LOS ANGELES CRIME REPORT FROM 2020 – 2023

INTRODUCTION

The Los Angeles crime data provides a detailed overview of reported criminal activities within the city from 2020 to 2023. This report presents an exploratory data analysis of the dataset, aimed at uncovering patterns and trends in crime, identifying hotspots, and examining the characteristics of criminal activities across Los Angeles. The goal of this analysis is to equip law enforcement agencies, researchers, policymakers, and the public with actionable insights to make informed decisions that enhance public safety and ensure the effective allocation of resources.

To achieve these objectives, Microsoft Excel was utilized for data cleaning and transformation, while Power BI Desktop was used to conduct the analysis and the visualizations. The insights derived from this analysis will serve as a valuable resource for addressing crime-related challenges and improving safety within the city.

PROBLEM STATEMENT

The increasing complexity and volume of criminal activities in Los Angeles presents significant challenges to law enforcement agencies, policymakers, researchers, and the community. To enhance public safety through optimized resource allocation, it is essential to conduct a thorough exploratory data analysis of crime data in Los Angeles. This analysis aims to uncover patterns, trends, and characteristics of criminal activities in the region, providing critical insights that can inform decision-making and improve crime prevention, policy formulation, and efficient deployment of law enforcement resources.

OBJECTIVES

1. **Crime Trend:** What is the rate of crimes over time (e.g., by month, year, or season)?

2. **Crime Types:** What are the most prevalent crimes (e.g., theft, assault, robbery, etc.)?

3. **Sex Distribution:** What is the distribution of crime by sex (e.g., male, female, unknown)?

4. **Crime Hotspots:** Where are the crime hotspots (e.g., specific neighborhoods, streets, or locations)?

5. **Weapons Used:** Which weapons are commonly used for criminal activities (e.g., firearms, knives, etc.)?

6. **Age Distribution:** What is the distribution of crime by age (e.g., juvenile, adult, senior)?

7. **Racial Distribution:** Which race/ethnicity has the most crime records (e.g., based on victim or offender data)?

8. **Crime Severity:** How serious are the crimes committed (e.g., violent, non-violent, property-related)?

9. **Solve Rate:** What is the solve rate of crimes committed (e.g., percentage of cases solved)?

10. **Victim Demographics:** What is the distribution of crime by victim demographics (e.g., descent, ethnicity, nationality)?

DATA OVERVIEW

The Los Angeles crime dataset consists of a single table with 28 columns, providing detailed information about reported crimes. The crime data table includes attributes such as:

- Date and time of occurrence
- Crime codes and descriptions
- Victim demographics (sex, age, and race/ethnicity)
- Premises description
- Crime status
- Location coordinates (Latitude and Longitude)

DATA RETRIEVAL

The dataset was retrieved from the Kaggle website at Kaggle.com and was loaded into Microsoft Excel for Data cleaning and pre-processing.

DATA CLEANING AND PRE-PROCESSING

The following actions were taken to clean the dataset to ensure accuracy and reliability of our analysis:

1. **Renaming column names**: the columns were renamed to increase clarity in our analysis and results.
2. **Deleted irrelevant columns**: columns considered irrelevant to our analysis were deleted to reduce confusion and increase analysis speed. Such columns included crime codes, area code, Mocodes, report district number, premise code, weapon used code, crime code 1, code 2, code 3, and code 4.
3. **Formatting datatypes**: some columns had the wrong datatypes which could result to error during analysis, these columns was formatted to the preferred datatype;
   - formatted the date column to short date
   - changed the time of occurrence column from number datatype to time datatype using the formular =--(TEXT(D2,"00\:00")

`=--TEXT(D2,"00\:00")`

ɔrted	Date of Occurrence	Time of Occurence		Area Name
1/08/2020	01/08/2020	2230	22:30	Southwest
1/02/2020	01/01/2020	330	03:30	Central
4/14/2020	02/13/2020	1200	12:00	Central
1/01/2020	01/01/2020	1730	17:30	N Hollywood
1/01/2020	01/01/2020	415	04:15	Mission
1/02/2020	01/01/2020	30	00:30	Central

4. Replaced N Hollywood to North Hollywood in the District column using the find and replace tab.
5. Deleted rows with negative values in the victim's age column.
6. Changed the victim's sex from M, F, X and H to Male, Female, Unknown, and Heterosexual respectively using the find and replace tab.
7. Renamed the victim's descent column from letters to their full names using the find and replace tab.
8. Change the incident column and weapon used column from upper case to proper case using the formular =PROPER(K2).

`=PROPER(K2)`

K	L
Incident Location	
Single Family Dwelling	Single Family Dwelling
Sidewalk	Sidewalk
Police Facility	Police Facility

9. Deleted rows that contained UNK in incident status.
10. Filtered out rows that were recorded as Null and Unknown, 0, and Null in the victim's sex, victim's age and weapon used columns respectively.
11. Deleted rows with invalid street address in the incident location column.

