cluster. However, for this analysis, the clusterings will be defined by the majority demographic present within each. Clustering depends greatly on the dataset, and here it can be observed that the clusterings are not wholly complete or separated. The groupings are:

(0) 38 y/o Hispanic Males
(1) 25 y/o Hispanic Females
(2) 56 y/o White Males
(3) 82 y/o White Females
(4) 17 y/o Hispanic Females
(5) 50 y/o Hispanic Males
(6) 32 y/o Hispanic Males
(7) 70 y/o White Males
(8) 63 y/o White Males
(9) 44 y/o Hispanic Males

Looking at these clusterings intuitively, we can them by similar demographic "macro-clusters": young Hispanic females 17-25 (clusters 1 & 4), middle-aged Hispanic males 32-50 (clusters 0, 5, 6, & 9), senior white males 56-70 (clusters 2, 7, & 8), and elderly white females (cluster 3).

Surprisingly, there does not appear to be any significant connection between the demographics of the victim and the crimes they suffer (see Top Five Crime Codes for each K-Means Clusters in the Appendix). Many of the top 10 crimes for each cluster are shared by a majority of the other clusters:

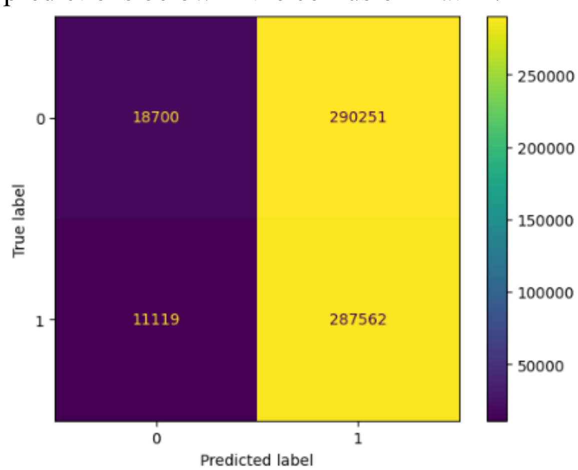
| CRM_CD | Cluster Count |
|---|---|
| BURGLARY FROM VEHICLE | 10 |
| THEFT PLAIN - PETTY ($950 & UNDER) | 10 |
| BATTERY - SIMPLE ASSAULT | 10 |
| Code Undefined | 10 |
| BURGLARY | 9 |
| VANDALISM - FELONY ($400 & OVER, ALL CHURCH VANDALISMS) | 9 |
| THEFT OF IDENTITY | 9 |
| ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT | 8 |
| VANDALISM - MISDEAMEANOR ($399 OR UNDER) | 7 |
| INTIMATE PARTNER - SIMPLE ASSAULT | 6 |
| THEFT-GRAND ($950.01 & OVER)EXCPT,GUNS,FOWL,LIVESTK,PROD | 5 |
| ROBBERY | 2 |
| BUNCO, GRAND THEFT | 1 |
| CRM AGNST CHLD (13 OR UNDER) (14-15 & SUSP 10 YRS OLDER) | 1 |
| CHILD ABUSE (PHYSICAL) - SIMPLE ASSAULT | 1 |
| BATTERY WITH SEXUAL CONTACT | 1 |
| THEFT FROM MOTOR VEHICLE - PETTY ($950 & UNDER) | 1 |

This would suggest that instead of being subject to specific types of crime, there are different groups who are more likely to be victims of crime generally.

**Classification**

Here, a Naïve Bayes classifier will be used as a confirmatory measure for the above clustering analysis. The goal is to predict who might be the victims of Part 1 crimes – more severe offenses according to FBI groupings, including homicide, rape, robbery, and burglary.

The demographic information was the same as what was used above for clustering, including victim age and one-hot-encoded gender and descent attributes. A training and testing split was created, with 75% of samples in the training set and the other 25% in the test set. You can view the results of the trained classifier's predictions below in the confusion matrix.




Using accuracy, precision, recall, and f1 score to evaluate this classifier, the following scores were achieved:

- Accuracy Score: 0.5040254627801037
- Precision Score: 0.4976731226192557
- Recall Score: 0.9627729919211466
- F1 Score: 0.6561642178953878

With an accuracy and precision of almost exactly 0.5, it provides no benefit in a binary classification. This result would appear to confirm the results of the clustering analysis above. The demographic information present here has no utility in predicting if a certain demographic will be the victim of a particular type of crime.

**Key Results**

When we consider all of the complete years represented in the dataset and exclude the incomplete years of 2024 and 2025 due to the LAPD's system change, we can see that the crime rate in Los Angeles is increasing.

Evaluating the data between the years of 2010 and 2023, annualized trends can be send in the crime data, where criminality drops across the Winter to a minimum in February before escalating across the Spring and Summer to a peak between August and October.

