La Cr

asd6413klq

Quick Submit

Quick Submit

Twin Valley High School

Document Details

Submission ID

trn:oid:::1:2742488195

Submission Date

Nov 6, 2023, 6:24 PM EST

Download Date

Nov 6, 2023, 6:29 PM EST

File Name

Crime_report_LA.docx

File Size

1.2 MB

34 Pages

7,635 Words

43,635 Characters

How much of this submission has been generated by AI?

0%

of qualifying text in this submission has been determined to be generated by AI.

Caution: Percentage may not indicate academic misconduct. Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Frequently Asked Questions

What does the percentage mean?
The percentage shown in the AI writing detection indicator and in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was generated by AI.

Our testing has found that there is a higher incidence of false positives when the percentage is less than 20. In order to reduce the likelihood of misinterpretation, the AI indicator will display an asterisk for percentages less than 20 to call attention to the fact that the score is less reliable.

However, the final decision on whether any misconduct has occurred rests with the reviewer/instructor. They should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in greater detail according to their school's policies.

How does Turnitin's indicator address false positives?
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be AI-generated will be highlighted blue on the submission text.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

What does 'qualifying text' mean?
Sometimes false positives (incorrectly flagging human-written text as AI-generated), can include lists without a lot of structural variation, text that literally repeats itself, or text that has been paraphrased without developing new ideas. If our indicator shows a higher amount of AI writing in such text, we advise you to take that into consideration when looking at the percentage indicated.

In a longer document with a mix of authentic writing and AI generated text, it can be difficult to exactly determine where the AI writing begins and original writing ends, but our model should give you a reliable guide to start conversations with the submitting student.

Analysis of Crime in Los Angeles

Student name

SRN

Course Name

Program name

University name

METHODOLOGY

Research Design

Basically, the research design is Exploratory Data Analysis (EDA) with Descriptive Research and Secondary Data Validation The study is characterized by its primarily descriptive and exploratory nature. It does not entail experimental hypotheses testing but rather centers on the analysis of extant data to derive insights and address research inquiries. The dataset has been procured from the Los Angeles Police Department's repository of crime records, constituting a primary data source. The dataset assumes a CSV file format, and the analysis thereof has been conducted through the utilization of Tableau, an established tool for data exploration and visualization. Various data variables have been scrutinized, encompassing gender, ethnicity, Modus operandi (MO), crime code, location, area name, neighborhood, type of weapon, geographical coordinates in the form of latitude and longitude, the age of victims, their descent, and gender.

The study employs exploratory data analysis (EDA) as its fundamental approach. This involves the application of data visualization and summary statistics to discern patterns, characteristics, and trends inherent to the provided crime data, with Tableau serving as the principal instrument for this endeavor. The research questions formulated are integral to the research design, as they constitute the compass guiding the examination and elucidation of the data. The research design features a component dedicated to the validation of findings through secondary data sources, which may include pertinent literature, extant research, and crime statistics, thus enhancing the comprehensiveness and robustness of the analysis.

The research's principal emphasis is on EDA, denoting the practice of summarizing, visualizing, and comprehending the intrinsic characteristics, trends, and patterns encapsulated within the dataset. The research adheres to the tenets of descriptive research, with its primary objective being the articulation of the prevailing state of crime data within the Los Angeles region, devoid of any active manipulation of variables. The research design embraces the practice of corroborating its findings through reference to secondary sources. Such secondary data sources serve the purpose of

substantiating or contextualizing the research's findings, thereby affording them a greater degree of reliability and situatedness within the wider discourse.

Data collection methods

The dataset employed for analysis was procured from "data.gov," an official repository of the United States Government. Established in 2009, "data.gov" is under the management and hosting purview of the U.S. General Services Administration, specifically the Technology Transformation Service division. The dataset featured on "data.gov" was originally sourced and contributed by the Los Angeles Police Department (LAPD). Notably, LAPD undertook the systematic accumulation of this dataset, with each entry being meticulously recorded on a daily basis, wherein each record was individually appended upon receipt of the corresponding incident report. Consequently, the method of data collection is fundamentally predicated on the diligent and routine data entry efforts conducted by LAPD, rendering each entry a distinct and meticulously documented record.

Sampling Strategy

The sampling strategy employed in this study is, temporal stratified sampling, for the analysis. The original dataset, encompassed records spanning from 2010 to 2023. To align with the research objectives, a judicious temporal selection was performed, leading to the retention of data from the years 2017 to 2023, thereby encompassing the contemporary period under examination. This selection decision facilitated the acquisition of data that was contextually relevant to the research at hand.

Subsequent to the temporal delineation, a comprehensive data refinement procedure was enacted to enhance the dataset's quality and relevance for analysis. Primarily, data cleansing was executed through the utilization of Python programming. The cleaning process was initiated by conducting a meticulous evaluation of the dataset's integrity. Key components of this evaluation included the identification of missing values and a descriptive analysis, which provided a comprehensive overview of the dataset's structure.

From this assessment, it became evident that certain columns were deemed superfluous to the research objectives and were subsequently eliminated from the

dataset. The removed columns include "Part 1-2," "Premis Cd," "Weapon Used Cd," "Status," "Crm Cd," "Crm Cd 2," "Crm Cd 3," "Crm Cd 4," "LOCATION," "AREA," "Mocodes," "Premis Desc," and "Cross Street." This deliberate culling of columns was instrumental in streamlining the dataset, resulting in a more concise and focused dataset, ultimately comprising 15 pertinent columns.

Furthermore, to ensure data integrity and statistical rigor, targeted removal of specific empty cells was performed. These deletions were not indiscriminate; rather, they were executed judiciously to eliminate cells that held significance in terms of data quality. Notably, cells with missing values, particularly within the "Victim Sex" column, which hindered the determination of the victim's gender, were systematically removed. This meticulous process of data cleansing and removal of empty cells was indispensable in mitigating the potential impact of incomplete or erroneous data on the research outcomes.

Consequently, the refined dataset, meticulously curated to encompass records from 2017 to 2023, devoid of superfluous columns and featuring judiciously managed missing values, culminated in a dataset comprising 1,303,052 rows and 15 columns. This dataset, as a result of the systematic sampling strategy and data cleansing procedures, is poised to serve as a robust foundation for the ensuing data analysis and research inquiries.

Data analysis methods

Upon the completion of data preparation and cleaning, the dataset was subsequently subjected to a comprehensive analytical exploration within the Tableau data visualization platform. The ensuing phase of data analysis was guided by the research objectives and a framework designed to address the research hypothesis and research questions. The analytical process encompassed several key visualizations, each tailored to illuminate distinct aspects of the dataset:

- A time series line chart was generated to delineate the temporal evolution of the city's overall crime rate from 2017 to 2023. This visualization harnessed "Date Occurred" as the x-axis and "Crime Count" as the y-axis, affording a longitudinal perspective on crime trends throughout the specified period.

- Bar charts were harnessed to explore the distribution of crimes with respect to specific demographic variables, namely "Victim Age," "Victim Sex," and "Victim Descent." The objective was to uncover potential correlations between these demographic factors and crime trends, thereby contributing to a deeper understanding of the dataset.
- The geographical distribution of crime incidents was elucidated through the deployment of a scatter plot. Leveraging latitude and longitude data, this visualization facilitated the geographic identification of crime occurrences. The color-coded differentiation of data points by "Crime Code" enriched the understanding of crime typologies across distinct locales.
- A heatmap was constructed to visualize crime density as it relates to geographical coordinates, specifically latitude and longitude. This depiction was instrumental in the identification of areas characterized by elevated crime rates, thus informing potential resource allocation decisions for law enforcement agencies.
- A stacked bar chart was employed to elucidate the distribution of crime report statuses ("Status") and their corresponding clearance rates. The utilization of "Status" on the x-axis, with stacked bars representing clearance rates, delivered insights into the progression of reported crimes through different clearance stages.
- The time required to resolve or clear cases was explored through a histogram, allowing for the identification of trends pertaining to case resolution times. This facilitated an assessment of potential delays in the resolution process, which may carry implications for law enforcement efficiency.
- A line chart, spanning the years from 2019 to 2023, was complemented by an overlay of economic indicators, most notably the "Unemployment Rate." This visualization sought to discern potential correlations between economic factors and crime rates, shedding light on any concurrent trends or dynamics.
- The relationship between crime rates (represented by the size of bubbles) and economic variables, with the "Unemployment Rate" as the x-axis variable, was elucidated through the deployment of a bubble chart. This dynamic representation aimed to identify areas of potential economic intervention by highlighting intersections between economic indicators and crime rates over time.

Each of these visualization strategies was meticulously tailored to distill pertinent insights from the dataset, ultimately contributing to a multifaceted and informed

analysis. By systematically addressing the research hypothesis and research questions through diverse visualization techniques, this analytical approach enhances the depth and comprehensiveness of the ensuing research investigation.

