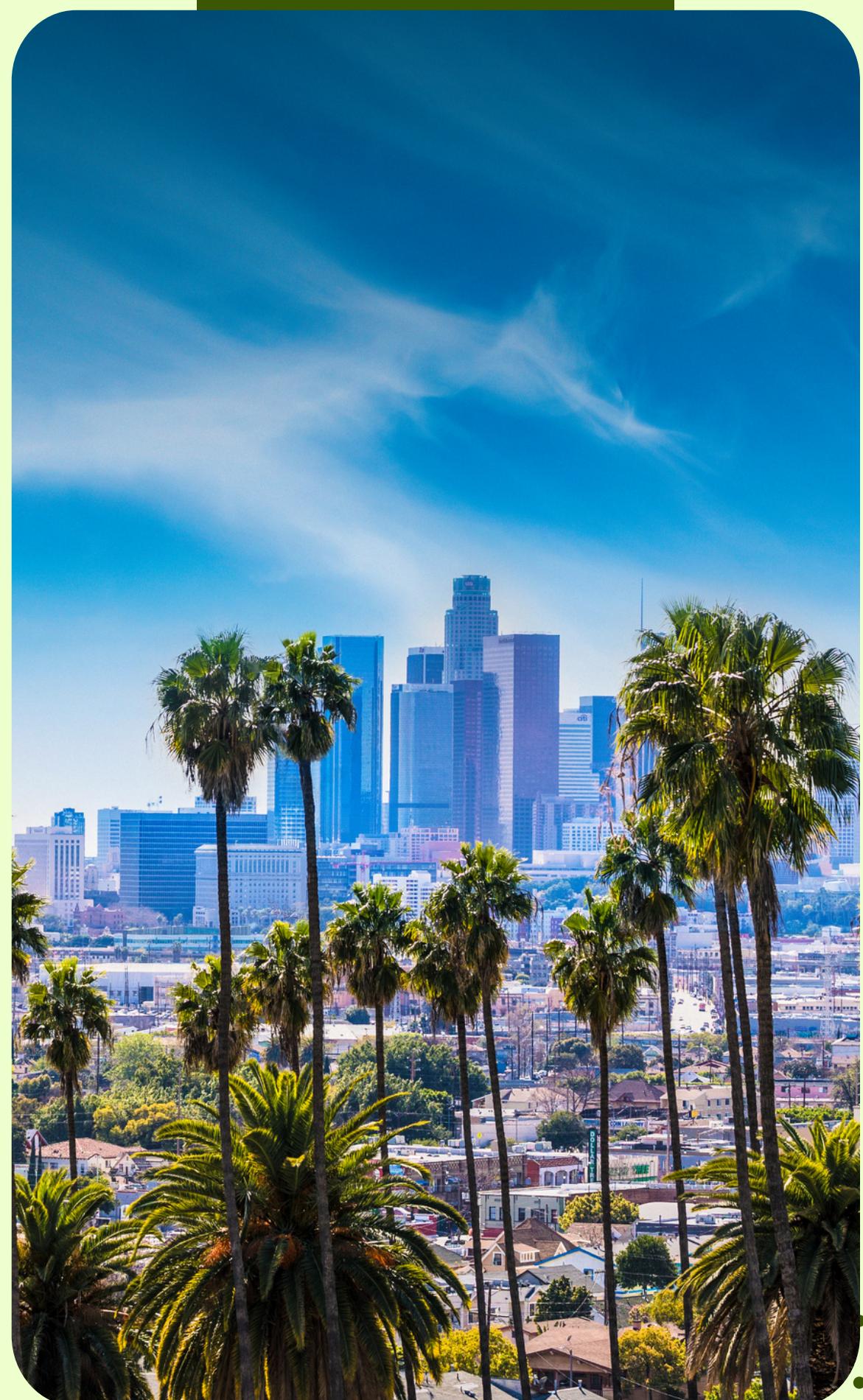


# Los Angeles County

Crime Data Analysis: 2020 to Present

By Anny Patino



# Content

01

Introduction

02

Methodology

03

Analysis & Findings

04

Discussion

05

Conclusion



# INTRODUCTION

**Objective:** This report provides a comprehensive analysis of crime data in Los Angeles County from 2020 to the present, aiming to understand patterns and trends in crime types, weapon usage, demographics, locations, and law enforcement response times.

