

# DAS732: Data Visualization Assignment 1

Subham Agarwala  
IMT2022110  
IIIT Bangalore  
Bangalore, India  
Subham.Agarwal@iiitb.ac.in

Sarvesh Kumar  
IMT2022521  
IIIT Bangalore  
Bangalore, India  
SarveshKumar.A@iiitb.ac.in

Ayush Gupta  
IMT2022546  
IIIT Bangalore  
Bangalore, India  
Ayush.Gupta546@iiitb.ac.in

## I. INTRODUCTION

This report is a compilation of our work on the "USA Big City Crime Data"[1] dataset, which includes data visualizations and inferences to answer the question "*What are the major crime trends in big cities, and how do factors such as location, type of crime, and demographic details influence these trends?*". We have divided our analysis into three tasks which are:

- **T1:** Analysis of crime trends based on location
- **T2:** Analysis of crime trends based on the type of crime and frequency
- **T3:** Analysis of crime trends based on victim demographics

## II. PREPROCESSING OF DATASET

Before performing any visualizations or analysis following are the steps we took to preprocess the dataset:

- We dropped the dataset for Chicago as it did not provide any more information than just the location of the crime and the crime type.
- Removed duplicate entries
- Removed rows with negative victim age
- Grouped null values for victim sex and victim descent column into a single category "X" to represent unknown values.
- Removed "DR\_NO", "Location", "CrossStreet" and "Mocodes" due to lack of useful information.
- Removed rows with Longitude value as "0" as it does not belong to LA.
- Removed "CrmCd1" as it was the same as "CrmCd".
- Computed the following new fields to ease and better the visualizations:
  - "AvgLong" and "AvgLat": grouped latitudes and longitudes by area bu computing their avg."
  - "Hour of the Day": extracted the hour from the "Time Occ" column
  - "Arrest\_Cat": specifies whether the criminal was arrested or not. Extracted from "status" column.
  - "Modified\_Part1-2": converted values to string.
  - "Modified\_Weapon\_Desc": replaces null values in the "Weapon Desc" column.
  - "Time\_Diff\_Occ\_Repo": The difference in number of days between "Date Occ" and "Date Rptd" columns

- "Criminal\_Age\_Category": classifies arrested criminals into adult and juvenile. Extracted from "Status" column.
- "Descent Name": expands the labels given in "Vict Descent" column to their actual meaning.

## III. TASKS

### A. *Task 1: Analysis of crime trends based on location*

The goal of this task is to analyze the distribution of crimes across different geographic areas in the dataset to identify which areas experience higher crime rates and potential crime hotspots. This analysis aims to provide insights into how crime is concentrated across different parts of the city and to help law enforcement or policymakers focus resources on the areas most affected.



Fig. 1. Geospatial map that shows the crime hotspots in LA.

First we plot the data on a geospatial map to identify crime hotspots in LA (Fig. 1). From the visualization we can see that areas like Central, 77th Street, Pacific and Southwest have the highest crime count among the rest of the areas in LA with Central having the highest crime count. Foothill and Hollenbeck seem to have the lowest crime count.

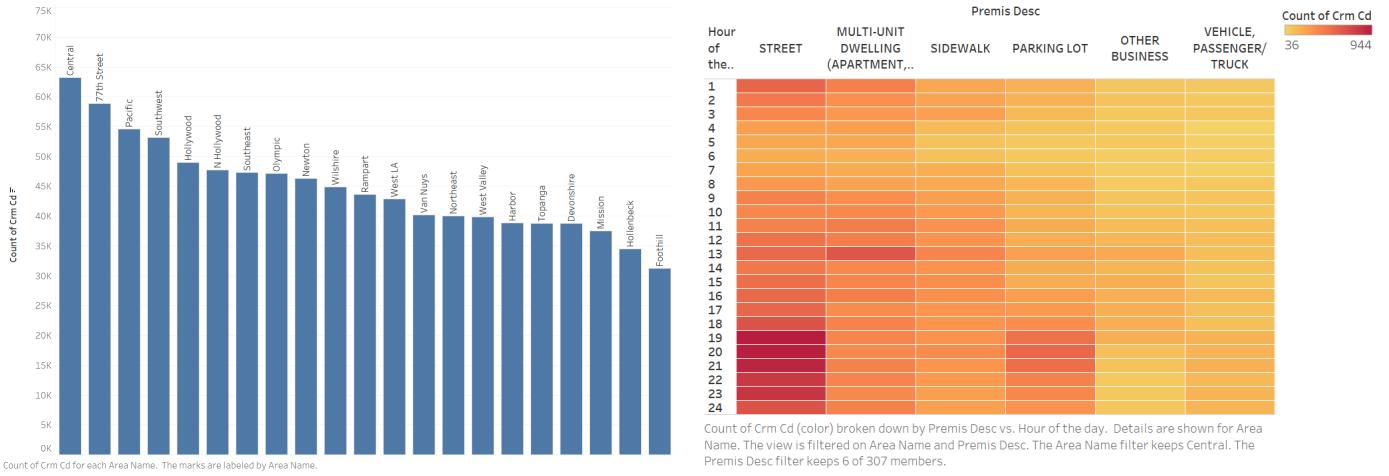


Fig. 2. Bar graph of crime count in the respective areas.

Fig. 2 gives us a better understanding of the crime count in the different areas of LA.

Now that we have identified that the Central area of LA is the biggest hotspot for crimes in the city, we will look deeper into the specifics of the crimes taking place there. We will now explore the premises in the Central region of LA where major volume of crimes take place.

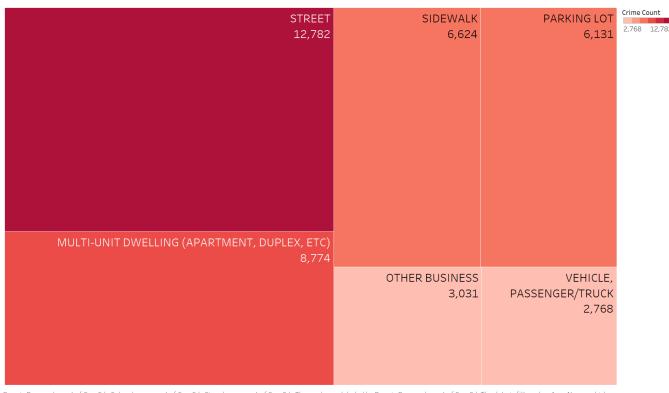


Fig. 3. Premises in Central region with maximum crime counts.

Fig. 3 shows the most common premises for crime occurrences in Central LA. As depicted, the majority of crimes take place on the street, which suggests that public spaces, where individuals are more exposed, are particularly vulnerable to criminal activity. Apartments/Duplexes rank as the second most frequent crime scene, indicating that residential areas are also significant targets for crime. This may highlight issues related to property crimes, domestic incidents, or security vulnerabilities in housing complexes.

To observe the hours of the day in which maximum crimes take place in the above premises in Central LA we have visualized the data in the form of a heat map (Fig. 4).

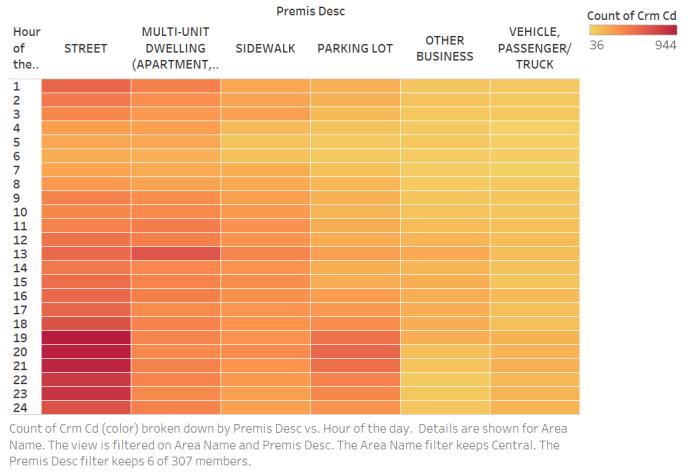


Fig. 4. Heat map showing density of crimes taking place at different hours of the day in Central LA.

We see that the maximum number of crimes in Central LA in all premises take place during the later hours of the day. This pattern is consistent across most premises, indicating that evening and nighttime are the most vulnerable periods for criminal activity in Central LA. This could be due to reduced visibility, fewer people in public spaces, or a higher likelihood of individuals being outside during these hours.