After analyzing the data in Tableau, Decision Tree and Random Forest regression models were built using python programming language. The LAPD's data was utilized to train these models separately and further on these models were tested, predictions were conducted. In this endeavor, accuracy, r squared values and mean square error were utilized for the validation of these models.

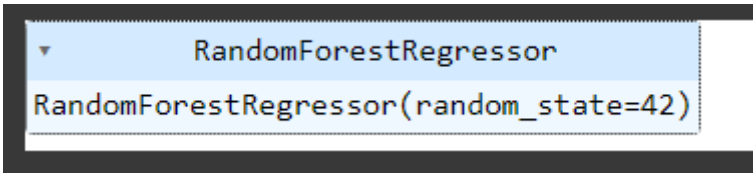

Figure 1: Random Forest Model Training visual

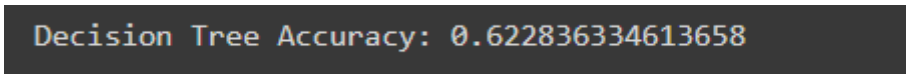

Figure 2: Decision Tree Model's accuracy after training

Ethical considerations

This research project strictly adheres to ethical principles and legal obligations, particularly with regard to the OPEN Government Data Act and the Foundations for Evidence-Based Policymaking Act. Rigorous measures are in place to protect sensitive information, respecting individual privacy rights within the dataset. Data fields that could reveal personal identities have been anonymized to mitigate privacy risks. Robust security protocols ensure the confidentiality and integrity of the data. Proper attribution and citation practices are followed when using secondary data sources. The research prioritizes data transparency, promoting accessibility and reproducibility. The project complies with relevant regulations, including the OPEN Government Data Act. The research aims to benefit society while ensuring it does no harm, respecting the rights and dignity of individuals and communities. These ethical considerations underscore the responsible use of government data, balancing knowledge generation and evidence-based policy with privacy, security, and legal adherence.

Reliability and validation

The data sourced from "Data.gov" adheres to a well-established framework for government data publication. Federal agencies, including the LAPD, are mandated to catalog and list their data assets on "Data.gov." This process entails an intricate chain of validation, beginning with each agency's designated open data point of contact. This liaison ensures the authenticity and accuracy of the data being published, acting as a guardian of reliability at the agency level. This inventory entails the meticulous cataloging of data assets, mirroring the care with which organizations inventory physical assets such as computers or office equipment. This systematic approach bolsters the reliability of the data, enhancing its manageability and utility. The specific dataset pertaining to crime data in Los Angeles City, supplied by LAPD, is noteworthy for its meticulous data collection methodology. Each row within this dataset represents an individual crime report. This granular approach to data collection underscores the reliability of the dataset, as it aligns with the established procedures of law enforcement agencies. These stringent protocols and processes collectively contribute to the robustness and trustworthiness of the data used in this research.

Data triangulation

Primary analysis employs LAPD data. Secondary sources like online research papers and LAPD's official website validate and authenticate assumptions, cross-verify findings, broaden contextual understanding, reduce bias, and enhance research credibility. Triangulation combines primary and secondary data, fortifying validity and trustworthiness for meaningful conclusions.

Limitations

This research project, while founded on LAPD data, exhibits certain inherent limitations that necessitate consideration:

- The research is contingent exclusively upon LAPD data. The omission of any additional data sources potentially restricts the comprehensive scope of the analysis. The reliance on a singular data source may inadvertently overlook nuances that could be illuminated by supplementary datasets.

- During data preparation, specific rows were removed due to data quality concerns. While these exclusions were meticulously executed to enhance data integrity, their elimination could have a marginal impact on the research outcomes. The absence of these rows may influence the representativeness of the dataset.
- The research fundamentally hinges on the understanding of human behavior and its correlation with crime trends. It is imperative to acknowledge that human behavior is inherently multifaceted and dynamic, rendering it challenging to assert with absolute certainty that the research results capture the entire spectrum of human actions and interactions. As such, the trends and patterns unveiled within the dataset may exhibit temporal variations, and the research outcomes may not be universally and perpetually applicable.

In recognition of these limitations, it is essential to approach the research findings with a degree of cautious interpretation. The acknowledgment of these constraints offers valuable context for the study's results and underscores the need for further research and ongoing validation to refine our comprehension of the intricate dynamics underpinning human behavior and crime trends over time.

RESULTS AND DISCUSSION

Data Analysis

As mentioned previously in methodology, that the dataset was stored in a comma separated file. Since, analyzing more than a million rows of data without any visualization tool, seems impractical and full of errors, thus Tableau was selected for the visualization purposes. Tableau no doubt was the primary analysis tool but prior to tableau just in order to get an over view of the dataset, Exploratory Data Analysis (EDA) was performed on the dataset using python. Afterwards, utilized the cleaned dataset and loaded it into the data source and started analyzing the dataset variable by variable. Comparison of various columns took place as well. In short, the data analyzed with the use of tableau's data visualization tools and as resultant, we obtained insightful visuals from the data. The results are mentioned in the next section.

Results

EDA

For the sake of exploratory data analysis, the data was accessed from the directory using the pandas library and stored into data frame. After necessary data transformation, the data is visualized using seaborn and matplot library. The results of EDA are described in the figure 1, 2 and 3.

Figure 3: EDA- =1

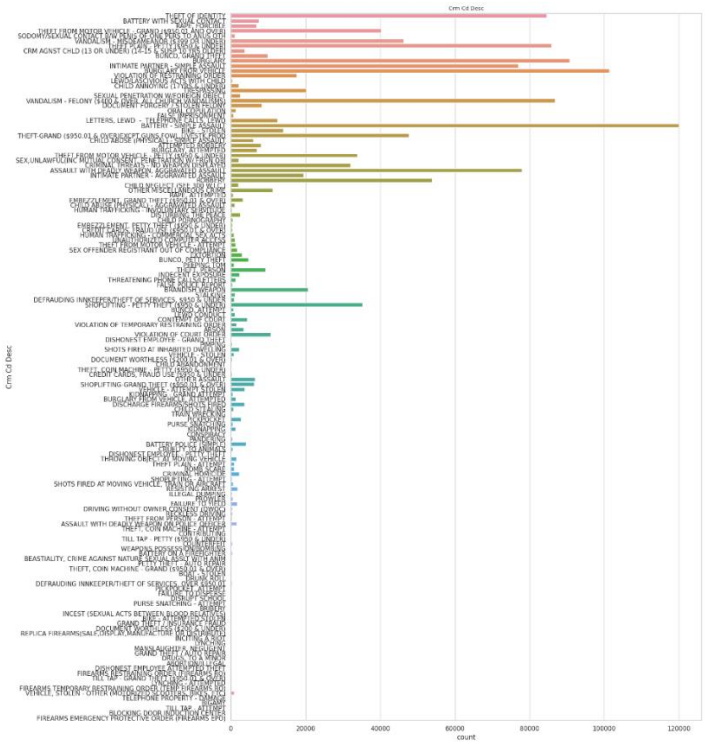

Figure 4: EDA 2 (crime type)

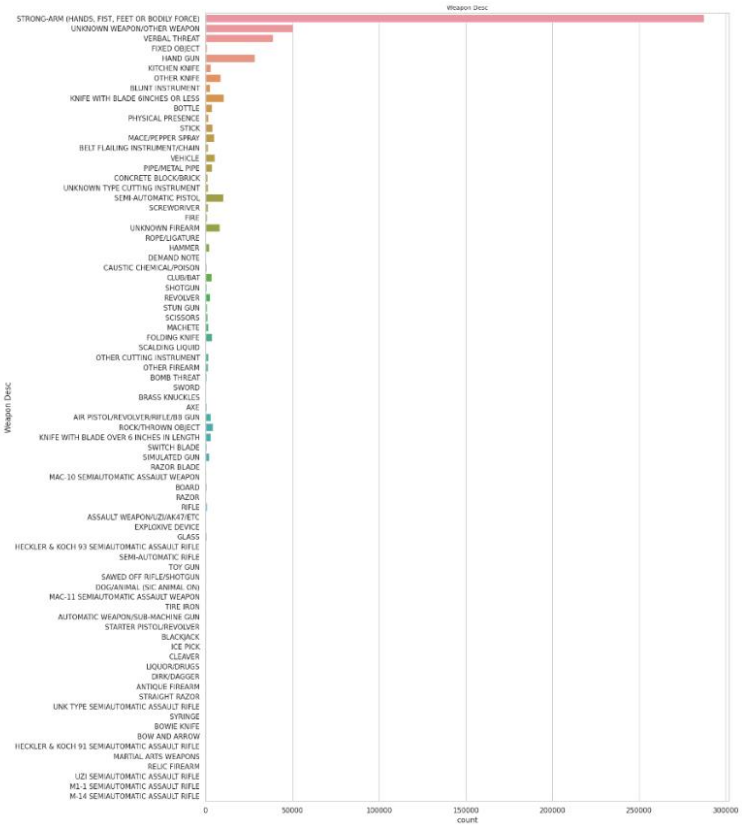

Figure 5: EDA-3 (weapon description)

Crime 2017-23

The analysis of the total number of crimes reported in Los Angeles from 2017 to 2023 offers valuable insights into the temporal trends and fluctuations within the dataset. The data is presented through a discrete line chart, which visually portrays the variations in reported crime incidents over this six-year period.