Custom Autofilter   ?   ✕

Show rows where:
Crime Location

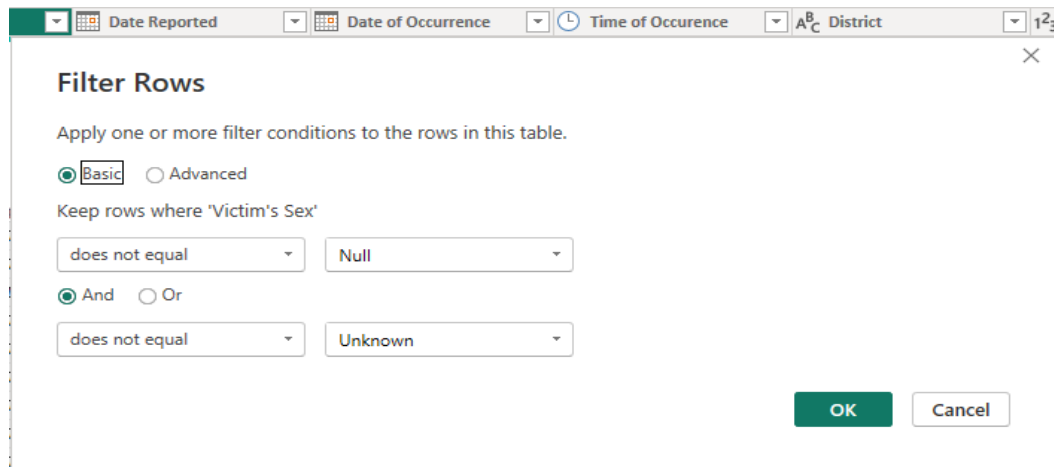| ends with | ▾ | 6th* | ▾ |

◉ And ○ Or

| | ▾ | | ▾ |

Use ? to represent any single character
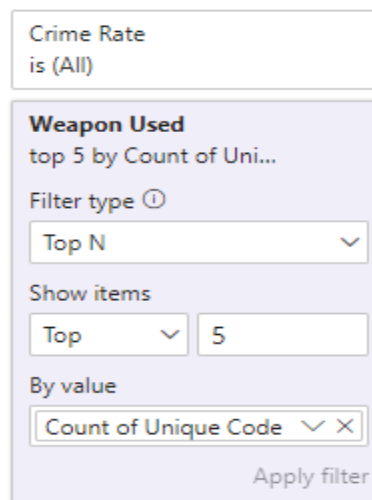Use * to represent any series of characters

OK   Cancel

After the cleaning process in Excel, the dataset was loaded to Power BI where some columns were categorized like the longitude and latitude columns were categorized to longitude and latitude.

12. In PowerBi, power query was used to filter unwanted data such as NULLS and UNKNOWNS in Victim's Age column and other discrepancies.

13. Top N was mostly used in summarizing the charts like Weapons used, Crimes Description, Victim's Descent, Incident Locations etc.

14. We have 1-99 age range, in other to accommodate the ages, class interval was introduced to the chart and this was done either with Switch statements in DAX or by grouping the ages with lists.