## Data Source:

Los Angeles Police Department. “Crime Data from 2020 to Present.” Public Safety | Los Angeles - Open Data Portal. Accessed on October 9, 2023. Available at:

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

DR_NO	Date Rptd	DATE OCC	TIME OCC	AREA	AREA NAME	Rpt Dist No	Part 1-2	Crm Cd	Crm Cd Desc	...	Status	Status Desc	Crm Cd 1
0	10304468	01/08/2020	12:00:00 AM	01/08/2020	12:00:00 AM	2230	3	Southwest	377	2	624	BATTERY - SIMPLE ASSAULT	...
1	190101086	01/02/2020	12:00:00 AM	01/01/2020	12:00:00 AM	330	1	Central	163	2	624	BATTERY - SIMPLE ASSAULT	...
2	200110444	04/14/2020	12:00:00 AM	02/13/2020	12:00:00 AM	1200	1	Central	155	2	845	SEX OFFENDER REGISTRANT OUT OF COMPLIANCE	...
3	191501505	01/01/2020	12:00:00 AM	01/01/2020	12:00:00 AM	1730	15	N Hollywood	1543	2	745	VANDALISM - MISDEAMEANOR (\$399 OR UNDER)	...
4	191921269	01/01/2020	12:00:00 AM	01/01/2020	12:00:00 AM	415	19	Mission	1998	2	740	VANDALISM - FELONY (\$400 & OVER, ALL CHURCH VA...	...

5 rows × 28 columns

## Libraries and Tools Used:

- **Pandas:** Data manipulation and analysis library for Python
- **Matplotlib:** Python plotting library
- **Seaborn:** Statistical data visualization library for Python
- **Folium:** Python library for creating interactive maps
- **Wordcloud:** text mining method to find the most frequently used words in a text

# METHODOLOGY

- Leveraging the Los Angeles Open Data, the analysis encompassed statistical calculations, data visualization using word clouds, bar charts, and pie charts, with an emphasis on temporal and spatial variables.

## ▼ Data Consistency Check

```
[ ] # 1. Date Consistency Check

# Date Consistency Check: Ensure 'DATE OCC' is on or before 'Date Rptd' and no future dates are present
date_consistency_check = (crime['DATE OCC'] <= crime['Date Rptd']).all() and (crime['DATE OCC'].max() <= pd.Timestamp.now())

# 2. Time Consistency Check: Ensure hours are within 0 to 23 and minutes are within 0 to 59
crime['Hour'] = crime['TIME OCC'].str.split(':').str[0].astype(int)
crime['Minute'] = crime['TIME OCC'].str.split(':').str[1].astype(int)

hours_check = (crime['Hour'] >= 0) & (crime['Hour'] <= 23)
minutes_check = (crime['Minute'] >= 0) & (crime['Minute'] <= 59)
time_consistency_check = hours_check.all() and minutes_check.all()

# 3. Age Consistency Check
# Check for any unreasonable age values (e.g., negative ages or extremely high ages)
age_consistency_check = crime['Vict Age'].dropna().between(0, 120).all()

# Print results
print("Date Consistency Check:", date_consistency_check)
print("Time Consistency Check:", time_consistency_check.all())
print("Age Consistency Check:", age_consistency_check)
```

```
Date Consistency Check: True
Time Consistency Check: True
Age Consistency Check: True
```