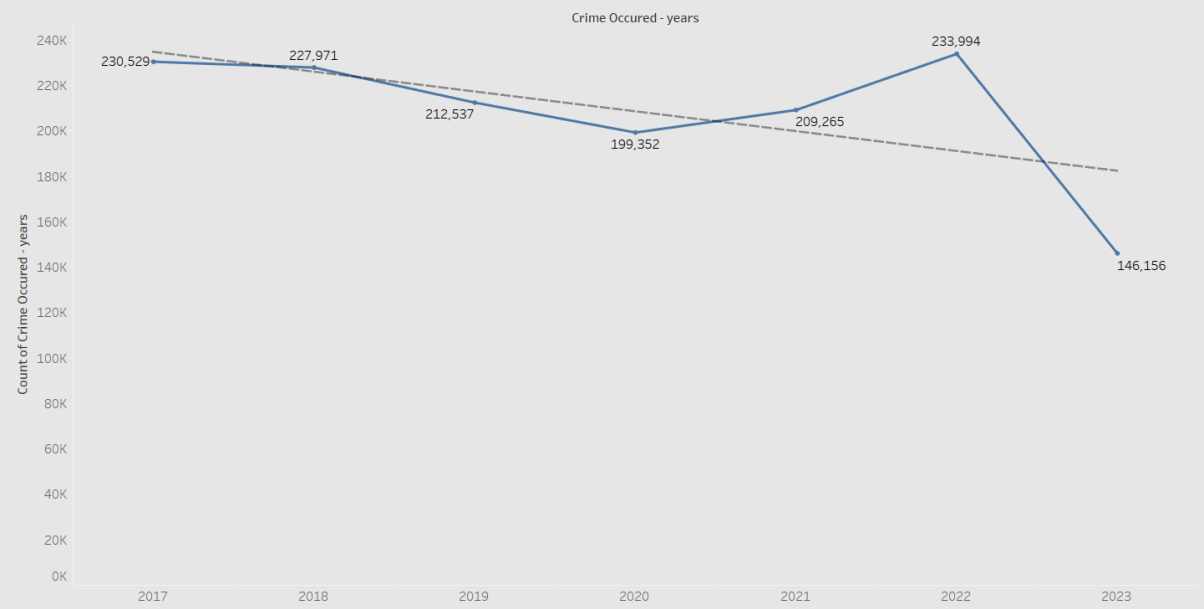

Figure 6: Total crime reported from 2017-2023

As depicted in the line chart, the number of reported crimes in Los Angeles exhibited notable variations from 2017 to 2023. The data reveals a distinct pattern of fluctuation, reflecting shifts in the city's crime landscape. In 2017, the dataset records a total of 230,529 reported crimes. Subsequently, in 2018, the reported crime incidents slightly decreased to 227,971. The year 2019 saw a further reduction in reported crimes, with 212,537 incidents documented. However, 2020 marked a substantial shift in this trend, as the number of reported crimes decreased to 199,352. The subsequent year, 2021, witnessed a reversal of the declining trend, with reported crime incidents increasing to 209,265. This trend continued into 2022, wherein the dataset documented 233,994 reported crimes. It is important to note that the dataset for 2023 comprises data up to August 28th, recording 146,156 reported crimes during this period.

Crime analysis ethnicity based 2017-23

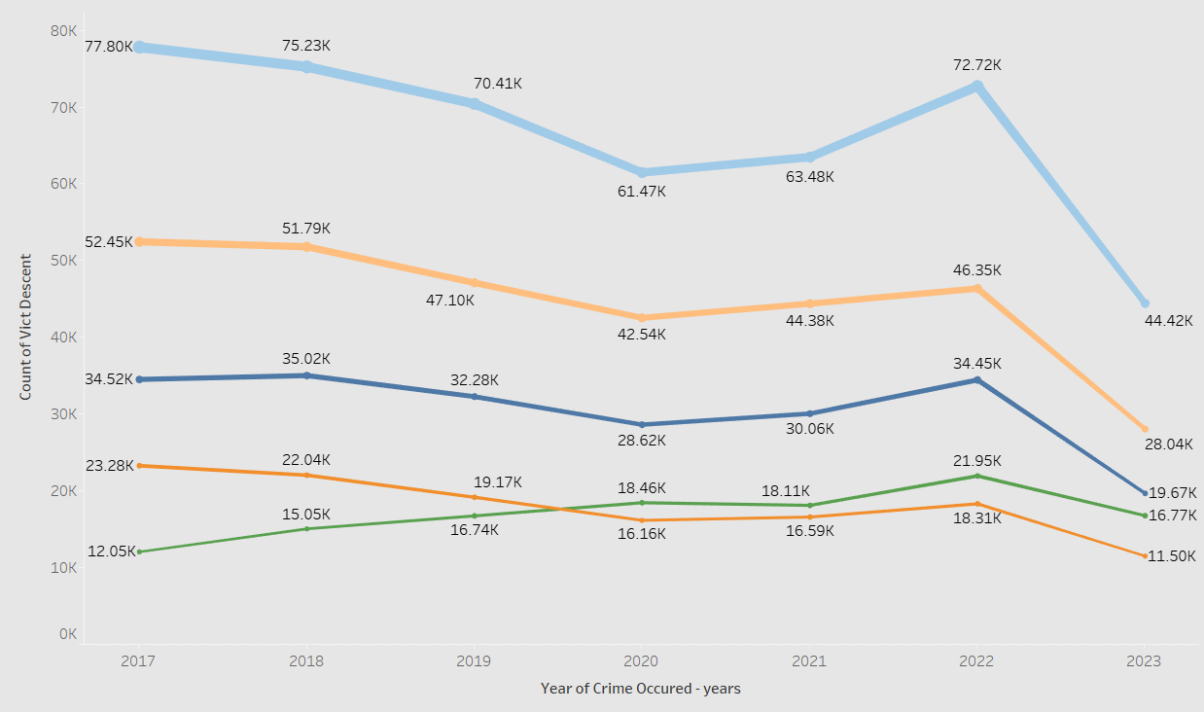

Figure 7: Crime - ethnicity timeline.

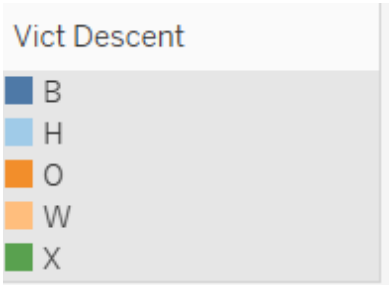

Figure 8: legend - ethnicity

The analysis of crime categorized by ethnicity, specifically white (W), black (B), Hispanic (H), other (O), and unknown (X), reveals notable trends and variations in reported crime incidents over the years 2017 to 2023. The data is presented through a comprehensive examination of crime incidents within each ethnic group:

Hispanic (H)

In 2017, the dataset records approximately 77,800 reported crimes within the Hispanic community. Subsequently, in 2018, the number of reported crimes decreased to approximately 75,230. The trend of declining crime incidents within the Hispanic

community continued into 2019, with 70,410 incidents. A notable decrease was observed in 2020, with approximately 61,470 reported crimes. However, in 2021, there was a slight increase to around 63,480 reported crimes. 2022 marked a further increase, with approximately 72,720 reported crimes. The dataset for 2023, spanning from January 1st to August 28th, records approximately 44,420 reported crimes within the Hispanic community.

White (W)

In 2017, approximately 52,450 reported crimes are attributed to the white ethnic group. In 2018, this number slightly decreased to approximately 51,790 reported crimes. A further decline was observed in 2019, with approximately 47,100 reported crimes. The number of reported crimes within the white community decreased to approximately 42,540 in 2020. In 2021, there was a slight increase to around 44,380 reported crimes. 2022 witnessed an additional increase, with approximately 46,350 reported crimes. The dataset for 2023, covering the period from January 1st to August 28th, records approximately 28,040 reported crimes within the white community.