   i. Using Switch statements

```
1. Age Range = SWITCH(
3.     TRUE(),
4.     Crime_Data_from_2020_to_Present[Victim's Age] <= 10, "1-10",
5.     Crime_Data_from_2020_to_Present[Victim's Age] <= 20, "11-20",
6.     Crime_Data_from_2020_to_Present[Victim's Age] <= 30, "21-30",
7.     Crime_Data_from_2020_to_Present[Victim's Age] <= 40, "31-40",
8.     Crime_Data_from_2020_to_Present[Victim's Age] <= 50, "41-50",
```
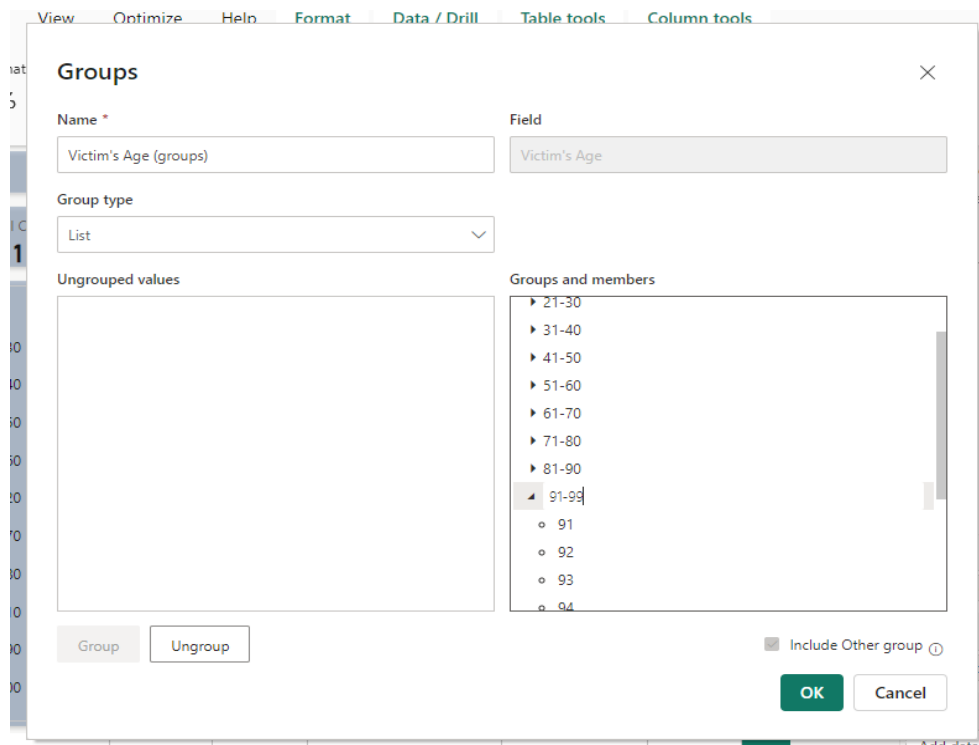
```
 9.     Crime_Data_from_2020_to_Present[Victim's Age] <= 60, "51-60",
10.     Crime_Data_from_2020_to_Present[Victim's Age] <= 70, "61-70",
11.     Crime_Data_from_2020_to_Present[Victim's Age] <= 80, "71-80",
12.     Crime_Data_from_2020_to_Present[Victim's Age] <= 90, "81-90",
13.     Crime_Data_from_2020_to_Present[Victim's Age] <= 100, "91-100",
14.     "Other"
15.                 )
```
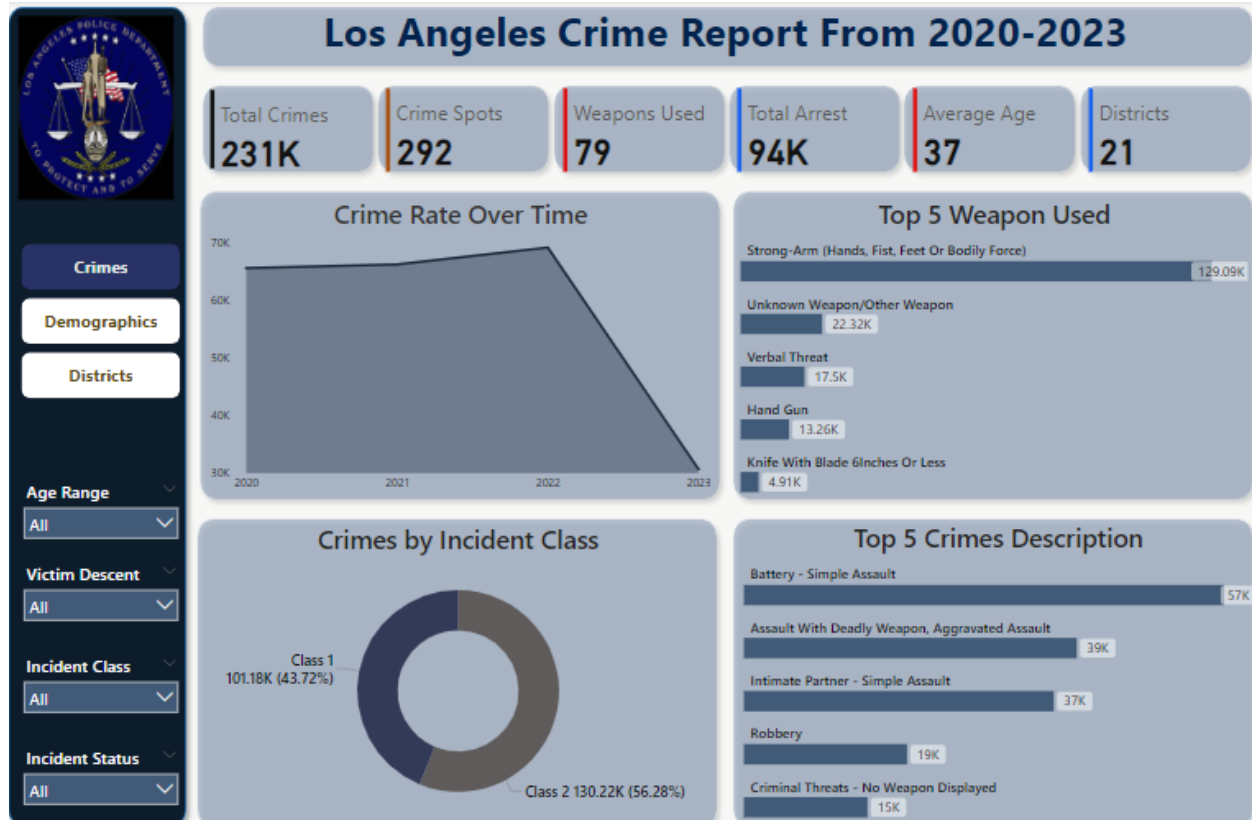
ii.   Using the grouping by list method
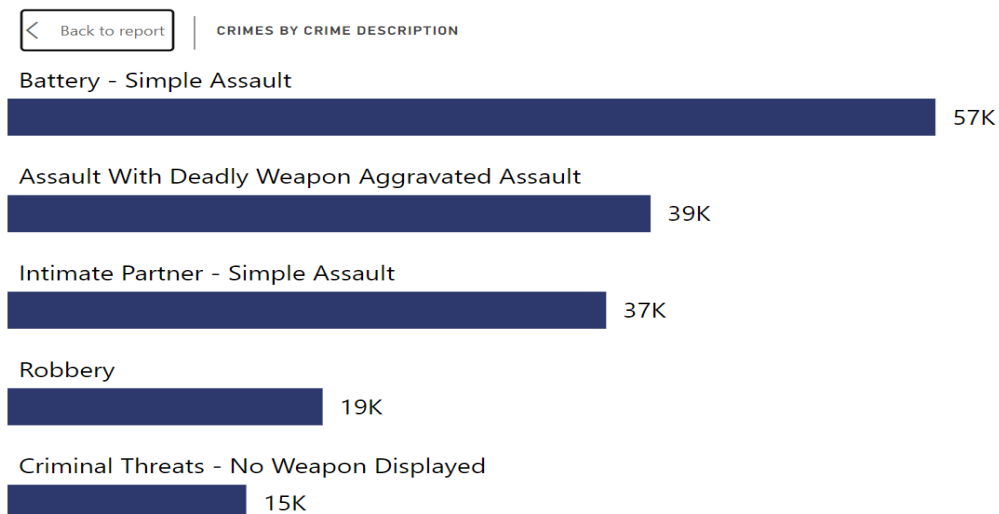
DATA ANALYSIS AND VISUALIZATION

The Analysis was grouped into three, Crime analysis, District analysis and Demographics analysis.