For the "Other Business" category, we observe a distinct peak in crime incidents between the 13th and 19th hour (1 PM to 7 PM). This could suggest that certain types of businesses, such as retail stores, restaurants, or entertainment venues, may experience more crime during business hours, possibly linked to theft, robbery, or other opportunistic crimes.

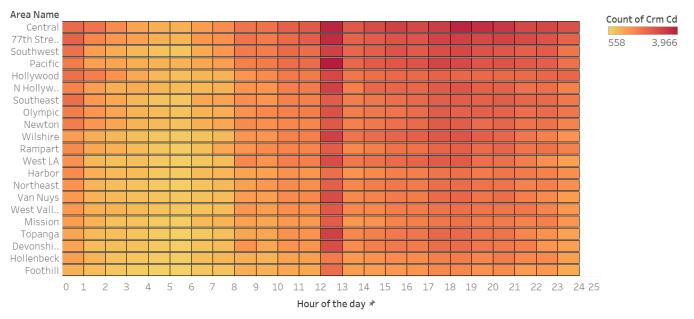


Fig. 5. Heat map illustrating density of crimes taking place at different hours of the day in LA.

Fig. 5 displays the distribution of crime incidents across different hours of the day throughout Los Angeles. The data reveals a similar pattern to that observed in Fig. 4 for Central LA, where the highest crime activity occurs during the evening and nighttime hours. However, we see an unusual spike in crime count during the first hour (12 AM) and the 12th hour (12 PM). This anomaly is present across all areas of LA and likely stems from 00:00 and 12:00 being used as

default placeholders for unknown or missing time values in the dataset, skewing the crime count at these times.

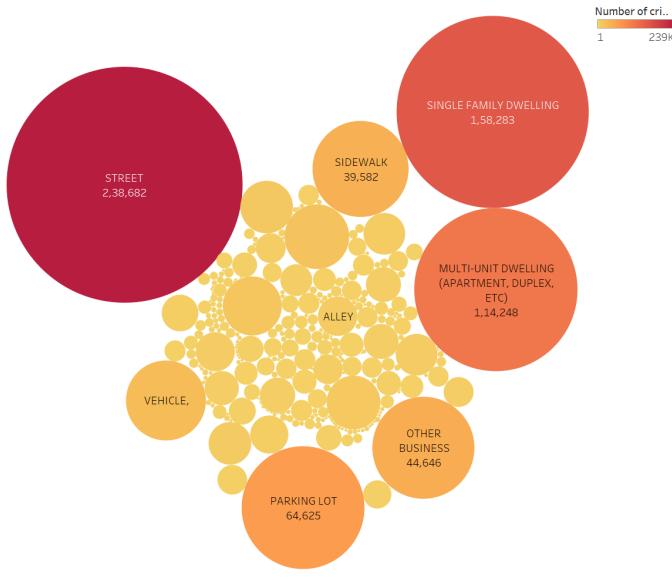


Fig. 6. Crime count by premises across all of LA

When analyzing crime count by premises across all of Los Angeles in Fig. 6, we observe a similar pattern to that seen in Central LA, with the street remaining the most frequent location for criminal incidents. However, in contrast to Central LA, single-family dwellings emerge as the second most frequent premises for crime in the broader LA region. This shift highlights the vulnerability of residential areas across the city.

The prevalence of crime in single-family dwellings suggests that crimes such as burglary, theft, and domestic violence may be more common in residential neighborhoods than in apartment complexes or multi-family housing units, which were more prominent in Central LA. Streets and public spaces continue to be hotspots, indicating that crimes in public or semi-public areas (e.g., parking lots, sidewalks) are still a significant concern.

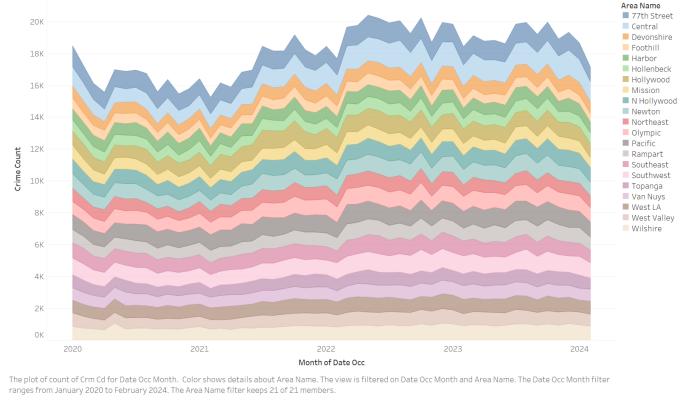


Fig. 7. Area chart showing crime count by area in LA over time

Fig. 7 is an area chart illustrating the trend in crime counts across different regions of Los Angeles over the years. The chart provides a visual representation of how the number of crimes has fluctuated or remained stable in various areas.

The most striking observation from the chart is the significant increase in crime count in Central LA. This region shows a notable and steady rise in criminal incidents, shown more clearly in Fig. 8, setting it apart from other areas in Los Angeles. In contrast, other regions either show a slight increase in crime over time or maintain a relatively constant rate. These areas appear to be less affected by the broader trends influencing crime in Central LA.

Fig. 8 effectively highlights the dominant contribution of Central LA to the overall crime increase in the city, suggesting that Central LA may be experiencing unique social or economic factors that are driving this upward trend.

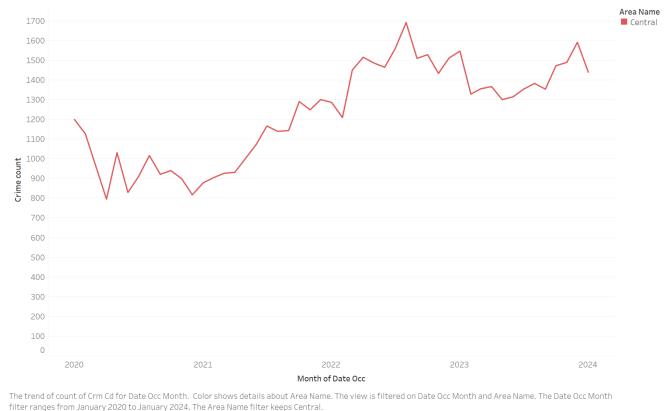


Fig. 8. Line chart showing the increase in crime in Central LA over the years

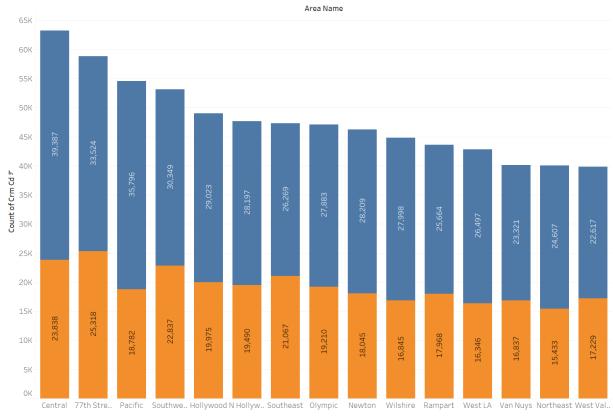


Fig. 9. Distribution of part 1 and part 2 crimes across top areas in LA with respect to crime count.

Crimes in the dataset are categorised into part 1 and part 2 crimes. Fig. 9 presents the distribution of Part 1 and Part 2 crimes for the top 15 areas in Los Angeles based on total crime count. The visualization highlights a clear trend: in all of these high-crime areas, Part 1 crimes are more prevalent than Part 2 crimes.

Part 1 crimes typically include serious offenses such as homicide, robbery, assault, and burglary, while Part 2 crimes encompass less severe offenses, like vandalism, public intoxication, and disorderly conduct. The fact that Part 1 crimes consistently outnumber Part 2 crimes across all these areas points to the gravity of criminal activity in these regions.

The areas with the highest crime count show particularly large disparities between Part 1 and Part 2 crimes, suggesting that these regions may be hotspots for violent or serious criminal offenses.