Black (B)

In 2017, the dataset documents approximately 34,520 reported crimes within the black community. In 2018, this number increased to approximately 35,020 reported crimes. In 2019, the number of reported crimes slightly decreased to approximately 32,280. A further decline was observed in 2020, with approximately 28,620 reported crimes. However, in 2021, there was an increase to approximately 30,060 reported crimes. 2022 marked an additional increase, with approximately 34,450 reported crimes. The dataset for 2023, covering the period from January 1st to August 28th, records approximately 19,670 reported crimes within the black community.

Other (O)

In 2017, approximately 23,280 reported crimes are attributed to the 'other' ethnic category. In 2018, this number decreased to approximately 22,040 reported crimes. A further decline was observed in 2019, with approximately 19,170 reported crimes. The number of reported crimes within the 'other' category decreased to approximately 16,160 in 2020. In 2021, there was a slight increase to around 16,590 reported crimes.

2022 witnessed an additional increase, with approximately 18,310 reported crimes. The dataset for 2023, spanning from January 1st to August 28th, records approximately 11,500 reported crimes within the 'other' category.

Unknown (X)

In 2017, the dataset documents approximately 12,050 reported crimes within the 'unknown' ethnic category. In 2018, this number increased to approximately 15,050 reported crimes. In 2019, the number of reported crimes continued to rise, with approximately 16,740. An additional increase was observed in 2020, with approximately 18,460 reported crimes. However, in 2021, there was a slight decrease to around 18,110 reported crimes. 2022 marked a further increase, with approximately 21,950 reported crimes. The dataset for 2023, covering the period from January 1st to August 28th, records approximately 16,770 reported crimes within the 'unknown' category.

Crime Analysis based on Gender 2017-23

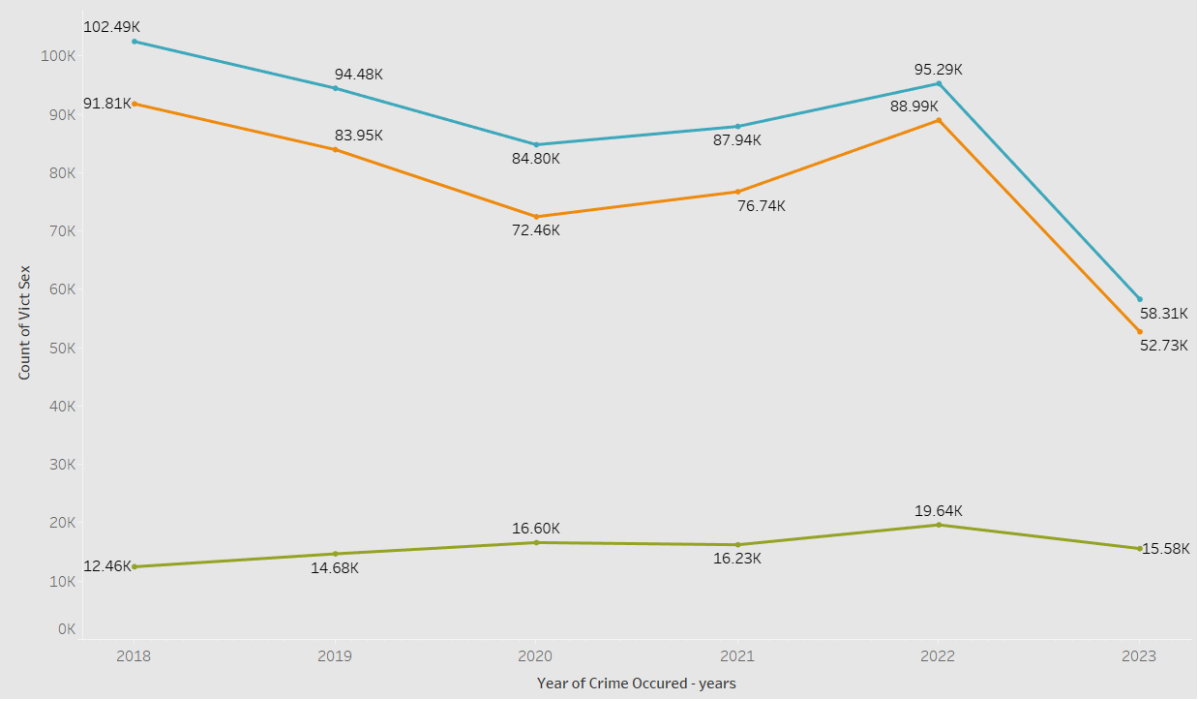

Figure 9:Figure 3: Crime by gender timeline

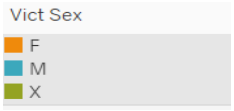

Figure 10: Legend Gender

The analysis of reported crimes categorized by gender, specifically male (M), female (F), and other (X), unveils significant patterns and fluctuations in crime incidents over the years 2017 to 2023. The data is presented through a comprehensive examination of crime incidents within each gender category:

Male (M)

In 2017, the dataset records approximately 104,420 reported crimes involving male individuals. In 2018, this number slightly decreased to approximately 102,490 reported crimes. The trend of declining male-involved crime incidents continued into 2019, with 94,480 incidents documented. A substantial decrease was observed in 2020, with approximately 84,800 reported crimes. However, in 2021, there was a slight increase to around 87,940 reported crimes. 2022 witnessed a further increase, with approximately 95,290 reported crimes. The dataset for 2023, spanning from January 1st to August 28th, records approximately 58,310 reported crimes involving male individuals.

Female (F)

In 2017, approximately 93,310 reported crimes are attributed to female individuals. In 2018, this number decreased to approximately 91,810 reported crimes. A further decline was observed in 2019, with approximately 83,950 reported crimes involving female individuals. The number of reported crimes involving female individuals decreased to approximately 72,460 in 2020. In 2021, there was a slight increase to around 76,740 reported crimes. 2022 witnessed an additional increase, with approximately 88,990 reported crimes involving female individuals. The dataset for 2023, covering the period from January 1st to August 28th, records approximately 52,730 reported crimes involving female individuals.

Other (X)

Approximately 9,590 recorded offenses in 2017 involved people who were classified as 'other.' This figure rose to over 12,460 recorded crimes in 2018. With over 14,680

recorded offenses involving "other" people in 2019, the number of such crimes rose further in 2019. 2020 saw a further uptick, with over 16,600 offenses involving "other" people being reported. Nonetheless, there was a minor decline to about 16,230 reported offenses in 2021. 2022 witnessed a further surge, with roughly 19,640 recorded offenses involving 'other' people. The 2023 dataset, which spans the dates January 1st to August 28th, includes information on almost 15,580 reported offenses involving people who were classified as "other."

Area wise crime analysis Los Angeles 2017-23

The analysis of crime incidents in different areas of Los Angeles over the years 2017 to 2023 provides valuable insights into the variations and trends within these regions. The data is presented through an examination of crime incidents within each area:

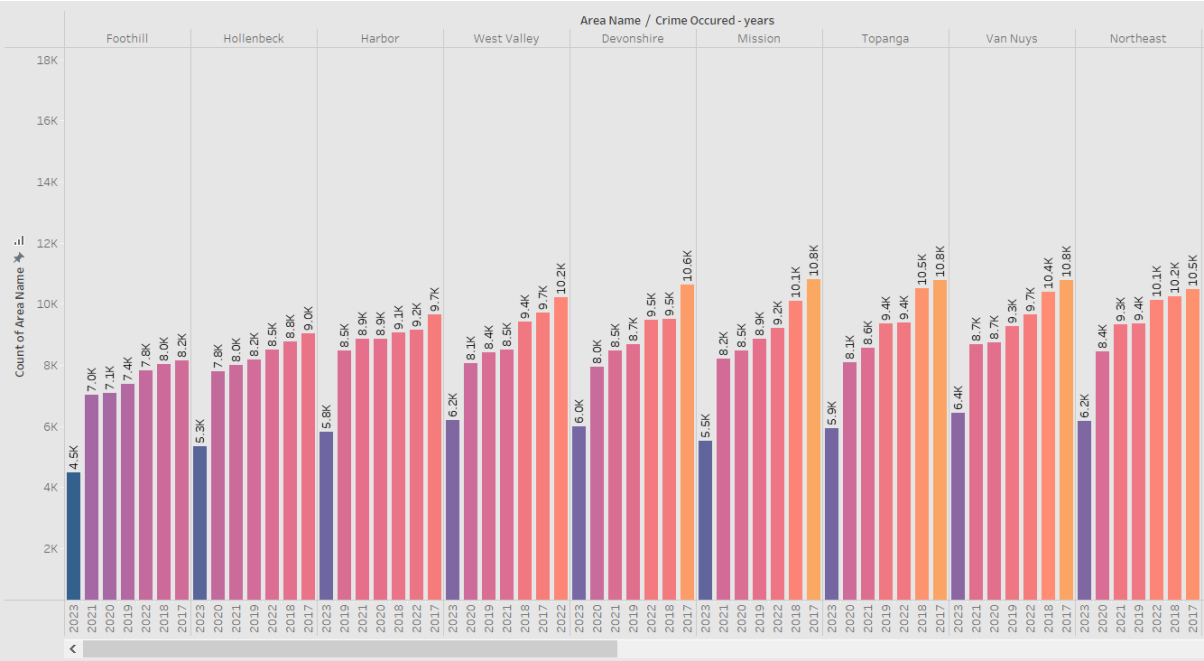

Figure 11: Area to crime analysis timeline - 1

Harbor Area

In the Harbor area, 9,662 criminal incidents were recorded in 2017. The number of reported crimes in the Harbor area decreased over time, with figures of 9,060 in 2018, 8,481 in 2019, 8,856 in 2020, 8,853 in 2021, and 9,155 in 2022. Information on 5,823 recorded criminal incidents in the Harbor area between January 1 and August 28 of 2023 is included in the dataset.