Crime Analysis

The following questions was answered to better understand the crimes carried out in the city:

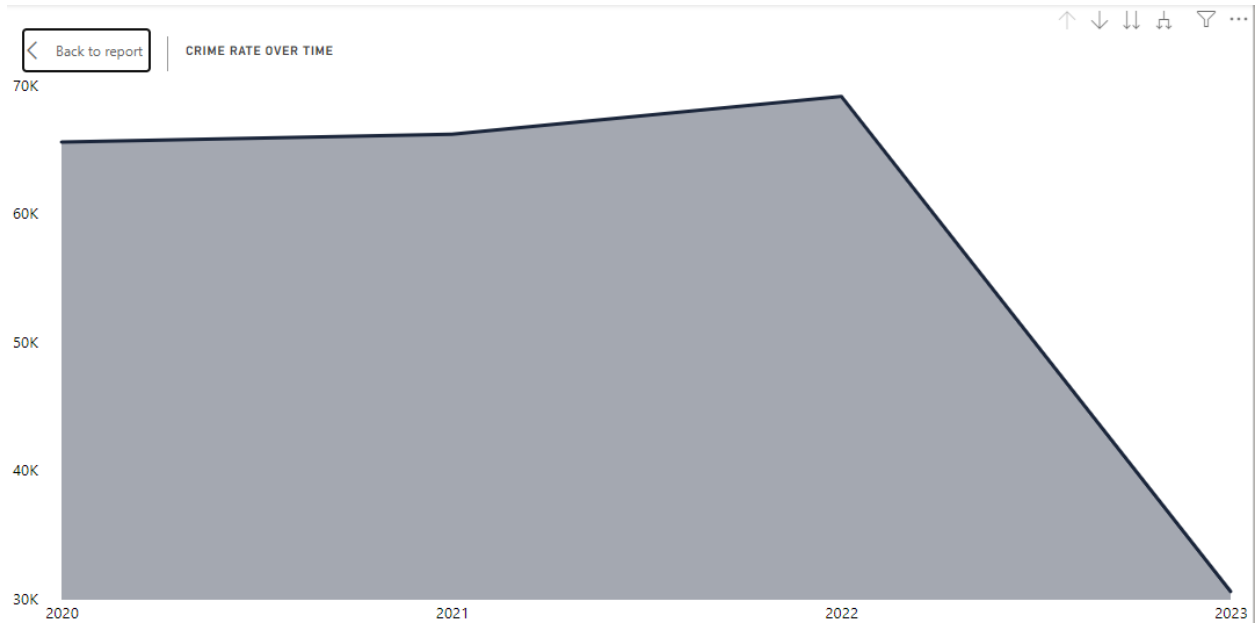1. ## What are the most prevalent crimes?

Using a bar chart, we plotted the crime description over the total crimes, we filtered the chart to show the top 5 most prevalent crimes in the city of Los Angeles.

In Los Angeles the Most common crime committed had to do with Assaults of different kinds, Robbery and Criminal threats with Simple assault topping the chart with 56,943.
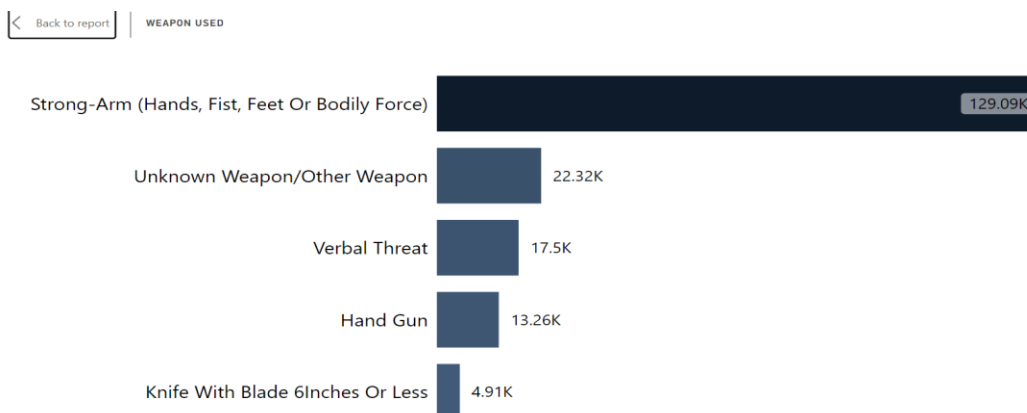
2. What is the rate of crimes over time?

Using an area chart, the crime rate was plotted against the date to show the trend in crimes over the four years recorded.

From the chart, we discovered that the crime rate was increasing steadily from 2020 to 2021 with total crime rate of 65,562 to 66,177 respectively and hit a peak in 2022 with 69,114 total crimes after which, there was a rapid decline in 2023 with total crime rate of 30,548.