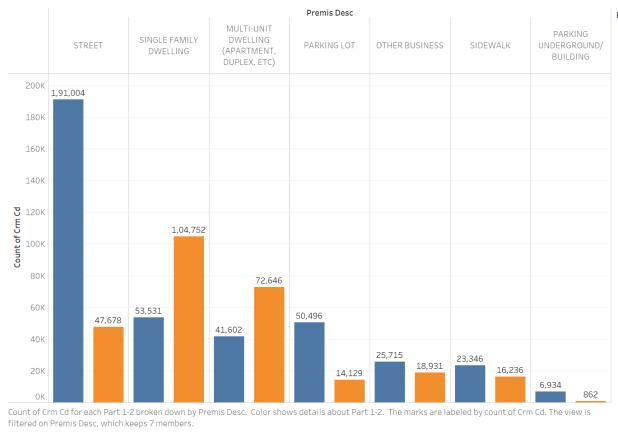


Fig. 10. Distribution of part 1 and part 2 crimes in top premises in LA with respect to crime count

Now that we have seen the distribution of Part 1 and Part 2 crimes across different areas in LA, let's examine their distribution across the top premises in LA.

In Fig. 10, we observe that streets, parking lots, businesses, and sidewalks have a significantly higher number of Part 1 crimes compared to Part 2 crimes. This aligns with the nature of these locations, where offenses like robbery, assault, and burglary (classified under Part 1 crimes) are more likely to occur.

However, we see a contrasting trend when it comes to single-family dwellings and multi-unit dwellings. Residential areas show a higher concentration of less severe crimes (Part 2), possibly reflecting neighborhood disputes, minor offenses, or violations that are less violent in nature.

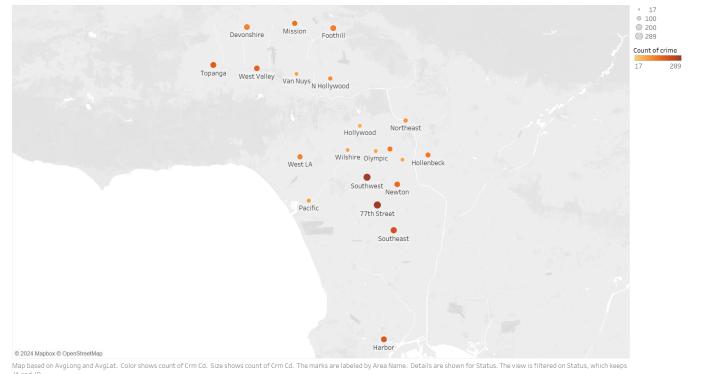


Fig. 11. Geographical plot showing number of juvenile criminals in the different areas of LA.

We also utilized the "Status" column of the dataset, which provides information about actions taken against offenders, to gain insight into the number of juvenile criminals in different areas of Los Angeles. By filtering the data for statuses related to juvenile actions, we were able to estimate the number of juvenile offenders in each area, as depicted in Fig. 11.

Despite having the highest overall crime count, the Central LA area shows a very low number of juvenile criminals. This may indicate that the crimes committed in Central LA are more likely to involve adult offenders or that youth-related criminal activity is minimal in this area. The low juvenile crime rate in Central LA, despite its high overall crime rate, suggests that crime prevention measures for juveniles might already be effective or that adult-related crimes are the primary concern in this region. On the other hand, areas such as 77th Street and Southeast LA report the highest numbers of juvenile offenders. These regions appear to have a greater concentration of youth involvement in criminal activities compared to other areas.

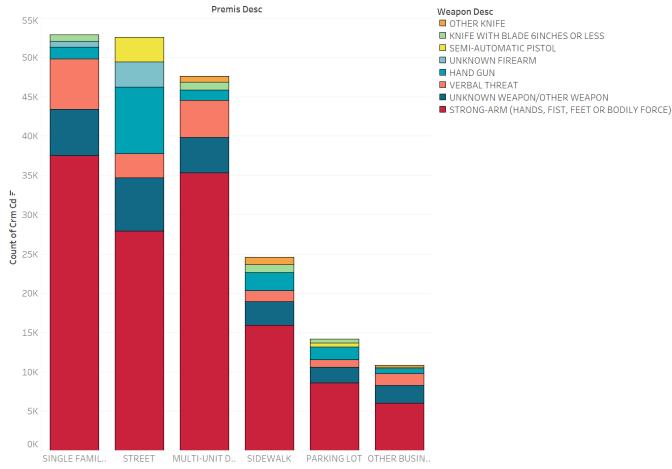


Fig. 12. Top weapon used in the premises with maximum number of crimes.

Our final crime analysis focuses on the top weapons used in the premises with the highest crime counts across Los Angeles. As shown in Fig. 12, we observe distinct patterns in weapon usage based on the type of location.

Strong-arm force (which includes hands, fists, feet, or bodily force) is the most commonly used method to commit crimes across all premises, indicating a high prevalence of physical assault in both public and residential areas. In residential areas like single-family dwellings and multi-unit dwellings, verbal threats rank as the second most common method used, followed by the use of handguns and unknown firearms. This suggests that while physical force dominates, verbal intimidation and firearms also play a role in residential crimes. On streets, handguns are used more frequently than verbal threats, indicating a higher tendency for gun-related crimes in public areas. This is followed by the use of unknown firearms and semi-automatic pistols, showing that gun violence is a major concern in these locations. In sidewalks and parking lots, aside from strong-arm tactics, knives with blades 6 inches or less are also commonly used, adding a layer of concern for weapon-related incidents in these public spaces.

#### B. Task 2: Analysis of crime trends based on the type of crime and frequency

The goal of this task is to analyze the distribution of crimes by type and frequency within Los Angeles. This analysis aims to identify which types of crimes are most prevalent and how their frequency varies across different time periods. By examining these patterns, we seek to gain insights into the predominant crime issues affecting the city and understand trends that may inform targeted interventions.

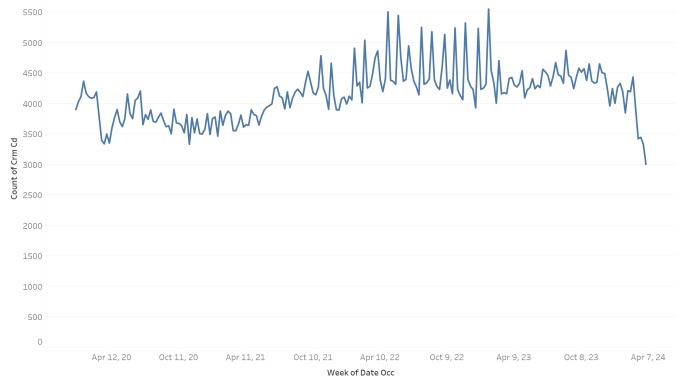


Fig. 13. Weekly Crime Trends in Los Angeles

The time series graph [Fig. 13] of weekly crime data in Los Angeles reveals a notable spike in criminal activity during 2022 and 2023. This significant rise may be linked to various factors such as economic changes, shifts in policing, or other influences. The data underscores the need to investigate the causes behind this surge and to develop targeted strategies to address the increase in crime. Notably, the graph shows a dip in crime rates after 2023, which could be due to incomplete data for 2024.

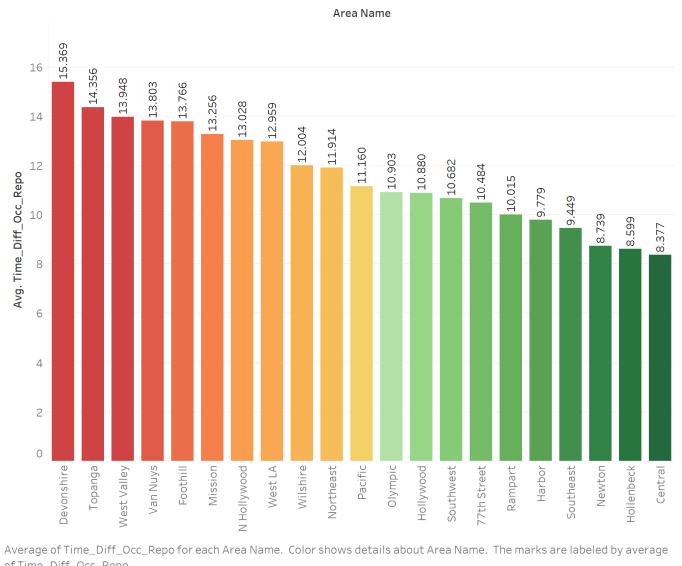


Fig. 14. Average Crime Reporting Delays Across LA Regions

[Fig. 14] shows that crime reporting times vary significantly across different regions of Los Angeles. Central, Hollenbeck, and Newton report crimes the quickest, with an average delay of only 8 days. In contrast, Devonshire and Topanga exhibit the longest delays, averaging over two weeks. These differences may indicate variations in community responsiveness, police efficiency, or local infrastructure. The disparity in reporting times raises important questions about the underlying causes

and potential interventions to improve crime reporting efficiency in slower regions.