Hollenbeck Area

9,034 criminal occurrences were reported in the Hollenbeck region in 2017. With 8,769 reported crimes in 2018, 8,188 in 2019, 7,797 in 2020, 8,009 in 2021, and 8,490 in 2022, Hollenbeck reported less crimes on average each year. The dataset for 2023 contains information on 5,332 reported criminal occurrences in the Hollenbeck region, covering the period from January 1st to August 28th.

Topanga Area

10,772 crime incidents were reported in the Topanga area in 2017. There were fluctuations in the number of crimes that were reported in Topanga: 10,524 in 2018, 9,354 in 2019, 8,091 in 2020, 8,566 in 2021, and 9,398 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 5,940 reported crime incidents in the Topanga area.

Van Nuys Area

In 2017, the Van Nuys area reported a total of 10,769 crime incidents. The number of reported crimes in Van Nuys displayed fluctuations, with 10,396 in 2018, 9,274 in 2019, 8,741 in 2020, 8,677 in 2021, and 9,656 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 6,423 reported crime incidents in the Van Nuys area.

Foothill Area

In the Foothill area, 8,156 criminal incidents were recorded in 2017. The number of crimes that were reported decreased in the ensuing years: 8,025 in 2018, 7,372 in 2019, 7,093 in 2020, 7,023 in 2021, and 7,813 in 2022. The dataset for 2023 spans the days of January 1 through August 28 and contains 4,496 recorded criminal incidents in the Foothill region.

Devonshire Area

In the Devonshire region, 10,642 criminal incidents were recorded in 2017. Devonshire's reported crime rate showed a declining trend, with 9,503 recorded crimes in 2018, 8,683 in 2019, 7,955 in 2020, 8,478 in 2021, and 9,485 in 2022. The 2023 dataset records 6,005 criminal incidents that were reported in the Devonshire region between January 1 and August 28..

Mission Area

10,810 crime incidents were reported in the Mission area in 2017. There were variations in the number of crimes reported in the Mission area: 10,092 in 2018, 8,868 in 2019, 8,458 in 2020, 8,208 in 2021, and 9,210 in 2022. The dataset for 2023 includes 5,526 reported crime incidents in the Mission area, covering the period from January 1st to August 28th.

Northeast Area

In 2017, the Northeast area reported a total of 10,493 crime incidents. The number of reported crimes in the Northeast area showcased variations, with 10,242 in 2018, 9,353 in 2019, 8,442 in 2020, 9,333 in 2021, and 10,141 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 6,159 reported crime incidents in the Northeast area.

West Valley Area

9,723 criminal occurrences were reported in the West Valley area in 2017. In the West Valley, there have been varying numbers of recorded crimes over the years: 9,412 in 2018, 8,423 in 2019, 8,066 in 2020, 8,515 in 2021, and 10,220 in 2022. This dataset covers 6,194 reported criminal occurrences in the West Valley area between January 1st and August 28th of 2023.

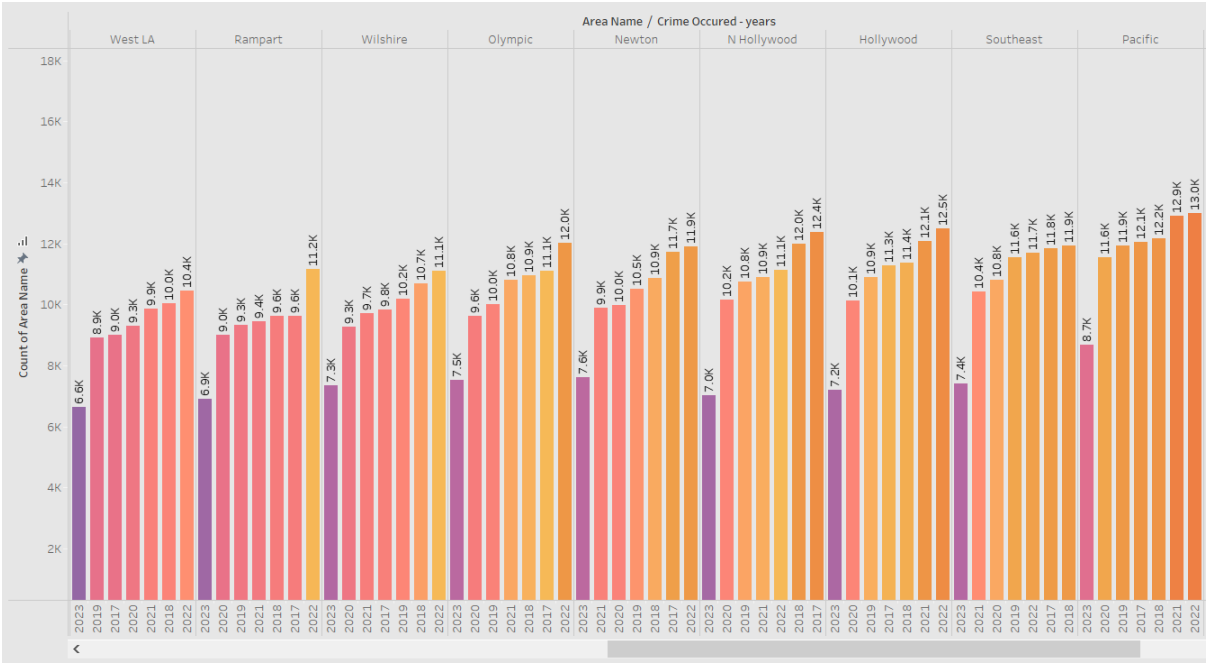

Figure 12: Area to crime analysis timeline - 2

West LA Area

In 2017, the West LA area reported a total of 9,008 crime incidents. Subsequently, the number of reported crimes showed fluctuations, with 10,036 in 2018, 8,923 in 2019, 9,288 in 2020, 9,873 in 2021, and 10,443 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 6,646 reported crime incidents in the West LA area.

Rampart Area

In 2017, the Rampart area reported a total of 9,623 crime incidents. The number of reported crimes in Rampart exhibited variations, with 9,620 in 2018, 9,326 in 2019, 9,005 in 2020, 9,441 in 2021, and 11,157 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 6,897 reported crime incidents in the Rampart area.

Wilshire Area

In 2017, the Wilshire area reported a total of 9,844 crime incidents. The number of reported crimes in Wilshire showcased fluctuations, with 10,689 in 2018, 10,187 in 2019, 9,272 in 2020, 9,716 in 2021, and 11,103 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 7,335 reported crime incidents in the Wilshire area.

Olympic Area

In 2017, the Olympic area reported a total of 11,112 crime incidents. The number of reported crimes in the Olympic area displayed variations, with 10,943 in 2018, 10,005 in 2019, 9,621 in 2020, 10,797 in 2021, and 12,004 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 7,533 reported crime incidents in the Olympic area.

Newton Area

In 2017, the Newton area reported a total of 11,733 crime incidents. The number of reported crimes in Newton exhibited fluctuations, with 10,852 in 2018, 10,523 in 2019,

9,974 in 2020, 9,899 in 2021, and 11,898 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 7,602 reported crime incidents in the Newton area.

North Hollywood Area

In 2017, the North Hollywood area reported a total of 12,384 crime incidents. The number of reported crimes in North Hollywood showcased variations, with 11,988 in 2018, 10,751 in 2019, 10,151 in 2020, 10,897 in 2021, and 11,120 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 7,039 reported crime incidents in the North Hollywood area.

Hollywood Area

In 2017, the Hollywood area reported a total of 11,266 crime incidents. The number of reported crimes in Hollywood exhibited fluctuations, with 11,371 in 2018, 10,905 in 2019, 10,140 in 2020, 12,077 in 2021, and 12,494 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 7,208 reported crime incidents in the Hollywood area.

Southeast Area

In 2017, the Southeast area reported a total of 11,839 crime incidents. The number of reported crimes in the Southeast area displayed variations, with 11,924 in 2018, 11,556 in 2019, 10,817 in 2020, 10,410 in 2021, and 11,696 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 7,418 reported crime incidents in the Southeast area.