According to the k-means clusters developed in this analysis, there are a few "macro-cluster" groups that, while not the targets of specific crimes, are the most likely to be victimized generally. These include young Hispanic females 17-25, middle-aged Hispanic males 32-50, senior white males 56-70, and elderly white females.

The result that there is no connection between victim demographics and particular crimes was confirmed with a Naïve Bayes classifier, which found no predictive utility in the demographic information in the dataset.

**Applications**

This research provides some valuable insights to policymakers and civil servants within the city of Los Angeles. Firstly, the slow upward trend that continued through 2023 means that public resources ought to be allocated for a greater police presence within the city. Burglary, theft, battery, robbery, and vehicle thefts are common, and the city needs to address the trend.

The seasonality of crime rates has some implications for policy. New officers should be on-boarded during the Winter months each year, ahead of the increases seen in the Spring and the peak in the late Summer.

Additionally, greater investments in community education and more focused police presence could help to protect the groups that generally fall victim to criminality in the city.

**References**

[1] X. Alphonse Inbaraj and A. Seshagiri Rao. 2018. Hybrid Clustering Algorithms for Crime Pattern Analysis. In *International Conference on Current Trends towards Converging Technologies (ICCTCT). IEEE, New York, NY, USA.* https://ieeexplore.ieee.org/document/8551120

[2] Yingjie Du and Ning Ding. 2023. A Systematic Review of Multi-Scale Spatio-Temporal Crime Prediction Methods. In *International Journal of Geo-Information. MDP,. Basel, Switzerland.* https://www.mdpi.com/2220-9964/12/6/209

[3] Youngsub Lee, Ben Bradford and Krisztian Posch. 2024. The Effectiveness of Big Data-Driven Predictive Policing: Systematic Review. In *Justice Evaluation Journal. ACJS, Greenbelt, MD, USA, 34 pages.* https://doi.org/10.1080/24751979.2024.2371781

[4] Los Angeles Police Department. *LAPD, Los Angeles, California. https://data.lacity.org/Public-Safety/Crime-Data-from-2020-to-Present/2nrs-mtv8/about_data*

**Appendix**

**Description of Dataset Attributes (from official documentation):**

- DR_NO – Division of Records Number; official file number made up of a 2 digit year, area ID, and 5 digits
- Date Rptd – Date Reported, a timestamp
- DATE OCC – Date Occurred, a timestamp
- TIME OCC – Time Occurred, in military time
- AREA – The LAPD has 21 Community Police Stations referred to as Geographic Areas within the department. These Geographic Areas are sequentially numbered from 1-21.
- AREA NAME – The name designation for the 21 geographic areas or patrol divisions, which reference a landmark or the surrounding country that it is responsible for
- Rpt Dist No – A four-digit code that represents a sub-area within a Geographic Area. All crime records reference the "RD" that it occurred in for statistical comparisons.
- Part 1-2 – Groupings based on the FBI's Uniform Crime Reporting (UCR) program, with Part 1 representing more severe crimes (e.g. homicide, rape, aggravated assault), and Part 2 representing less severe crimes (e.g. gambling, forgery, disorderly conduct)
- Crm Cd – Crime Code; indicates the crime committed
- Crm Cd Desc – Defines the Crime Code provided
- Mocodes – Modus Operandi: activities associated with the suspect in commission of the crime
- Premis Cd – Premise Code, the type of structure, vehicle, or location where the crime took place
- Premis Desc – Premise Description, defines the Premise Code provided
- Weapon Used Cd – The type of weapon used in the crime
- Weapon Desc – Defines the Weapon Used Code provided
- Status – Status of the case (IC is the default)
- Status Desc – Defines the Status Code provided
- Crm Cd 1 – Indicates the crime committed (very frequently matches Crm Cd; however, sometimes not). Crime Code 1 is the primary and most serious crime. Crime Code 2,3, and 4 are respectively less serious offenses. Lower crime class numbers are more serious.

- Crm Cd 2 – May contain a code for an additional crime, less serious than Crime Code 1.
- Crm Cd 3 – May contain a code for an additional crime, less serious than Crime Code 1.
- Crm Cd 4 - May contain a code for an additional crime, less serious than Crime Code 1.
- LOCATION – Street address of crime incident rounded to the nearest hundred block to maintain anonymity
- Cross Street – Cross Street of rounded Address
- LAT – Latitude
- LON – Longitude

**Davis-Bouldin Index**:
A lower Davies-Bouldin Index relates to a model with better separation between the clusters. The index signifies the average 'similarity' between clusters, where the similarity is a measure that compares the distance between clusters with the size of the clusters themselves. Zero is the lowest possible score. Values closer to zero indicate a better partition.