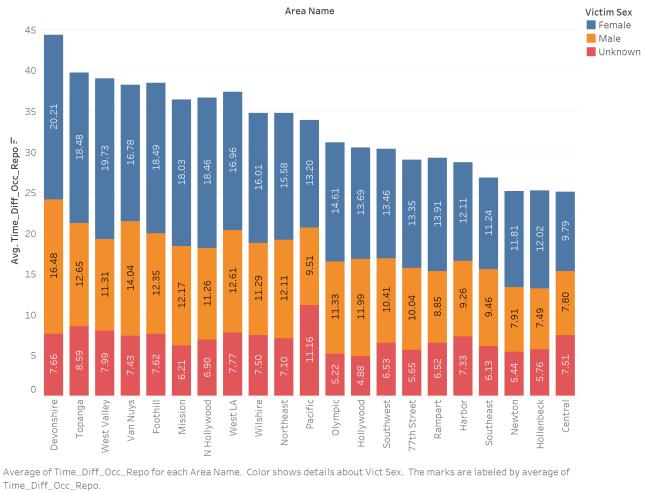


Fig. 15. Gender Disparity in Crime Reporting Across LA Regions

[Fig. 15] highlights the average number of days to report a crime across LA regions, segmented by gender. On average, crimes with male victims are reported 4-5 days earlier than those involving female victims. This trend is consistent across regions, with Hollenbeck, and Newton having the shortest delays, but still showing a notable gender gap in reporting times.

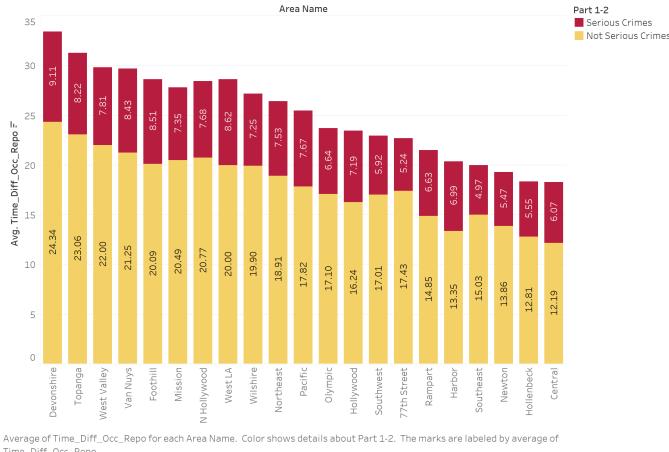
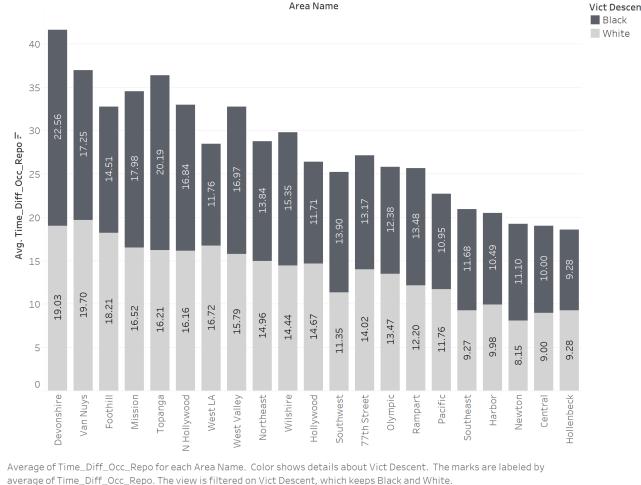


Fig. 16. Reporting Time Delays by Crime Seriousness across LA Regions

[Fig. 16] shows the average number of days to report a crime in LA, categorized by crime seriousness. It reveals that serious crimes are reported over a week faster than less serious ones, demonstrating a prompt response to urgent cases. This insight can be used to evaluate the effectiveness of crime reporting and response mechanisms across different regions.

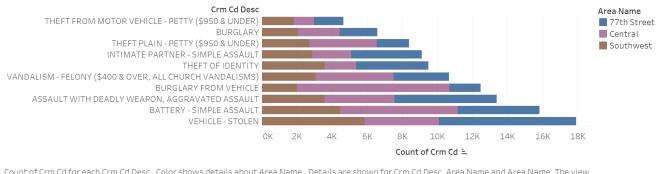




Vict Sex and Cat. Time. Occ. Color shows count of Crm Cd. Size shows count of Crm Cd. The marks are labeled by Vict Sex and Cat. Time. Occ. The view is filtered on Vict Sex, which keeps Female and Male.

Fig. 19. Crime Distribution by Hour and Gender

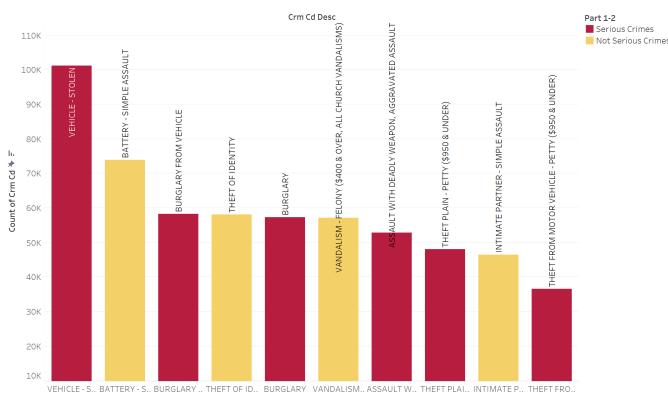
[Fig. 19] presented as a tree map, shows the distribution of crimes by hour of the day, segmented by gender. The data reveals a notable spike in crime reports at 12:00, which may indicate potential data inaccuracies or a default time entry. Overall, the count of crimes is nearly equal between men and women, with higher incidents occurring during late hours and fewer during earlier times of the day.



Count of Crm Cd for each Crm Cd Desc. Color shows details about Area Name. Details are shown for Crm Cd Desc, Area Name and Area Name. The view is filtered on Crm Cd Desc and Area Name. The Crm Cd Desc filter has multiple members selected. The Area Name filter keeps 77th Street, Central and Southwest.

Fig. 21. Crime Trends in Violent LA Areas

[Fig. 21] highlights the most common types of crimes in the violent areas of LA—Central, 77th Street, and Southwest. It shows that Burglary from Vehicle is notably prevalent in Central, while Vehicle Stolen is most common in 77th Street. The data provides insights into the specific crime patterns in these high-violence areas, which can be crucial for targeted policing and crime prevention strategies.



Count of Crm Cd for each Crm Cd Desc. Color shows details about Part 1-2. The marks are labeled by Crm Cd Desc. Details are shown for Crm Cd Desc. The view is filtered on Crm Cd Desc, which has multiple members selected.

Fig. 20. Top 10 Most Common Crimes in LA

[Fig. 20] displays the top 10 most common crimes in LA, categorized by crime seriousness. Vehicle theft emerges as the most frequent and serious crime. This visualization is useful for identifying prevalent crime types and their severity, aiding in resource allocation and targeted interventions to address the most critical issues.

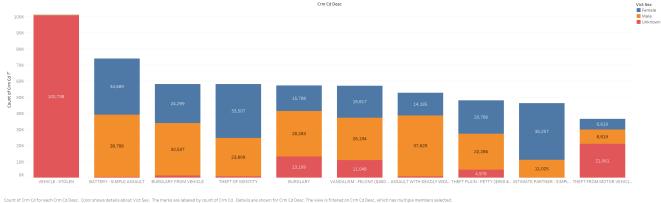


Fig. 22. Gender-Specific Crime Patterns in LA

[Fig. 22] illustrates the top 10 crimes in LA segmented by gender. The data reveals that the victim's sex for Vehicle Stolen is often unknown. Crimes like Intimate Partner Violence and Theft of Identity predominantly affect females, while Assault with Deadly Weapon primarily impacts males.

This visualization is useful for identifying gender-specific crime patterns and understanding which crimes disproportionately affect each gender.