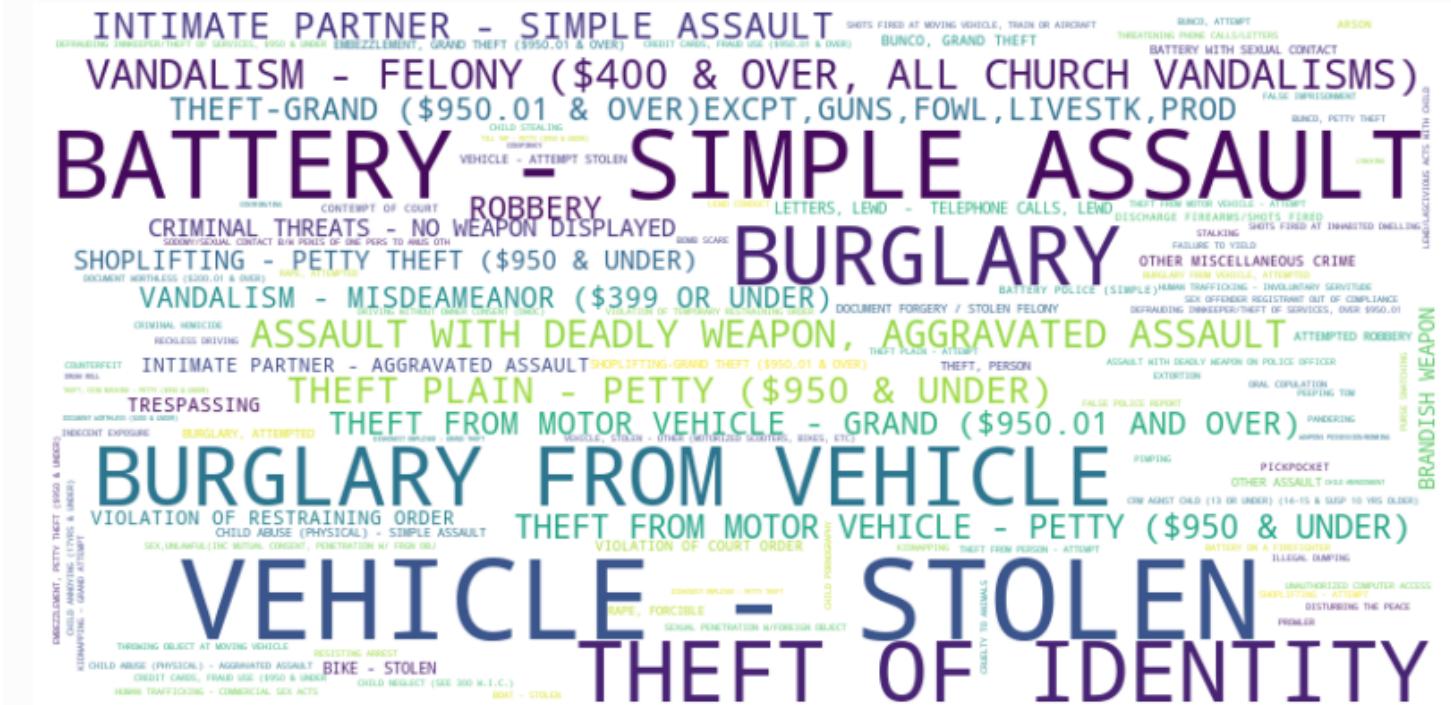