Pacific Area

In 2017, the Pacific area reported a total of 12,057 crime incidents. The number of reported crimes in the Pacific area showcased fluctuations, with 12,152 in 2018, 11,925 in 2019, 11,556 in 2020, 12,907 in 2021, and 12,993 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 8,669 reported crime incidents in the Pacific area.

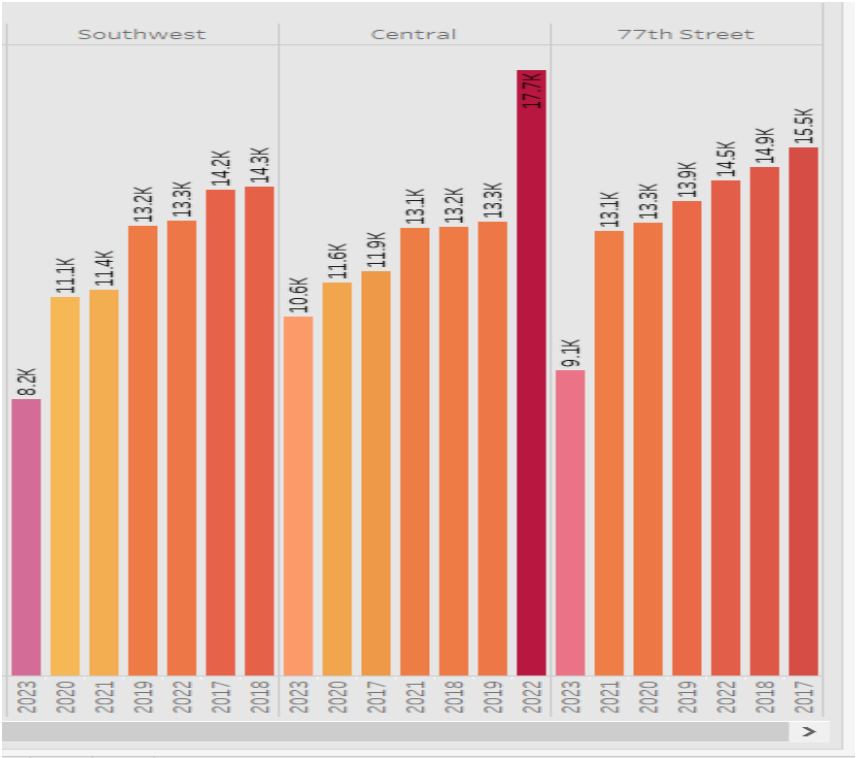

Figure 13: Area to crime analysis timeline - 3

Southwest Area

In 2017, the Southwest area reported a total of 14,246 crime incidents. Subsequently, the number of reported crimes showed fluctuations, with 14,314 in 2018, 13,208 in 2019, 11,149 in 2020, 11,376 in 2021, and 13,345 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 8,247 reported crime incidents in the Southwest area.

Central Area

In 2017, the Central area reported a total of 11,898 crime incidents. The number of reported crimes in the Central area exhibited fluctuations, with 13,172 in 2018, 13,325 in 2019, 11,586 in 2020, 13,145 in 2021, and 17,658 in 2022. The dataset for 2023, spanning from January 1st to August 28th, records 10,611 reported crime incidents in the Central area.

77th Street Area

In 2017, the 77th Street area reported a total of 15,458 crime incidents. The number of reported crimes in the 77th Street area showcased fluctuations, with 14,887 in 2018, 13,907 in 2019, 13,294 in 2020, 13,065 in 2021, and 14,515 in 2022. The dataset for

2023, spanning from January 1st to August 28th, records 9,053 reported crime incidents in the 77th Street area.

Analysis of Report Status 2017-23

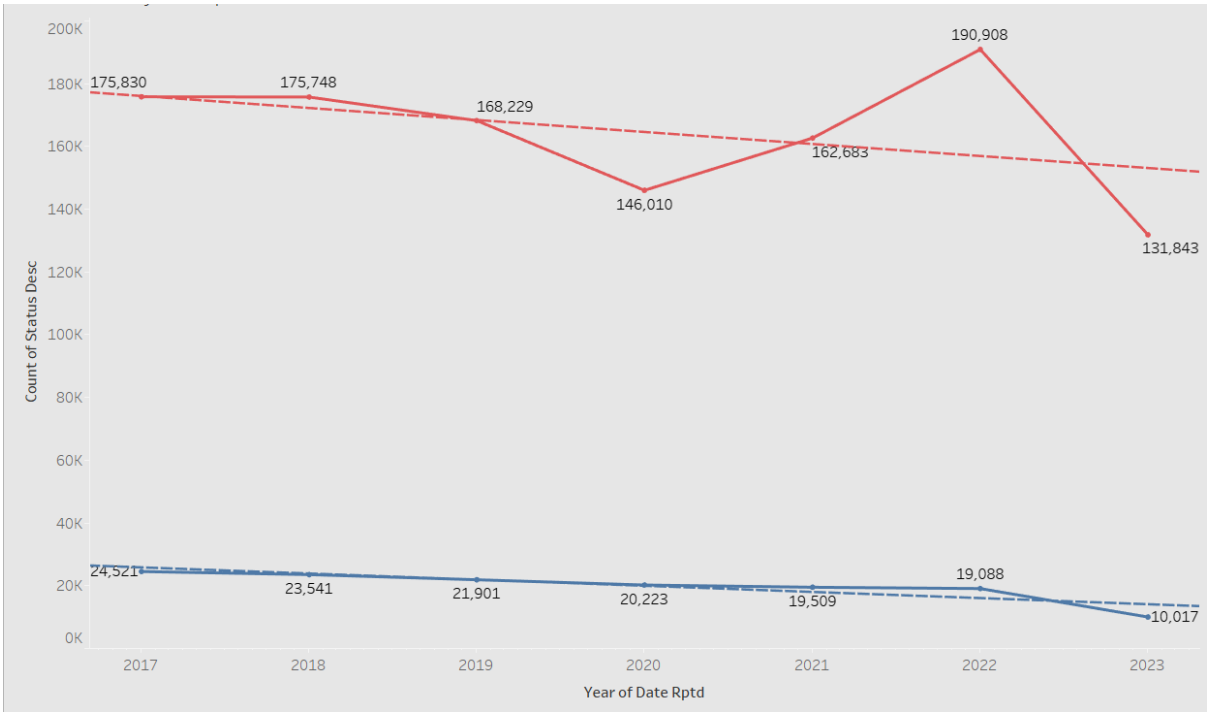

Figure 14: Status analysis timeline

Figure 15: Legend - Status

The analysis of the status of reported crime incidents in Los Angeles for the years 2017 to 2023 provides insights into the disposition of cases. The data is categorized into two main status descriptions investigation continue and adult arrest. In 2017, there were 175,830 cases where the investigation was still in progress. This trend continued with 175,748 cases in 2018, 168,229 cases in 2019, 146,010 cases in 2020, 162,683 cases in 2021, 190,908 cases in 2022. The 2023 dataset, which covers the period from January 1st to August 28th, includes 131,843 cases that are still being investigated. Adult arrests were made in 24,521 cases in 2017, indicating the conclusion of these cases. The following years saw a decline in the number of adult arrest cases: 23,541 in 2018, 21,901 in 2019, 20,223 in 2020, 19,509 in 2021, and

19,088 in 2022. The dataset for 2023 includes 10,017 cases where adult arrests were made and runs from January 1 to August 28.