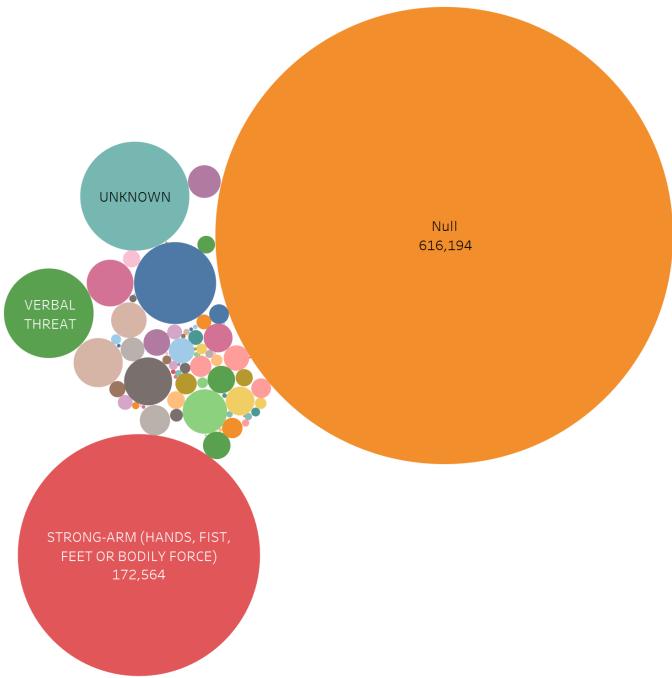


Fig. 23. Weapon Usage in LA Crimes

[Fig. 23] highlights the most common weapons used in crimes across LA. The data shows that most crimes involve no weapon, while in cases where a weapon is used, strong arms, body, or fists are the most frequent.

This visualization is useful for understanding the nature of physical violence in crimes.

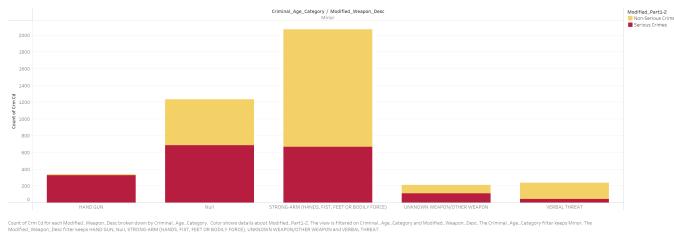


Fig. 24. Weapons Used by Minors in LA Crimes

[Fig. 24] showcases the most common weapons used by minors arrested for crimes in LA, segmented by the seriousness of the offense. The data reveals that minors frequently rely on strong arms and fists, but serious crimes have also been committed using handguns, fists, and even no weapons at all.

This visualization is useful for understanding the role of minors in both non-lethal and serious crimes, aiding in the development of focused intervention strategies for youth crime prevention and rehabilitation.

In addition to minors, I visualised that adults arrested in LA display similar patterns in weapon usage. However, adults

commit crimes at a significantly larger volume, approximately 30-35 times more than minors.

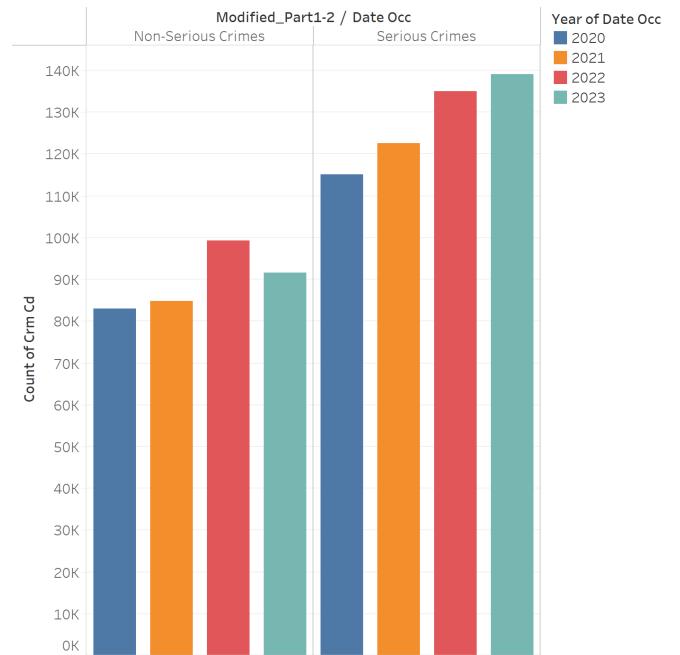


Fig. 25. Yearly Crime Trends in LA (2020-2023) Categorized by Crime Seriousness

[Fig. 25] illustrates the trends in crime seriousness from 2020 to 2023 on a yearly basis. Serious crimes have consistently increased year by year, while non-serious crimes rose until 2022, then saw a decline in 2023. The data for 2024 is excluded due to incomplete months.

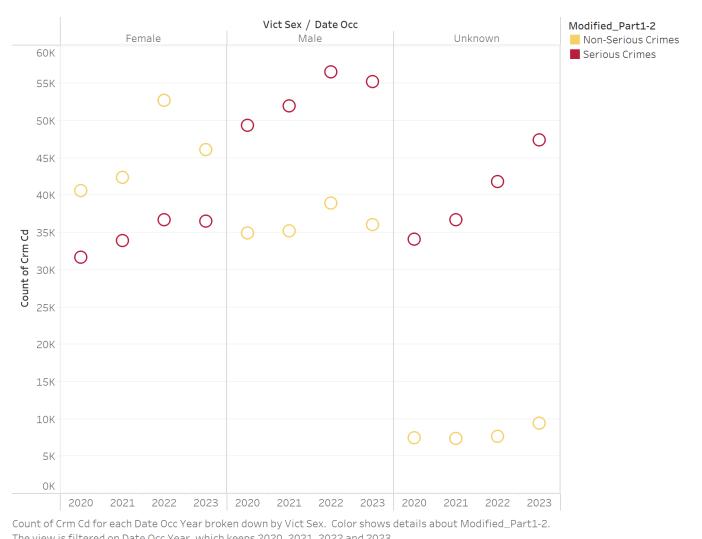


Fig. 26. Gender-Specific Trends in Crime Seriousness (2020-2023)

[Fig. 26] shows the yearly trends in crime seriousness from 2020 to 2023, segmented by victim sex. The data reveals that more male victims are affected by serious crimes than females, while non-serious crimes are more common among female victims. For males, serious crimes outnumber non-serious ones, whereas non-serious crimes on females peaked in 2022.

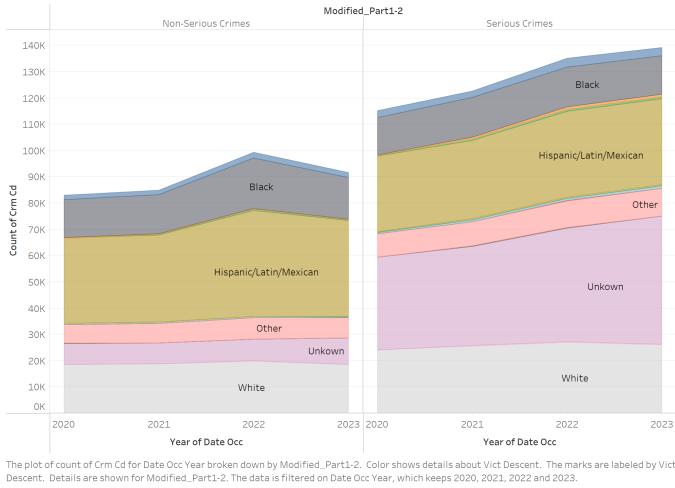


Fig. 27. Yearly Trends in Crime Seriousness by Victim Descent (2020-2023)

[Fig. 27] displays crime seriousness from 2020 to 2023, segmented by victim descent. Hispanic/Latin/Mexican victims are disproportionately affected by both serious and non-serious crimes compared to other descent groups. This pattern persists across the years, highlighting the vulnerability of this community.

The above visualization is useful for recognizing descent-based crime patterns, aiding in the development of targeted crime prevention and victim support initiatives for affected groups.