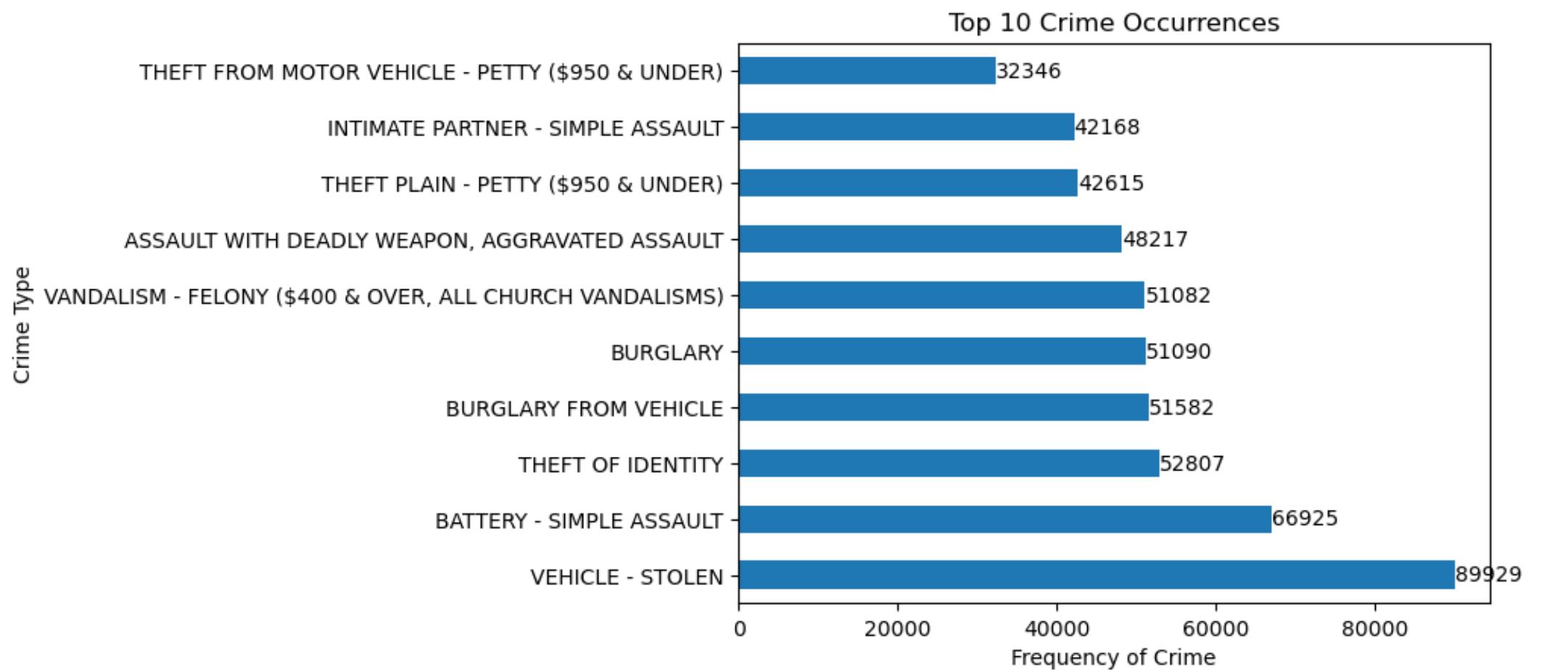
# ANALYZIS & FINDINGS