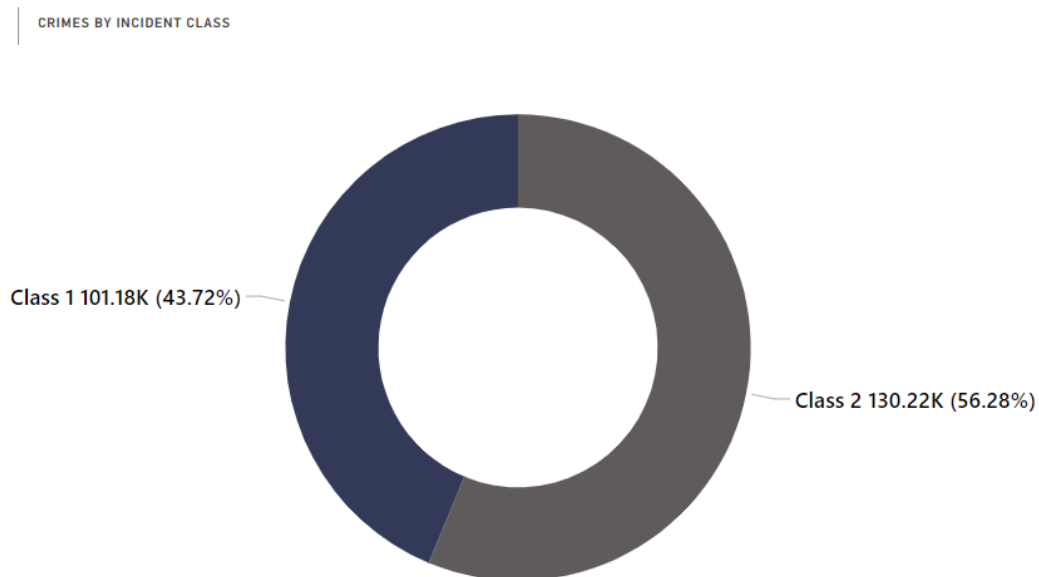## 3. Which weapons are mostly used for criminal activities?

Narrowing the analysis to the top 5 weapon used for committing crimes, it was discovered that the weapon used to perpetrate most of the crimes is Strong-Arm with a total of 129,087. Use of verbal threat, handgun, and knife were also indicated as common weapons in descending order.

4. How serious are the crimes committed?

The seriousness of the crime was indicated in the incident class column, which classified crimes into classes 1 and 2. According to the data, class 1 crimes include criminal homicide, Forcible rape, Robbery, Aggravated assault, Burglary, Larceny theft, Grand theft auto, and Arson. Class 2 crimes include forgery, Fraud and nsf checks, Sex offenses felonies, Sex offenses misdemeanor, non-aggravated assaults, Weapon laws, offenses against family, Drunk/alcohol/drugs etc.

Using a doughnut chart, the crime status was observed.

CRIMES BY INCIDENT CLASS

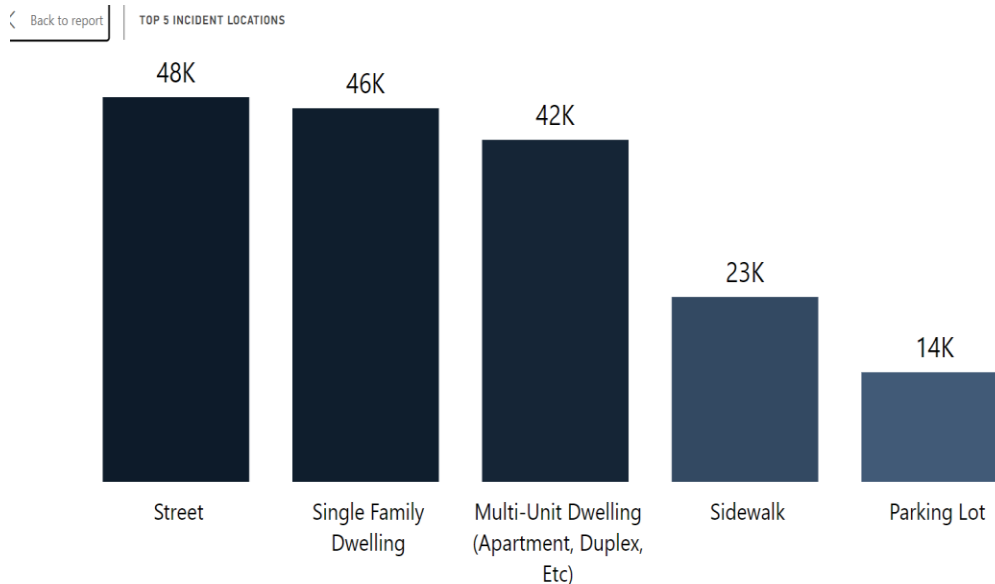Class 1 101.18K (43.72%)

Class 2 130.22K (56.28%)