Crime forecast 2017-2025

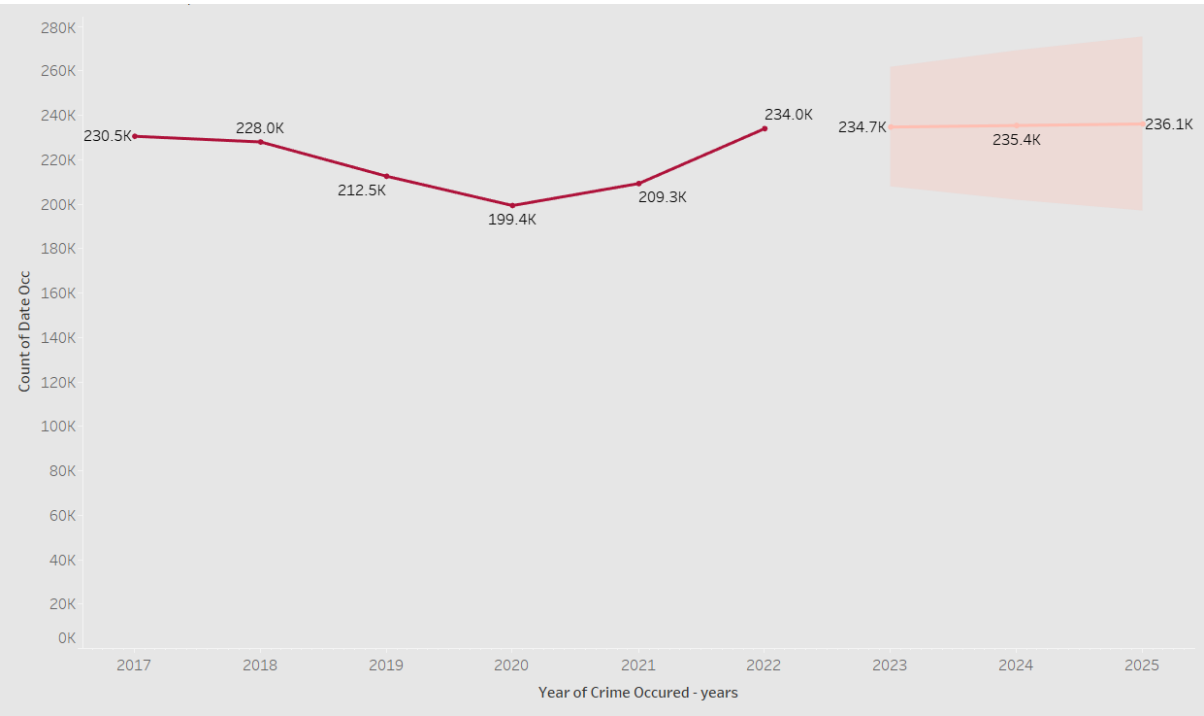

Figure 16: Crime forecast

The Los Angeles crime incident forecast for the years 2023 to 2025 offers important information about anticipated trends in crime rates. Actual crime statistics from 2017 to 2022 are included in the data, as are projected statistics for 2023 to 2025. There were 230,529 real crime incidents that were reported in 2017. The real number for 2018 was 227,971. With 212,537 in 2019 and 199,352 in 2020, the pattern persisted. In 2021, there was an increase to 209,265 reported incidents, followed by 233,994 in 2022. In 2023, it is estimated that there will be 234,687 reported crime incidents. The forecast for 2024 predicts 235,380 incidents, and in 2025, it is estimated to be 236,073.

Prediction

As mentioned earlier, the data was utilized for training of predictive models in order to perform prediction. Predictive model is necessary to get future predictions for crime, also in case some of the information is missing, this model can fulfill the missing data through prediction. Moreover, these predictions can help in assisting the unresolved cases. The accuracies of these models are as under:

```
Decision Tree Accuracy on Test Set: 0.622836334613658
Classification Report on Test Set:
              precision    recall  f1-score   support

           1       0.57      0.58      0.58    113169
           2       0.00      0.00      0.00        25
           3       0.61      0.61      0.61    126465
           4       0.00      0.00      0.00         2
           5       0.95      0.95      0.95     20950

    accuracy                           0.62    260611
   macro avg       0.43      0.43      0.43    260611
weighted avg       0.62      0.62      0.62    260611

Confusion Matrix on Test Set:
 [[65605    10 47289     0   265]
 [   12     0    12     0     1]
 [48790     9 76848     3   815]
 [    2     0     0     0     0]
 [  259     0   826     0 19865]]
```

Figure 17: Decision tree accuracy and confusion matrix

```
Mean Squared Error (MSE): 246.81779431670523
```

Figure 18: MSE value of Random Forest Model

```
R-squared (R2): 0.3570848960886319
```

Figure 19: R-squared value of Random Forest

Discussion

The analysis of reported crime data from the Los Angeles Police Department (LAPD) for the years 2017-2023 revealed distinct temporal trends. Notably, the data exhibited a consistent pattern from 2017 to 2020, with a gradual decrease in the number of reported cases (Lopez & Rosenfeld, 2021). However, a notable shift in the data occurred from 2021 onwards marking the beginning of a new trend characterized by rising crime rates. This upward trajectory continued into 2022, with 233,000 reported

cases, signifying a sustained increase in crime. This projection, based on the observed data, indicates a continued increase in crime rates only, if necessary, measures are not taken and deployed in the city ASAP.

The findings from this analysis confirm a noteworthy shift in the dynamics of reported crime in Los Angeles City over the examined period. The initial years, from 2017 to 2020, saw a consistent decline in the number of reported cases. In contrast, the subsequent years, starting from 2021, demonstrated a marked increase in crime rates (Piquero et al., 2021). The significant and sustained rise in reported crime cases warrants further investigation to identify potential causal factors.

The observed trends align with the assumption that crime rates in Los Angeles City have increased post-2020. However, this analysis does not explore the underlying reasons for the observed patterns. It serves as an empirical foundation for future research and policy considerations related to addressing and mitigating crime in the region. The data's transition from declining to ascending trends underscores the complexity of crime dynamics and the need for ongoing monitoring and analysis to inform effective crime prevention and law enforcement strategies.

Changing demographic composition of Los Angeles

There is a strong correlation between changes in Los Angeles's general crime rate between 2017 and 2023 and the city's shifting demographic makeup. Different patterns are revealed by analyzing gender and ethnicity data (Glaeser et al., 2018). While white, black, and other ethnic groups exhibited different patterns, Hispanic and "unknown" ethnic groups reported an increase in crime events. Crime rates increased for men and those classified as "other," while they fluctuated for women. There is a correlation between crime trends and some demographic characteristics, like gender and ethnicity. This suggests that comprehending swings in crime rates requires taking population dynamics into account.

Geographical distribution of crime incidents

The examination of crime incidents in different parts of Los Angeles between 2017 and 2023 shows unique trends and patterns in each area. The data, which is presented as a percentage change from the previous year, offers crucial background information for comprehending the geographical distribution of criminal occurrences. With a few minor

exceptions, the Foothill Area's crime incidents followed a generally consistent trend. The Hollenbeck Area saw a continuous decrease in reported crime, while the Harbor Area saw a similar trend with a few minor variations. Variations were observed in the West Valley Area, while the Devonshire Area showed signs of decline. There were differences in the Topanga Area and the Mission Area in terms of crime incidents.

Crime reports were erratic in the Van Nuys Area and variable in the Northeast Area. There were swings in the Rampart Area and fluctuations in the West LA Area. The Wilshire Area reported fluctuations in crime incidents, and the Olympic Area displayed variations. The Newton Area exhibited fluctuations, and the North Hollywood Area showcased variations. The Hollywood Area reported fluctuations, and the Southeast Area displayed variations. The Pacific Area showed fluctuations, the Southwest Area reported fluctuations, and the Central Area exhibited variations. The 77th Street Area showcased fluctuations in reported crime incidents.

The analysis of these crime trends in different areas of Los Angeles can have significant implications for law enforcement resource allocation. Understanding the variations and patterns in crime incidents allows for more targeted and efficient allocation of resources to areas with increasing or consistent crime rates (Baumer et al., 2018). This data-driven approach can aid law enforcement agencies in devising strategies to address specific crime challenges in different regions of the city, ultimately contributing to enhanced public safety.

Impact of the status of a crime report on the clearance rate

The analysis of reported crime incidents in Los Angeles from 2017 to 2023 sheds light on the status of these cases and their impact on the clearance rate. The data primarily categorizes cases into two statuses: "investigation continues" and "adult arrest." The analysis reveals several trends in these two categories. The number of cases with "investigation continues" status displayed a relatively stable trend from 2017 to 2019, with minor fluctuations and a decrease in the number of cases. In 2020, there was a notable drop in cases, signifying potential challenges in case resolution. The number of cases with the status "investigation continues" increased significantly from 2021 to 2022, indicating the accumulation of cases over time. As a measure of case closure, however, the "adult arrest" status showed a declining trend between 2017 and 2022.

This implies a decrease in the quantity of cases that are settled by adult arrests. On the other hand, the 2023 dataset, which includes data up to August 28th, shows a rise in the number of adult arrest cases. This increase might be the result of continued case resolution efforts.

A crucial component of the research question is the examination of how these status categories affect the clearance rate. Recognizing the causes of cases that are postponed or remain unresolved is crucial. These patterns may be caused by variables like the distribution of resources, the intricacy of the case, the methods used in the investigation, and outside influences (Corsaro & Wilson, 2018). Further in-depth analysis is required to identify the specific determinants affecting the clearance rate and the timely resolution of reported crime incidents in Los Angeles (Baradaran-Baughman, 2020). Factors contributing to prolonged investigations and potentially lower clearance rates in criminal cases include the complexity of cases, limited resources, high case volumes, witness cooperation, delays in forensic analysis, legal processes, suspect identification, assessing witness credibility, multi-jurisdictional cases, external pressures, cold cases, and technological challenges (LoFaso, 2020). Efficient resource allocation, improved training, streamlined evidence processing, and case prioritization can help address these challenges and expedite investigations.