**Clustering Definition by Most Frequent Demographic Members:**

(0) 38 y/o Hispanic Males
(1) 25 y/o Hispanic Females
(2) 56 y/o White Males
(3) 82 y/o White Females
(4) 17 y/o Hispanic Females
(5) 50 y/o Hispanic Males
(6) 32 y/o Hispanic Males
(7) 70 y/o White Males
(8) 63 y/o White Males
(9) 44 y/o Hispanic Males

**Top Five Crime Codes for each K-Means Cluster:**

| | CLUSTER | CRM_CD | count |
|---|---|---|---|
| 16 | 0 | BURGLARY FROM VEHICLE | 31223 |
| 65 | 0 | BATTERY - SIMPLE ASSAULT | 28579 |
| 27 | 0 | THEFT OF IDENTITY | 25324 |
| 67 | 0 | INTIMATE PARTNER - SIMPLE ASSAULT | 22650 |
| 14 | 0 | BURGLARY | 21814 |
| 212 | 1 | BATTERY - SIMPLE ASSAULT | 53950 |
| 164 | 1 | BURGLARY FROM VEHICLE | 53891 |
| 214 | 1 | INTIMATE PARTNER - SIMPLE ASSAULT | 49982 |
| 187 | 1 | THEFT PLAIN - PETTY ($950 & UNDER) | 36740 |
| 155 | 1 | ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT | 32675 |
| 358 | 2 | BATTERY - SIMPLE ASSAULT | 21396 |
| 323 | 2 | THEFT OF IDENTITY | 17253 |
| 310 | 2 | BURGLARY | 16984 |
| 312 | 2 | BURGLARY FROM VEHICLE | 14379 |
| 335 | 2 | THEFT PLAIN - PETTY ($950 & UNDER) | 13853 |
| 450 | 3 | BURGLARY | 7446 |
| 462 | 3 | THEFT OF IDENTITY | 5749 |
| 473 | 3 | THEFT PLAIN - PETTY ($950 & UNDER) | 3809 |
| 489 | 3 | BATTERY - SIMPLE ASSAULT | 3503 |
| 454 | 3 | THEFT-GRAND ($950.01 & OVER)EXCPT,GUNS,FOWL,LI... | 2947 |
| 623 | 4 | BATTERY - SIMPLE ASSAULT | 35566 |
| 567 | 4 | ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT | 19224 |
| 565 | 4 | ROBBERY | 16833 |
| 710 | 4 | Code Undefined | 15877 |
| 600 | 4 | THEFT PLAIN - PETTY ($950 & UNDER) | 14438 |
| 774 | 5 | BATTERY - SIMPLE ASSAULT | 25832 |
| 724 | 5 | BURGLARY | 19877 |
| 737 | 5 | THEFT OF IDENTITY | 19853 |
| 726 | 5 | BURGLARY FROM VEHICLE | 19534 |
| 750 | 5 | THEFT PLAIN - PETTY ($950 & UNDER) | 16619 |
| 870 | 6 | BURGLARY FROM VEHICLE | 49521 |
| 917 | 6 | BATTERY - SIMPLE ASSAULT | 40781 |
| 919 | 6 | INTIMATE PARTNER - SIMPLE ASSAULT | 37974 |
| 881 | 6 | THEFT OF IDENTITY | 33060 |
| 893 | 6 | THEFT PLAIN - PETTY ($950 & UNDER) | 30867 |
| 1017 | 7 | BURGLARY | 12344 |
| 1029 | 7 | THEFT OF IDENTITY | 11435 |
| 1057 | 7 | BATTERY - SIMPLE ASSAULT | 8714 |
| 1037 | 7 | THEFT PLAIN - PETTY ($950 & UNDER) | 7810 |
| 1075 | 7 | VANDALISM - FELONY ($400 & OVER, ALL CHURCH VA... | 5744 |
| 1183 | 8 | BATTERY - SIMPLE ASSAULT | 15730 |
| 1138 | 8 | BURGLARY | 15319 |
| 1150 | 8 | THEFT OF IDENTITY | 15011 |
| 1161 | 8 | THEFT PLAIN - PETTY ($950 & UNDER) | 11275 |
| 1140 | 8 | BURGLARY FROM VEHICLE | 9948 |
| 1317 | 9 | BATTERY - SIMPLE ASSAULT | 26577 |
| 1272 | 9 | BURGLARY FROM VEHICLE | 24263 |
| 1283 | 9 | THEFT OF IDENTITY | 21726 |
| 1270 | 9 | BURGLARY | 20388 |
| 1294 | 9 | THEFT PLAIN - PETTY ($950 & UNDER) | 17869 |