From the chart above, 56.28% of the crimes committed are class 2 crimes, with 43.72% are class 1 crimes. This shows that Los Angeles city is more exposed to less dangerous criminal activities which may not be life threatening but makes the city to be in an unrest situation. Drilling down the incident chart with victim sex, it was found that for incident class 1, the victims were mostly male with a total number of 61,786(61.07%), the female victims with 39,382 (38.92%). For incident class 2, it was observed that female was the most affected sex with a total of 75,559 (58.02%), followed by male with 54,651 (41.97%), then heterosexual with 0.01% crime rates respectively.
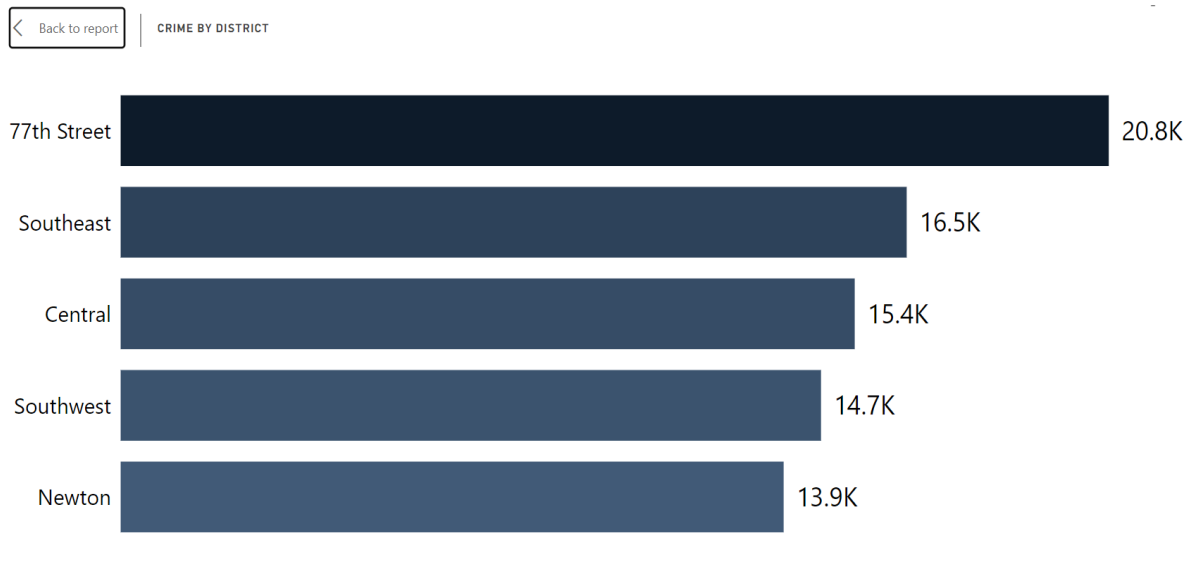
DISTRICT ANALYSIS

1. Which location recorded the most crimes?

TOP 5 INCIDENT LOCATIONS

- Street — 48K
- Single Family Dwelling — 46K
- Multi-Unit Dwelling (Apartment, Duplex, Etc) — 42K
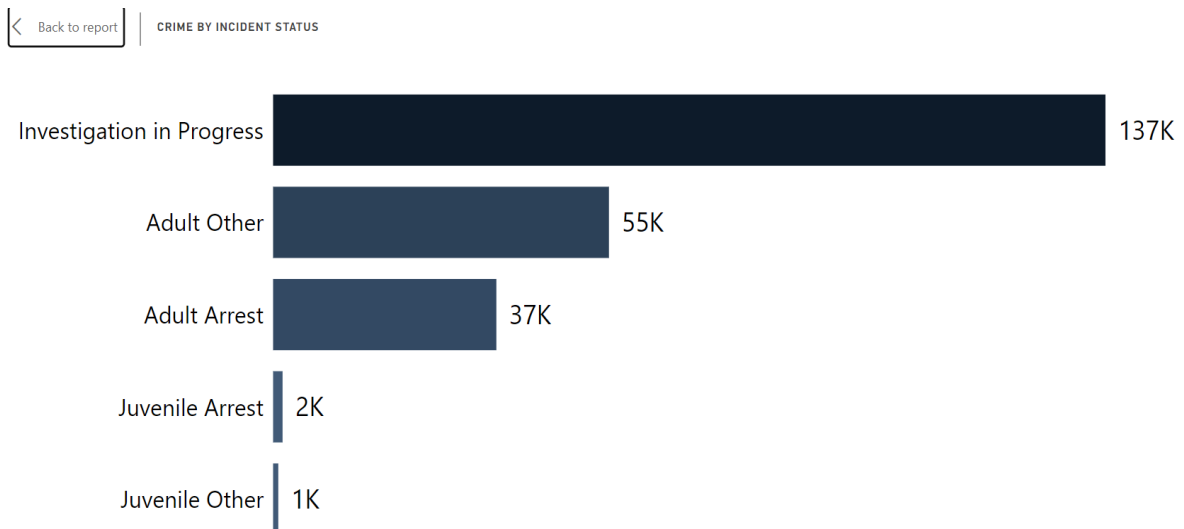- Sidewalk — 23K
- Parking Lot — 14K

From the analysis, most of the crimes happened in the street with a total of 47,621 crimes, followed by Single family dwelling with a total of 46,246 crimes, multi-unit dwelling (apartment, duplex, etc.) with a total of 42,338 crimes, Sidewalks with a total of 22,888 crimes and parking lot with a total of 13,567 crimes.