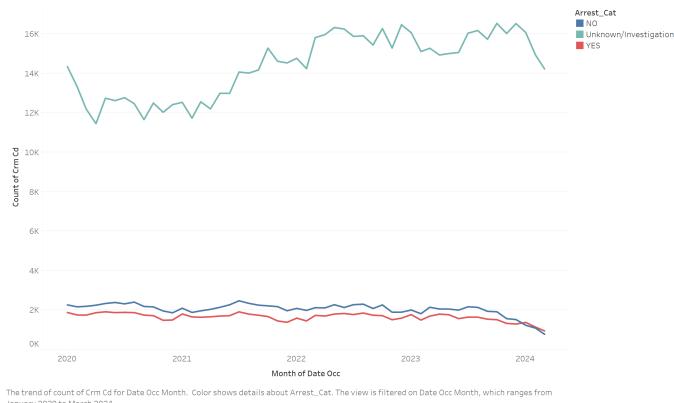


Fig. 28. Monthly Crime and Arrest Statistics

[Fig. 28] tracks the number of crimes and the status of criminals from January 2020 to March 2024 on a monthly

basis, highlighting the breakdown of cases where arrests have been made, where investigations are ongoing, and where no action is recorded. The data shows that approximately 2,000 criminals have been arrested monthly, with a similar number still not arrested. A significant proportion of cases remain under investigation or lack clarity on the action taken.

This visualization is crucial for understanding the dynamics of crime resolution and identifying areas where investigative processes may be lagging.

Delaying case resolutions can lead to prolonged risks for public safety, as unresolved cases might continue to pose threats. Efficient case management is essential for maintaining trust in the criminal justice system and ensuring that offenders are held accountable promptly.

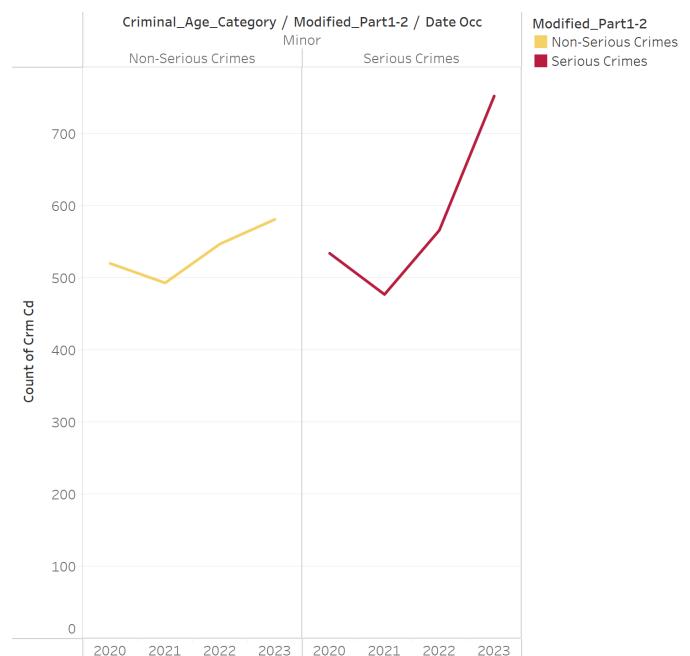


Fig. 29. Trend in Seriousness Crimes Committed by Minors(2020-2023)

[Fig. 29] displays the trend in the seriousness of crimes committed by minors from 2020 to 2023 on a yearly basis. The data indicates that serious crimes committed by minors hit their lowest point in 2021 but have been on the rise since then, with a notably steeper increase compared to non-serious crimes, which also show a rising trend.

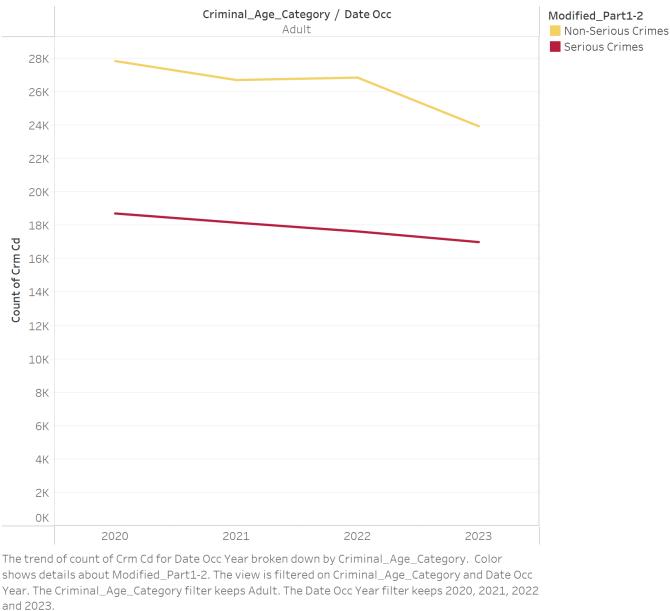


Fig. 30. Yearly Trends in Crime Seriousness Committed by Adults (2020-2023)

[Fig. 30] tracks the trend in the seriousness of crimes committed by adults from 2020 to 2023 on a yearly basis. It reveals that non-serious crimes significantly outnumber serious crimes committed by adults. Additionally, there is a noticeable decline in the number of both serious and non-serious crimes committed by adults over the years.

The series of visualizations and plots presented offer a comprehensive analysis of crime trends in Los Angeles, highlighting key insights into the seriousness of crimes, reporting patterns, weapons used and demographic impacts.

1. Reporting Timeliness: The analysis reveals that crimes are generally reported quickly, with most reports made within the first week. However, some cases experience significant delays, which can skew averages and impact the perception of crime trends.

2. Crime Seriousness and Demographics: Serious crimes are observed to affect males more significantly than females, with non-serious crimes showing a higher prevalence among females. Notably, Hispanic/Latin/Mexican victims are disproportionately impacted by both serious and non-serious crimes.

3. Weapon Usage: Most crimes involve no weapon, but when weapons are used, strong arms, body, or fists are the most common. This pattern is consistent among both minors and adults, although minors tend to use these methods more frequently in serious crimes.

4. Juvenile vs. Adult Crime Trends: The data indicates a troubling increase in serious crimes committed by minors since 2021, contrasted with a decrease in both serious and non-serious crimes committed by adults. This suggests a shift in crime patterns that requires targeted intervention strategies.

5. Crime Resolution and Investigation: There is a significant number of cases with ongoing investigations or unknown ac-

tion taken against criminals. The consistent number of arrests and unarrested criminals each month underscores the need for more effective case management and resolution strategies.

Overall, these insights are essential for understanding crime dynamics and developing strategies for crime prevention, victim support, and more effective law enforcement. Addressing the issues identified, such as the rise in serious crimes among minors and the need for better data on crime resolution, will be crucial for improving community safety and justice outcomes.

### C. Task 3: Victim Demographics Analysis

The goal of this task is to analyze the demographics of Los Angeles crime victims from 2000 to May 2024, focusing on age, gender, and descent. By examining these attributes, we aim to uncover trends in victimization across different demographic groups and how they are affected by various crimes. This section presents visualizations and statistics to highlight trends and correlations between demographics and crime types.

**1) Victim Age** *Analysing trends based on age of the victims:* We begin by analysing the correlation between victim age and the crimes committed. The correlation is not fully explored in this section and is further covered while looking at victim sex and descent.

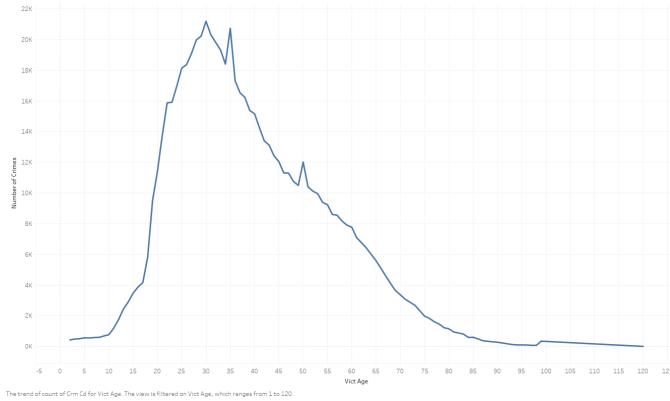


Fig. 31. Crime Victims by Age

Figure 31 is a line chart depicts the number of crimes in relation to the age of victims, revealing trends in victimization across various age groups. The sudden peaks at age 35 and 50 could be due to inaccurate reporting.