# CRIME TYPES

**Objective:** This analysis aims to discern and quantify the prevalent types of crimes in Los Angeles County.

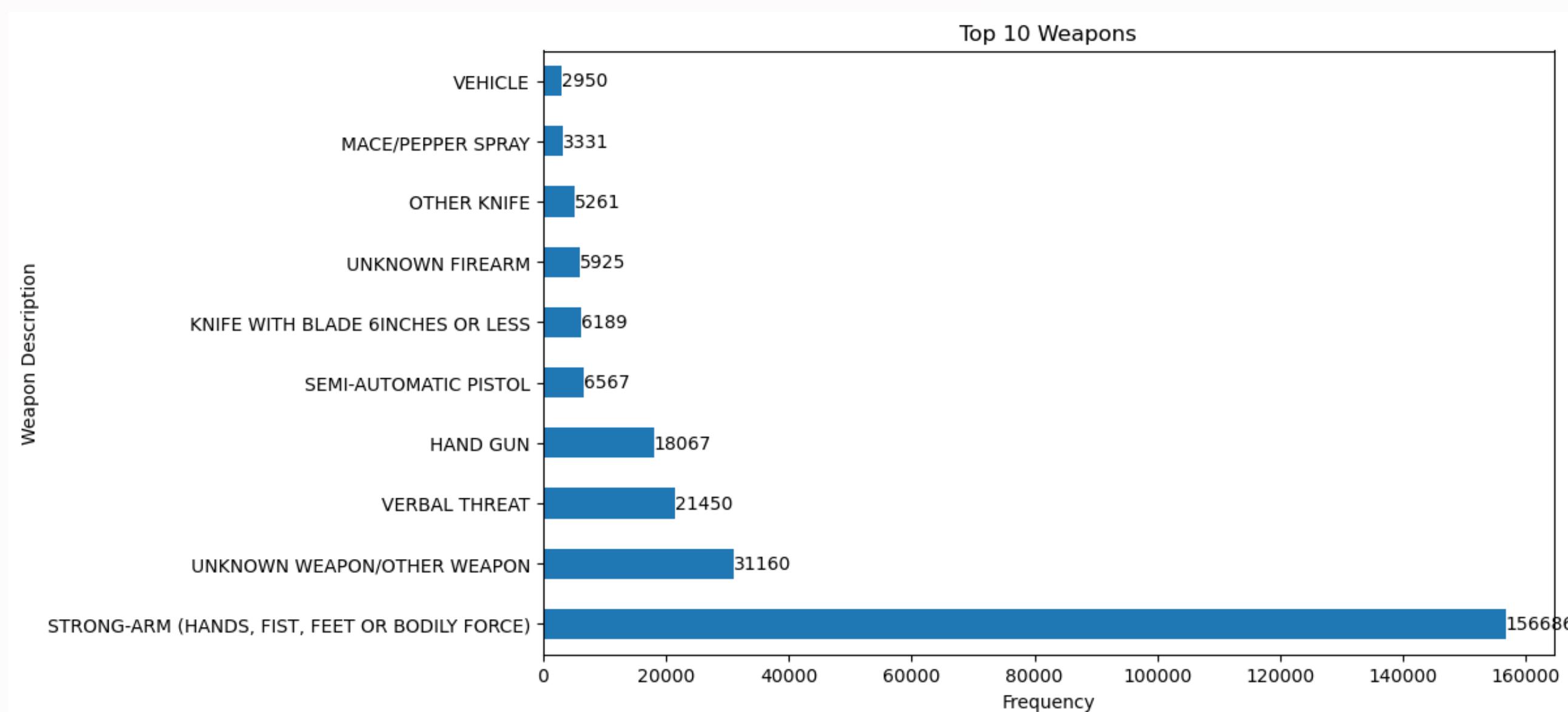
**Key Findings:** Theft-related offenses, particularly vehicle theft and burglary, dominate the crime landscape, with specific peaks observed in assault and vandalism.



# WEAPON USAGE

**Objective:** The objective of this analysis was to identify and quantify the types of weapons used in crimes, offering insights into the nature of weapon-related incidents.