2. Which district recorded the most crimes?

Based on the Los Angeles top 5 districts based on crime, 77th street recorded the highest crime rate of 20,758, followed by Southeast, Central, Southwest and Newton districts with a total of 16,513, 15,421, 14,713, 13,928 crimes respectively.
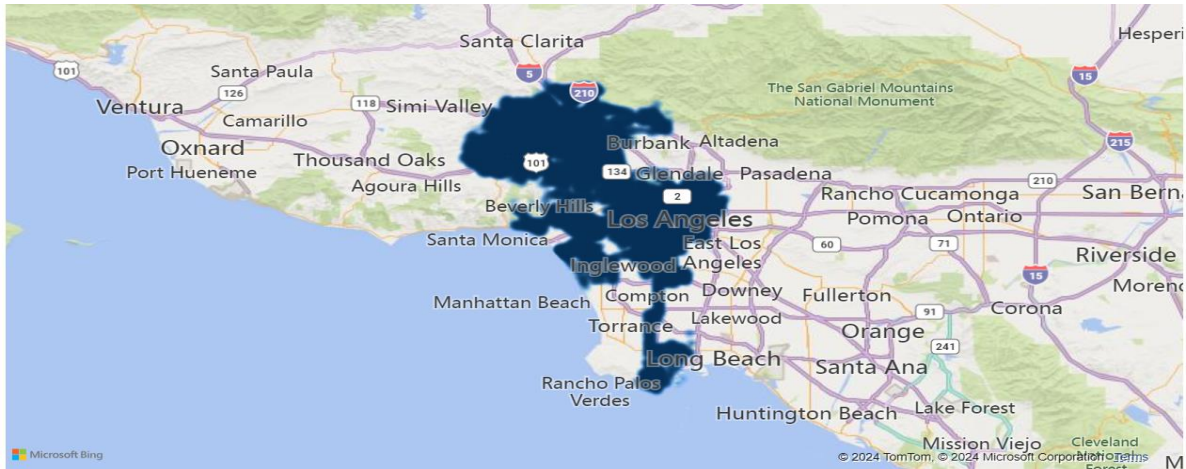
3. What is the solve rate of crimes committed?

From the illustration above, it was discovered that investigation in progress has a total of 136,976 cases, followed by adult other with a total of 55,267 cases, followed by adult arrest with a total

case of 36,742, Juvenile arrest and juvenile other with a total of 1,561 and 855 cases respectively. It can be observed that the rate at which offenders are being punished is low, which might encourage crime.

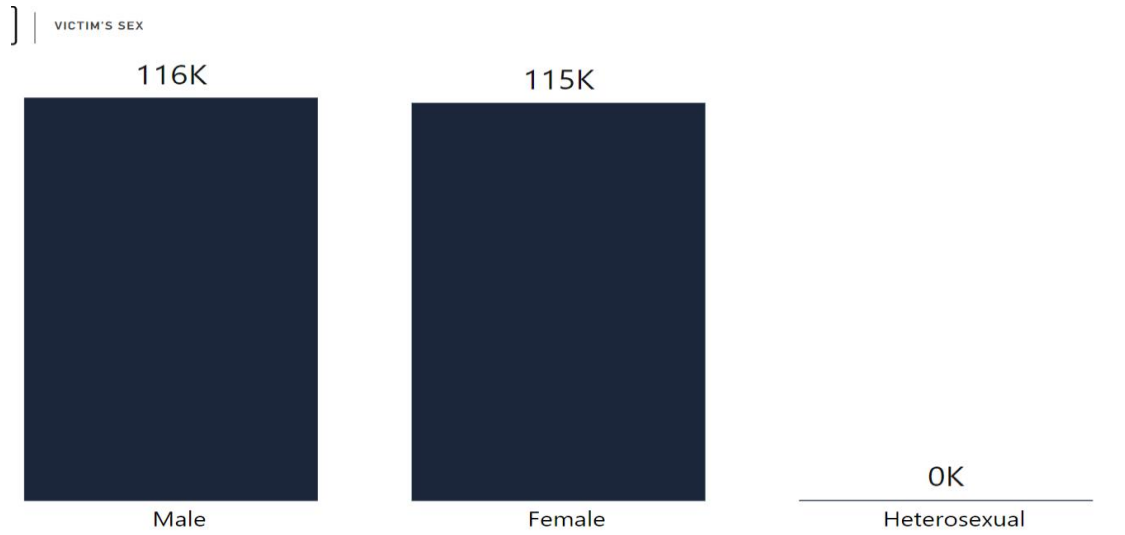## 4. Where are the crime hotspots?

Visualizing the crime hotspot using a heat map, it can be easily deduced that the crime rate is fast spreading in the cities.
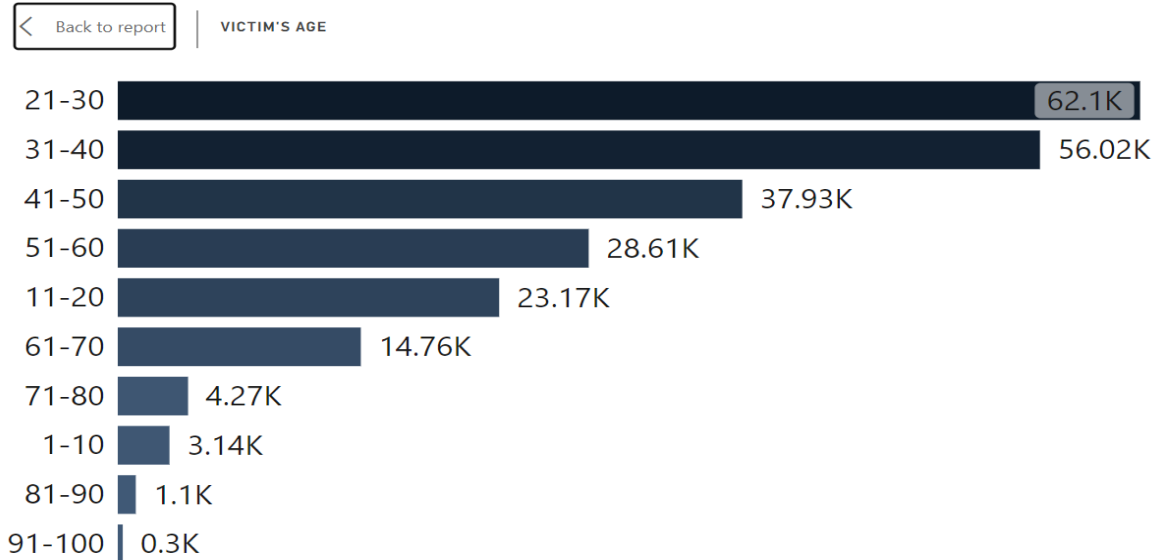
DEMOGRAPHICS ANALYSIS