Economic Factors, Unemployment and Forecast

In 2020, as the pandemic escalated, Los Angeles, like many other regions, implemented various lockdown measures and restrictions to curb the spread of the virus. These measures, including business closures, stay-at-home orders, and the shutdown of many public activities, had far-reaching economic consequences (Silverio-Murillo et al., 2023). The unemployment rate in Los Angeles dramatically surged to 13.0% in 2020, a level not seen in previous years. Concurrently, the city experienced a significant reduction in reported crime incidents in 2020, with 199,352 cases, which might be partially attributed to the lockdowns. The restrictions on public gatherings and the closure of businesses could have had a suppressive effect on certain types of crimes, such as property crimes and street-level offenses.

Even in 2021, with the pandemic's effects still being felt, the unemployment rate was still very high at 8.8%. Oddly, during this time there was also a rise in reported criminal

incidents (209,265). This seeming contradiction between a high rate of unemployment and an increase in criminal cases highlights the impact of various variables. It suggests that factors like stressors related to the pandemic, civil unrest, changes in law enforcement tactics, and changing community dynamics were important in determining the trends in crime. As 2022 rolled around, the COVID-19 pandemic's effects began to fade, and the jobless rate dropped to 5.0%. Nonetheless, the number of recorded criminal incidents increased, reaching 233,994 cases. This underlines the intricate nature of the relationship between unemployment rates and crime, demonstrating that the pandemic's ramifications extend beyond mere economic considerations (Miller et al., 2020). In the anticipated years 2023–2025, the unemployment rate showed little change, but the rate of crime showed a persistent upward trend. This pattern points to the presence of other contributing factors, which could include things like socioeconomic inequality, changes in policing strategies, and social unrest.

Although the COVID-19 pandemic and related lockdown measures probably had an impact on crime trends in Los Angeles, it is clear that there is much more going on between unemployment and crime rates than a simple cause-and-effect connection (Lasso-de-la-Vega et al., 2021). The pandemic served as a unique and unprecedented backdrop, and to comprehensively grasp the dynamics influencing crime rates during this period, one must consider an array of factors beyond economic conditions (Schleimer et al., 2022). Consequently, further research is imperative to disentangle these intricate interactions, paving the way for informed and targeted interventions aimed at reducing crime in Los Angeles.

CONCLUSION AND RECOMMENDATIONS

Conclusion

A thorough examination of the Los Angeles crime data from 2017 to 2023 has provided insights into the variables affecting crime rates. The study sought to analyze trends in crime rates, pinpoint demographic factors that influence crime, and look into the relationship between COVID-19 and the economy. The study underlined how important it is to identify demographic differences in order to customize interventions for particular communities. The study found distinct patterns in crime rates, including declines in 2018 and 2019, a significant drop in 2020, and a subsequent rise in 2021-2023, coinciding with the COVID-19 pandemic. These disparities among racial and gender groups, showcasing varying trends among ethnic groups and noting that men reported more crimes on average than women, with some variations over time. Thus, the complexity of factors influencing Los Angeles's crime rates—such as the state of the economy, pandemics, and demographics—is highlighted by this analysis. It highlights how crucial it is to take these things into account for a more focused and successful approach to crime prevention and intervention, one that is suited to the varied and ever-changing population of the city.

The investigation ventured into the relationship between the status of crime reports (whether they were under ongoing investigation or had led to adult arrests) and the corresponding clearance rates. It emerged that a significant number of cases were still awaiting resolution, a testament to the intricacies of the legal process. Factors contributing to these delays encompassed a range of issues, from evidence collection to legal procedures and resource allocation. This underscored the pressing need for streamlined investigative procedures and strategic resource allocation to enhance clearance rates. Economic elements, notably the unemployment rate, were scrutinized as potential contributors to shifts in crime trends. The emergence of the COVID-19 pandemic and the ensuing lockdown measures in 2020 and 2021 introduced unique circumstances. Despite a surge in the unemployment rate to unprecedented levels in 2020, reported crime incidents experienced a substantial decrease. In 2021, as the pandemic's economic impact persisted, crime rates began to rise. This intricate interplay between unemployment and crime rates highlighted that economic conditions alone could not comprehensively elucidate the fluctuations in crime trends.

A forecast of expected criminal activity between 2023 and 2025 was part of the project. The projected trend showed that while the unemployment rate remained largely consistent, other factors, including socioeconomic disparities, shifts in law enforcement strategies, and societal unrest, continued to have an effect on crime rates. This study highlights the complexity and diversity of Los Angeles' crime trends, emphasizing the necessity for comprehensive, data-driven strategies to address the intricacies of variations in crime. It also recognizes that, in addition to demographics, social dynamics, and law enforcement tactics, economic factors and the pandemic are just a few of the many influences that must be taken into account in order to inform effective policies and interventions aimed at reducing crime in Los Angeles by 2025.

Recommendations

This in-depth analysis of Los Angeles' crime data raises several significant issues and recommends various crucial policies:

- A multimodal strategy is required because of Los Angeles's erratic crime rate. It is essential to consider the intricate interactions of several factors, such as socioeconomic, demographic, and situational elements, when developing efficient crime reduction strategies.
- It is imperative to customize interventions according to the distinct requirements of diverse ethnic communities and gender categories. Law enforcement and politicians can enhance outcomes by recognizing inequalities and concentrating on community-specific remedies.
- Improving clearance rates and case resolution timeliness can be achieved by optimizing resource allocation and investigative procedures. To speed up investigations, it could be necessary to make investments in people, technology, and training.
- Not only economic considerations but also a wide range of other impacts should be considered by policymakers when determining crime rates. Comprehending the intricate interplay among unemployment, pandemic, societal elements, and law enforcement tactics is vital in formulating all-encompassing crime mitigation measures.

- Given the evolving nature of crime trends, it is essential to maintain vigilant and adaptable monitoring of data to inform proactive responses to emerging challenges.
- Engaging communities in crime prevention efforts, fostering trust between law enforcement and residents, and addressing underlying social issues are critical components of any long-term crime reduction strategy.

Future improvements to Research

Future research can build on this project by extending the study period beyond 2025, incorporating spatial analysis, qualitative research, and causal inference. Comparative studies with other cities, machine learning, and predictive modeling can enhance the accuracy of crime trend forecasts. Analyzing behavioral patterns of criminals and evaluating community-based interventions are key areas for future research. Additionally, assessing policy interventions, leveraging real-time data and technology, and studying long-term socio-economic trends and public health factors can provide a more comprehensive understanding of crime trends.

References

Baradaran Baughman, S. (2020). How effective are police? The problem of clearance rates and criminal accountability. Ala. L. Rev., 72, 47.

Baumer, E. P., Vélez, M. B., & Rosenfeld, R. (2018). Bringing crime trends back into criminology: A critical assessment of the literature and a blueprint for future inquiry. Annual Review of Criminology, 1, 39-61.

Corsaro, N., & Wilson, J. M. (2018). The effects of police contracting on crime: An examination of Compton, California. Journal of Experimental Criminology, 14, 59-81.

Glaeser, E. L., Kim, H., & Luca, M. (2018, May). Nowcasting gentrification: using yelp data to quantify neighborhood change. In AEA Papers and Proceedings (Vol. 108, pp. 77-82). 2014 Broadway, Suite 305, Nashville, TN 37203: American Economic Association.

Lasso de la Vega, C., Volij, O., & Weinschelbaum, F. (2021). Can more police induce more crime?. Available at SSRN 3931440.

LoFaso, C. A. (2020). Solving homicides: The influence of neighborhood characteristics and investigator caseload. Criminal justice review, 45(1), 84-103.

Lopez, E., & Rosenfeld, R. (2021). Crime, quarantine, and the US coronavirus pandemic. Criminology & Public Policy, 20(3), 401-422.

Miller, A. R., Segal, C., & Spencer, M. K. (2020). Effects of the COVID-19 pandemic on domestic violence in Los Angeles (No. w28068). National Bureau of Economic Research.

Piquero, A. R., Jennings, W. G., Jemison, E., Kaukinen, C., & Knaul, F. M. (2021). Domestic violence during the COVID-19 pandemic-Evidence from a systematic review and meta-analysis. Journal of criminal justice, 74, 101806.

Schleimer, J. P., Pear, V. A., McCort, C. D., Shev, A. B., De Biasi, A., Tomsich, E., ... & Wintemute, G. J. (2022). Unemployment and crime in US cities during the coronavirus pandemic. Journal of urban health, 99(1), 82-91.

Silverio-Murillo, A., de la Miyar, J. B., & Hoehn-Velasco, L. (2023). Families under Confinement: COVID-19 and Domestic Violence. In Crime and Social Control in Pandemic Times (Vol. 28, pp. 23-41). Emerald Publishing Limited.