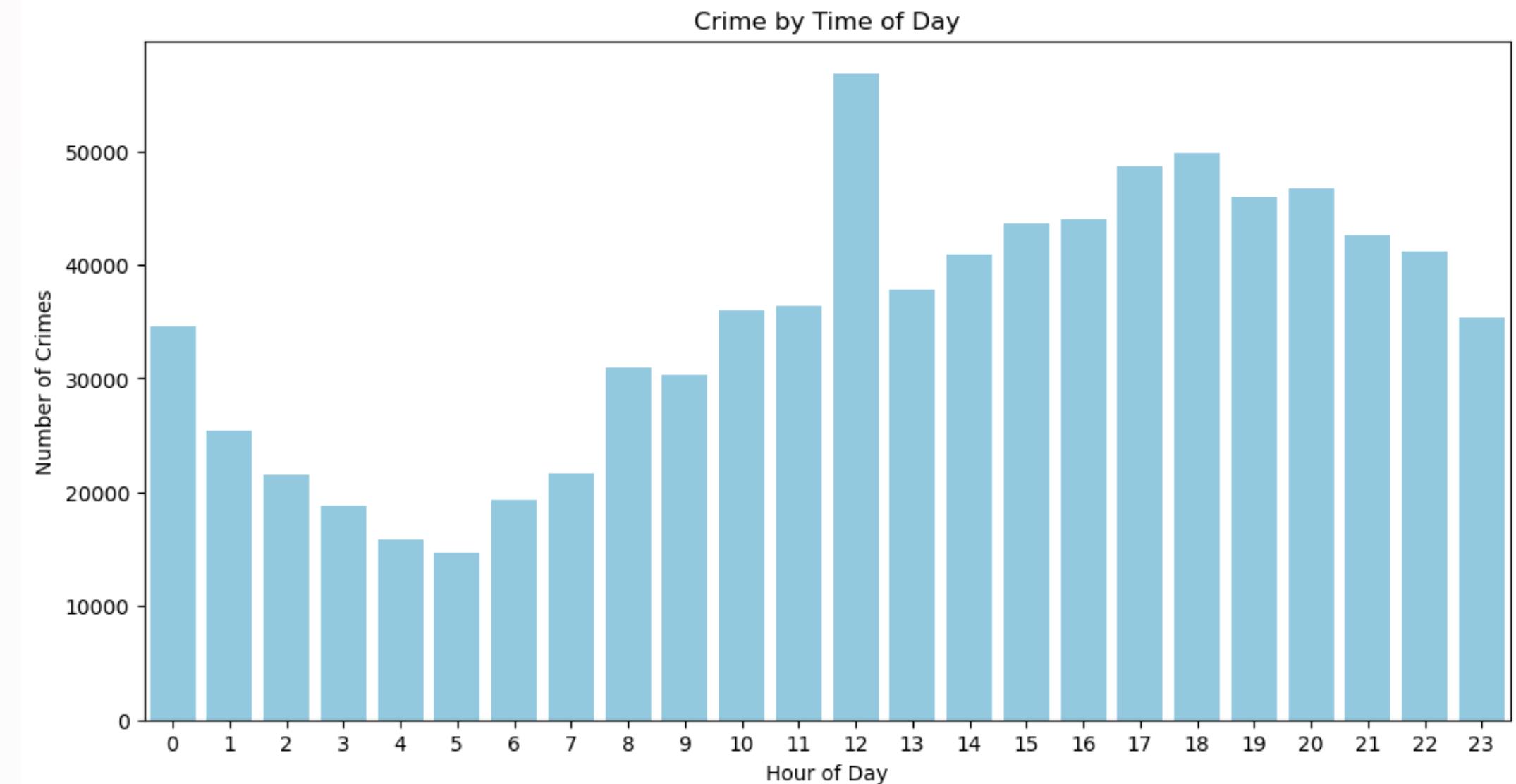
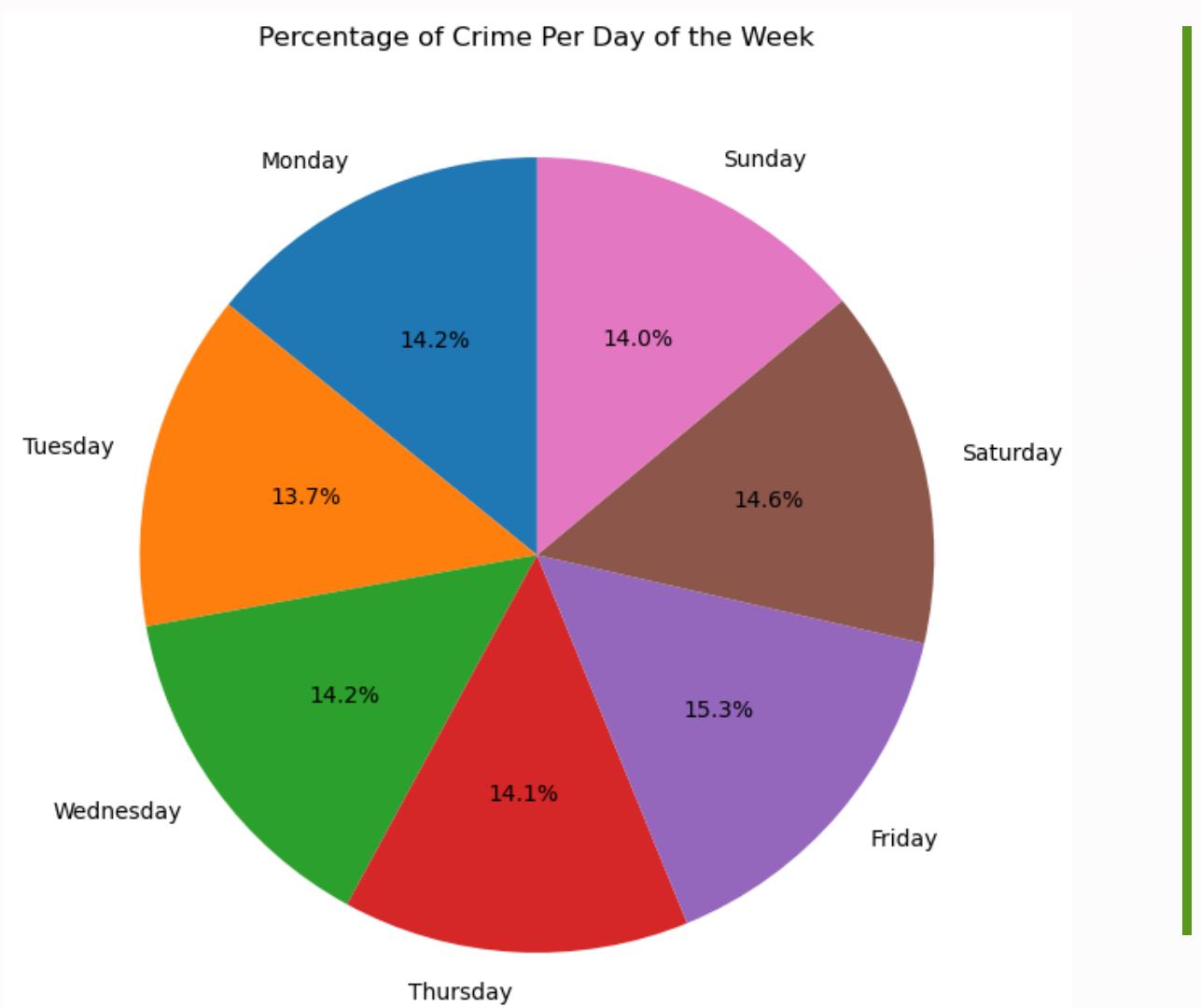
**Key Findings:** Physical force ("Strong Arm") and firearms are the most prevalent weapons used in crimes. 65% of all reported crimes were weapon-free with only 292,391 weapon-related crime cases. The data shows a notable reliance on physical force and firearms, with 'Strong Arm' used in approximately 20% of weapon-related crimes.