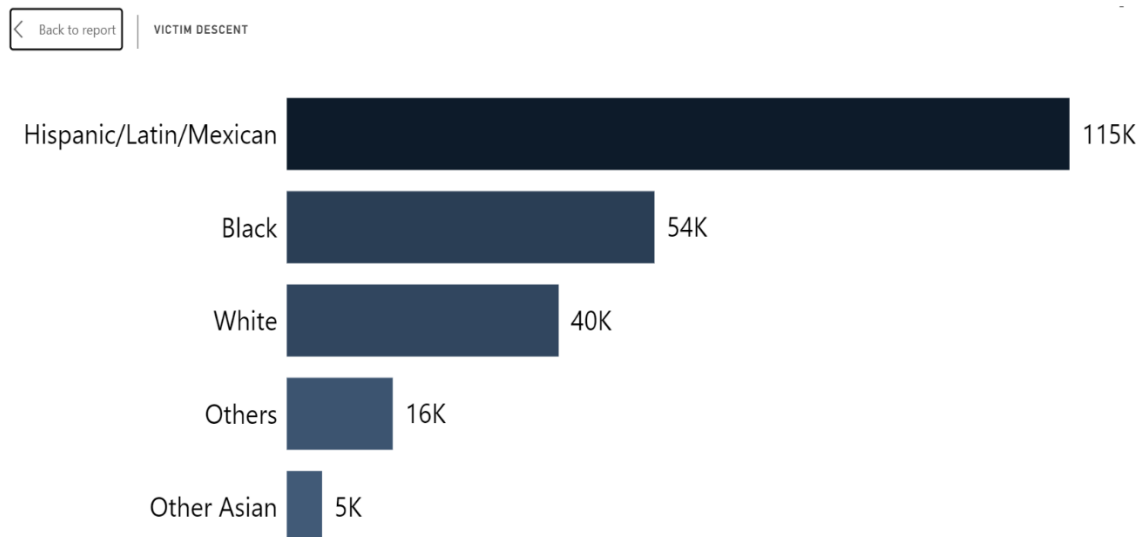## 1. What is the distribution of crime by sex?

From the analysis, the most affected victims were males with a total crime rate of 116,437, followed by females with a crime rate of 114,941, and heterosexual recorded 23 crimes.

2. What is the distribution of crime by age?

Age	Value
21-30	62.1K
31-40	56.02K
41-50	37.93K
51-60	28.61K
11-20	23.17K
61-70	14.76K
71-80	4.27K
1-10	3.14K
81-90	1.1K
91-100	0.3K

From the chart, it can be observed that ages between 21-30 were the most affected victim's age with a total of 62,100 followed by ages 31-40 with a total number of 56,018 and ages 41-50 with a total of 37,934 crime rates.

3. What is the distribution of crime by victim descent?

Descent	Value
Hispanic/Latin/Mexican	115K
Black	54K
White	40K
Others	16K
Other Asian	5K

From the chart above, the top 5 most affected races are the Hispanic/Latin/Mexican with a total of 114,937, followed by Black with a total of 53,989, White with a total of 39,922, others with a total of 15,564 and other Asians with a total of 5,174 crimes.

KEY INSIGHTS

1. Between 2020 and 2023, a total of 231,401 crimes were committed in Los Angeles, with the most common being with Assaults of different kinds, Robbery and Criminal threats which happen mostly streets, single and multi-family dwellings, sidewalks and parking lots.

2. Although investigations are still pending for 74,802 crimes, analysis of the solved cases reveals that the most used weapons were bodily force, verbal threats, handguns, pistols, and knives. Specifically, bodily force was often linked to simple assaults, while guns were frequently used in dangerous assaults and robberies. In summary, it shows the solve rate of cases in Los Angeles is low.

3. In Los Angeles, a staggering 43.72% of crimes are classified as dangerous Class 1 offenses, with a dismally low solve rate, while Class 2 crimes, making up 58.02% of cases, have a slightly higher solve rate, overall indicating a chaotic crime rate with most criminals remaining at large and likely to continue offending.

4. The crime rate in Los Angeles peaked in 2022 with a total of 69,114 crimes, and decreased significantly in 2023 to 30,548 crimes, with January being the month with the highest crime rate, particularly for identity theft, due to increased financial activities and online risks following the holiday season and start of tax season.

RECOMMENDATIONS

1. More security personnel, and check posts should be provided, CCTV cameras should be installed on the streets and parking lots.

2. More resources should be allocated to hasten investigation, and law enforcement agents should be trained to recognize crime patterns. Sensitization of victims on self-defense will help reduce vulnerability to assault.

3. To reduce Identity theft, cybersecurity measures should be improved like use of 2-factor authentication and use of secure websites for online shopping and transactions.

CONCLUSION

The crime statistics in Los Angeles from 2020 to 2023 reveal a concerning trend of high crime rates, particularly Assaults of different kinds, Robbery and Criminal threats, with a majority remaining unsolved. The low solve rate and high percentage of dangerous crimes highlight the need for enhanced security measures, improved investigation resources, and increased awareness and training for law enforcement and citizens. By addressing these issues, Los Angeles can work towards reducing crime rates and creating a safer environment for its residents.