Crm Cd Desc	Vict Age (bin)											
	7.4	14.8	22.2	29.6	37.0	44.4	51.8	59.2	66.6	74.0	81.4	88.8
BATTERY - SIMPLE ASSAULT	1,799	7,920	12,021	11,945	10,891	8,317	9,186	6,915	2,974	1,286	437	127
THEFT OF IDENTITY	62	2,624	10,034	12,921	10,933	6,437	5,968	3,801	2,245	1,327	475	159
BURGLARY FROM VEHICLE	12	3,527	13,531	13,597	10,098	6,862	5,043	2,684	1,337	563	135	26
ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT	696	6,424	9,596	9,419	8,401	6,654	5,382	3,143	1,267	486	158	39
INTIMATE PARTNER - SIMPLE ASSAULT	50	5,327	12,079	11,279	8,359	4,353	2,699	1,130	378	168	52	15
VANDALISM - FELONY (\$400 & OVER, ALL CHURCH VANDALISMS)	20	2,925	7,451	8,658	7,655	5,512	5,089	3,142	1,566	741	206	52
THEFT PLAIN - PETTY (\$950 & UNDER)	151	3,826	8,214	8,693	7,020	4,282	4,121	2,615	1,602	889	287	110
BURGLARY	25	1,957	4,131	6,622	6,800	5,288	5,288	3,953	2,745	1,703	646	258
THEFT FROM MOTOR VEHICLE - GRAND (\$950.01 AND OVER)	7	1,402	5,230	6,079	5,705	4,210	4,179	2,713	1,895	731	144	29
THEFT FROM MOTOR VEHICLE - PETTY (\$950 & UNDER)	4	877	3,004	2,957	2,630	1,864	1,746	1,090	634	343	103	29

Count of Crm Cd broken down by Vict Age (bin) vs. Crm Cd Desc. Colour shows count of Crm Cd. The marks are labelled by count of Crm Cd. The view is filtered on Vict Age (bin) and Crm Cd Desc. The Vict Age (bin) filter excludes 0.0 and 128.4. The Crm Cd Desc filter keeps 20 of 329 members.

Fig. 32. Heatmap of Age and Crime Types

Figure 32 is a heatmap which allows us to understand the relation between age and victim for 10 of the most frequent crimes. We can see for certain crimes such as "Intimate Partner - Sexual Assault", there is a highly concentrated victim age while for some such as "Burglary" it is spread out.

**2) Victim Sex** *Analysing trends based on sex of the victims:* Now that we have looked at how age influences various aspects of the crime, lets look at gender. As a general legend for all plots, "M" label stands for male, "F" stands for female and "X" is unknown.

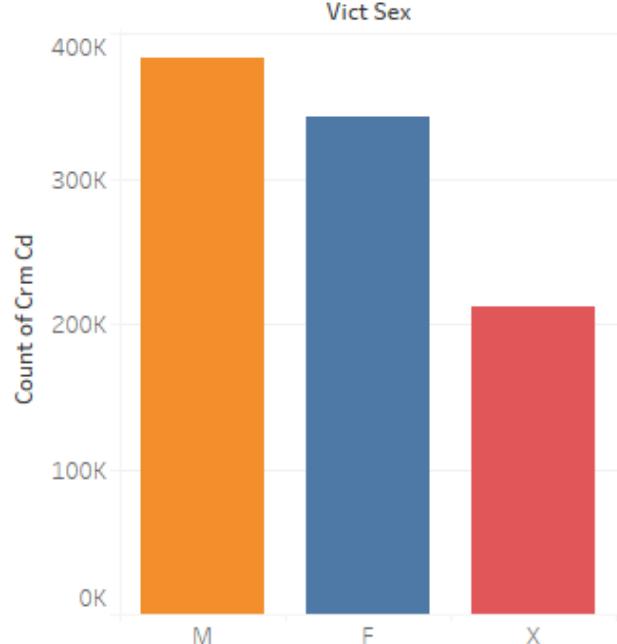


Fig. 33. Number of Crimes by Gender

In Fig. 33, the bar chart displays the number of crimes committed, categorized by the gender of the victims. Each bar

represents the total number of crimes associated with male, female, and unknown gender victims. We can see there were more crimes committed against Males than Females. Also we can see that there is a significant number unknown values which will need to be excluded for further visualisations.

Let us now look at the ratios of male to female crime in various areas. In many of the areas the ratio is almost 1:1 these points are excluded and we'll look at points with a measurable difference in ratio.

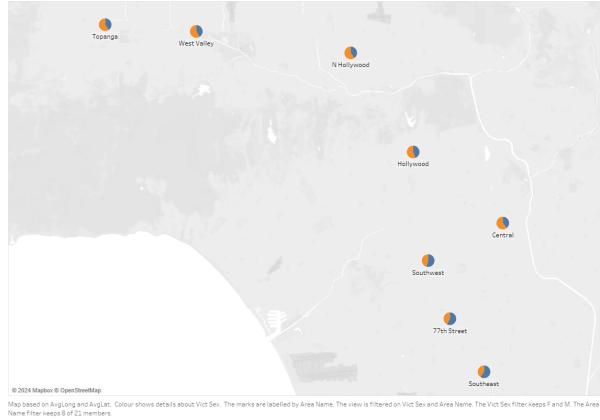


Fig. 34. Map with pie charts showing ratio of male and female victims

We see that there areas leaning towards both sides however, sides with more crimes against females are more significant as the overall dataset has more males. There can be many reasons for this imbalance between areas. This will depend on the factors that affect Male and Female crimes. We will now visualise other fields to find such factors.

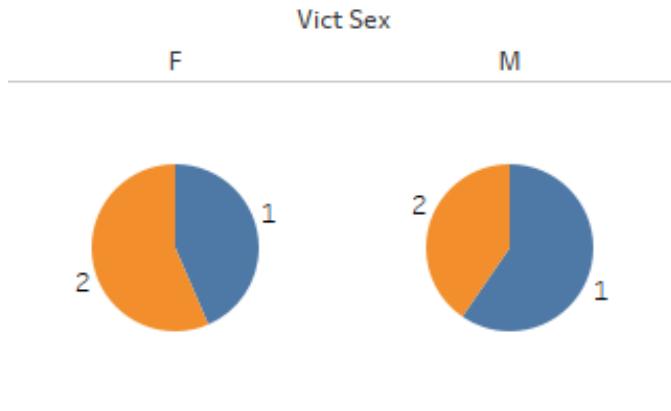


Fig. 35. Part of Crime by Gender

Fig. 35 is a pie chart displaying how many crimes of part 1 and part 2 are committed against Males and Females. Part 1 crimes refer to more serious and violent crimes. Part 2 crimes are less serious crimes as explained earlier. We see that males face a majority of serious crimes while its the opposite for

females. If we look at the Figure 9, we see that the areas with more female victims had more part 2 crimes. Now we can look at what part 1 and part 2 crimes are faced by males and females.

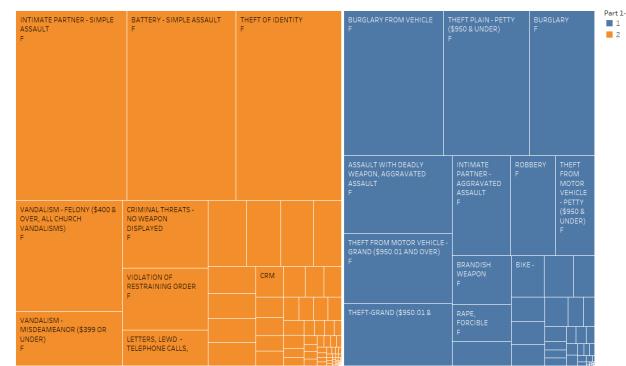


Fig. 36. Crimes Faced by Females in Part 1 and Part 2

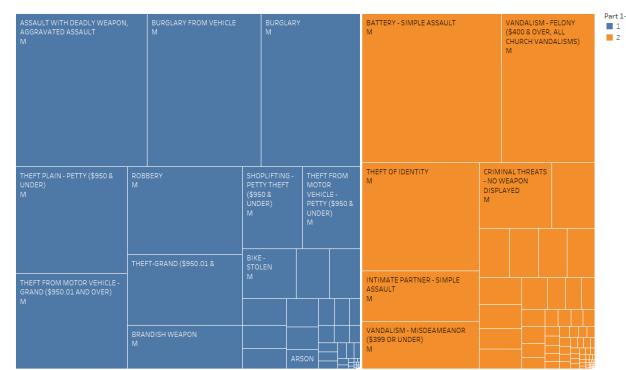


Fig. 37. Crimes Faced by Males in Part 1 and Part 2

We get some very interesting details from these two charts. We can see that females face more "Simple Assaults" Which refers to just threatening instead of a physical altercation. Males face more violent crimes such as Aggravated Assault.

We also see crimes that have a huge frequency in females but are barely present in the other. Such as various "Intimate Partner" crimes and "Rape - forcible".



Fig. 38. Gender of Victims over Time

The trend of count of Crm Cd for Date Occ Month. Colour shows details about Vict Sex. The view is filtered on Vict Sex and Date Occ Month. The Vict Sex filter keeps F and M. The Date Occ Month filter ranges from January 2020 to February 2024.

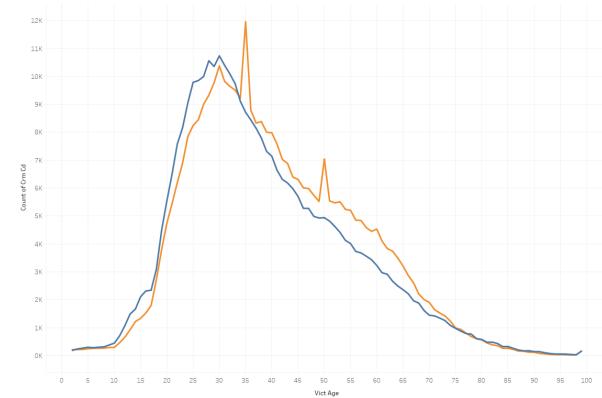


Fig. 40. Age of Victims by Gender

We can now analyse the trend of crime against both genders from 2020 to 2024. We see a dip in both during 2020. This can be reasoned by the pandemic. We then see an rise in both in 2022 although the rise is more in females. We can try to see why this unequal rise happens.



Fig. 39. Gender of Victims by Part over Time

The trend of count of Crm Cd for Date Occ Month. Colour shows details about Vict Sex. The marks are labelled by Part 1-2. The view is filtered on Vict Sex and Date Occ Month. The Vict Sex filter keeps F and M. The Date Occ Month filter ranges from January 2020 to February 2024. The view is highlighted where Part 1-2 contains "2".

In figure 39, we see that in the time when the spike happened, part 1 crimes against females remained stable but there was a huge spike in number of part 2 crimes. This spike is observed at a smaller level in males too. This could be the reason for the bigger rise in females after the dip.

In figure 40 we see the relation between age and gender of victims. We see that there's actually more female victims than male up to the age of 33.

We also see spikes in males at the ages of 35 and 50 which corresponds to the spikes in the combined age graph. We see that there are so many more male victims over the age of 33, they become greater in number overall.

### 3) Victim Descent Trends based on descent of the victims:

In this section we look at the cultural origin of the victims. This data could be helpful in seeing if certain communities are being targeted and what communities need to be protected better. We first look at the communities we have in the dataset. The word "Community" has been used to refer to the values of Descent.

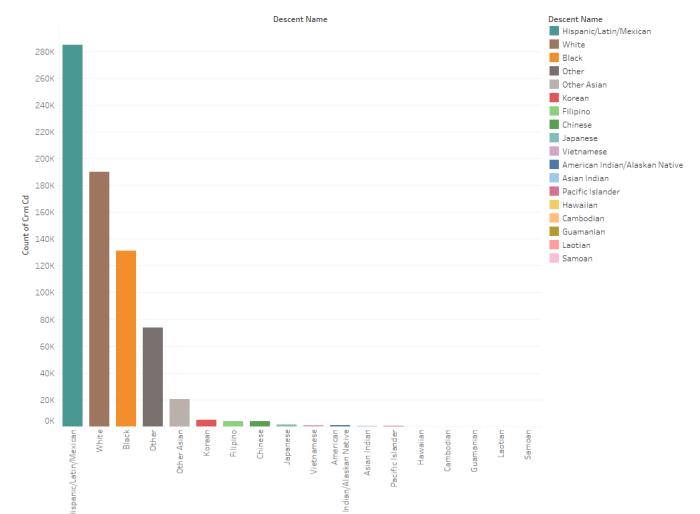


Fig. 41. Communities Present in the Dataset

In figure 41, we see the various communities present. There's 3 very dominant values here.

"Hispanic/Latin/Mexican", "White" and "Black". Further visualisation can be done on these 3.

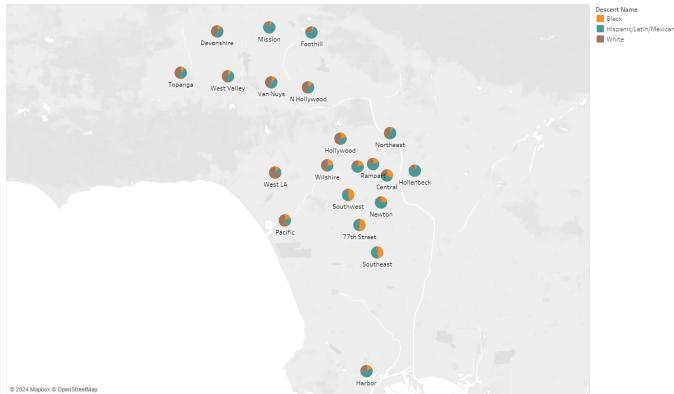


Fig. 42. Descent Based Analysis of Each Area

Figure 42 contains pie charts showing the split of the main 3 communities for each area. This could be due to different levels of population of each community in different areas. There are extremely different pie charts for different areas, for example "Southeast" and "Pacific". This shows the geographical and cultural variation of victims.

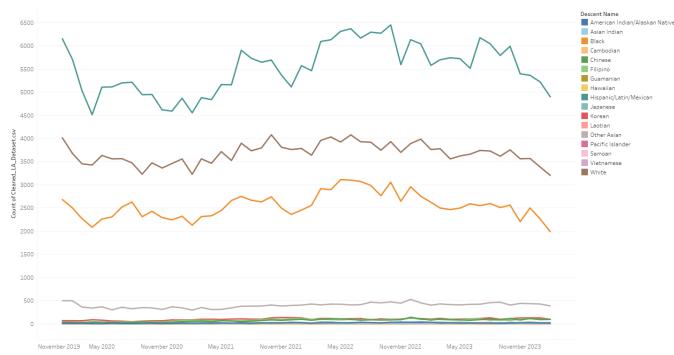


Fig. 43. Descent Based Analysis Over Time

In figure 43, we can see the number of victims by descent from 2020 to 2024. We see that there is a slump in "Hispanic/Latin/Mexican" victims that matches the slump in the overall data. However, the quantity of rest of the values were not affected or barely affected over that period. We also see a dip towards the end in all communities.



Fig. 44. Descent Based Analysis of Part

Figure 44 shows the pie charts for part 1 and 2 for various Descent values. We see that a lot of the communities have a heavy majority of part 1 i.e serious crimes. There are only 3 communities, "Black", "Hispanic/Latin/Mexican", "Laotian", that have a part 2 majority.

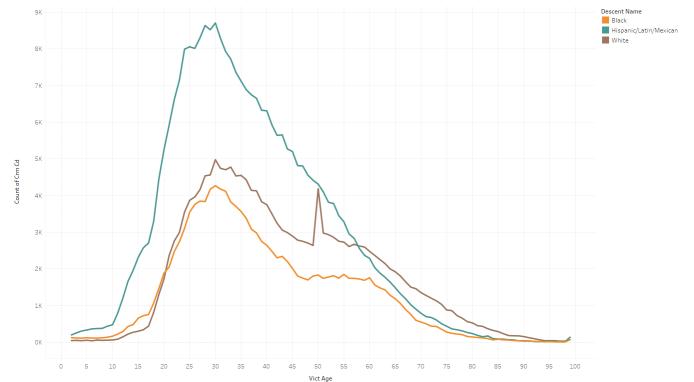


Fig. 45. Age based analysis of Descent

In figure 45, we can see the age distribution for the main 3 communities. We see that "Hispanic/Latin/Mexican" has a peak early then falls off. The other 2 communities are more spread out. We see a sudden peak at 50 for "White". This is seen in the overall age chart. However there is no peak in these communities at 30. Thus that peak must be cause by unknown or "Other" Descent Values.

#### IV. AUTHOR'S CONTRIBUTIONS

- **Subham Agarwala:** visualization and analysis for Task 1
- **Sarvesh Kumar:** visualization and analysis for Task 2
- **Ayush Gupta:** visualization and analysis for Task 3  
Preprocessing of data was a cumulative contribution by all team members.

#### V. REFERENCES

##### REFERENCES

- [1] Kaggle Crime Data. *Los Angeles and Chicago Crime Data from 2000*. Available at: <https://www.kaggle.com/datasets/middlehigh/los-angeles-crime-data-from-2000?resource=download&select=Chicago+Crime+Data.csv>.