## TEMPORAL TRENDS

**Objective:** Analyze crime trends in Los Angeles County with a focus on how crime occurrence varies by year, month, day of the week, and time of day.

**Key Findings:** Crime rates fluctuate throughout the year, with notable increases during the summer months and on Fridays.



```

crime['Year'] = crime['DATE OCC'].dt.year
crime['Month'] = crime['DATE OCC'].dt.month_name()

# Filter data for the years 2020, 2021, 2022, and 2023
filtered_crime = crime[crime['Year'].isin([2020, 2021, 2022, 2023])]

# Group by year and month, count the number of crimes
crime_by_year_month = filtered_crime.groupby(['Year', 'Month']).size().unstack()

# Ensure that the months are in the correct order
months_order = list(calendar.month_name)[1:] # Starting from January
crime_rate_by_year_month = crime_by_year_month.reindex(months_order, axis=1)

# Calculate crime rates per month
crime_rate_by_year_month = crime_rate_by_year_month.div(crime_rate_by_year_month.sum(axis=1), axis=0) * 1000

# Plot the line graph
plt.figure(figsize=(12, 6))
for year in [2020, 2021, 2022, 2023]:
    plt.plot(crime_rate_by_year_month.loc[year], label=str(year), marker='o')

# Highlight the last record date in 2023
last_record_2023 = filtered_crime[filtered_crime['Year'] == 2023]['DATE OCC'].max()
if last_record_2023:
    last_record_month = last_record_2023.month_name()
    last_record_value = crime_rate_by_year_month.at[2023, last_record_month]
    plt.scatter([last_record_month], [last_record_value], color='red', marker='*', s=100, label='Last Record in 2023')

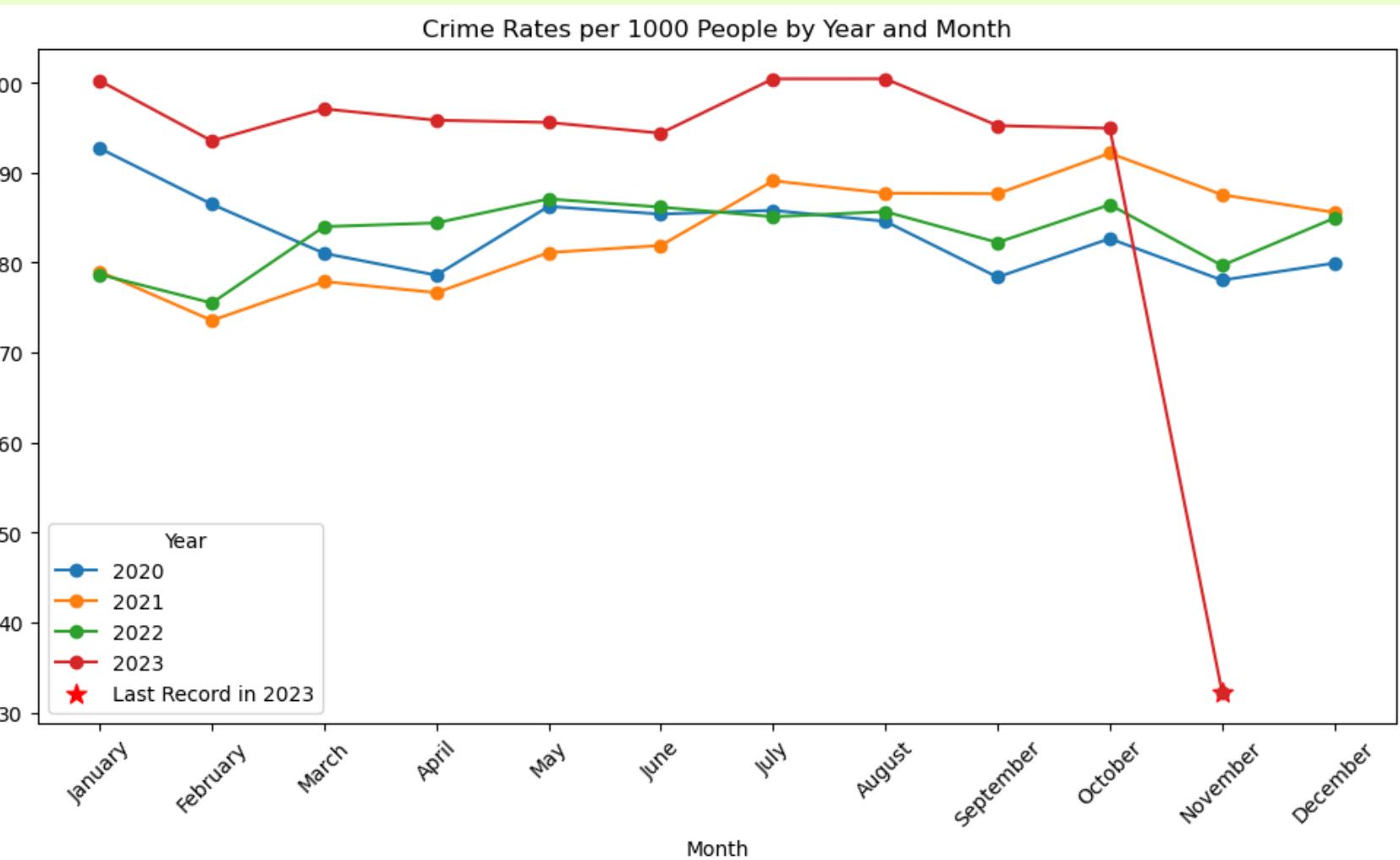
plt.title('Crime Rates per 1000 People by Year and Month')
plt.xlabel('Month')
plt.ylabel('Crime Rate per 1000 People')
plt.xticks(rotation=45) # Rotate x-axis labels for better visibility
plt.legend(title='Year')

# Adding a description to the bar chart
plt.text(0.5, -0.1, "Figure 9. This bar plot shows the crime rate by year.", ha='center', va='center', fontsize=16, transform=plt.gcf().transFigure)

# Save the plot as a PNG file
plt.savefig('../reports/figures/crime_rate_by_year_month.png', bbox_inches='tight')

plt.show()

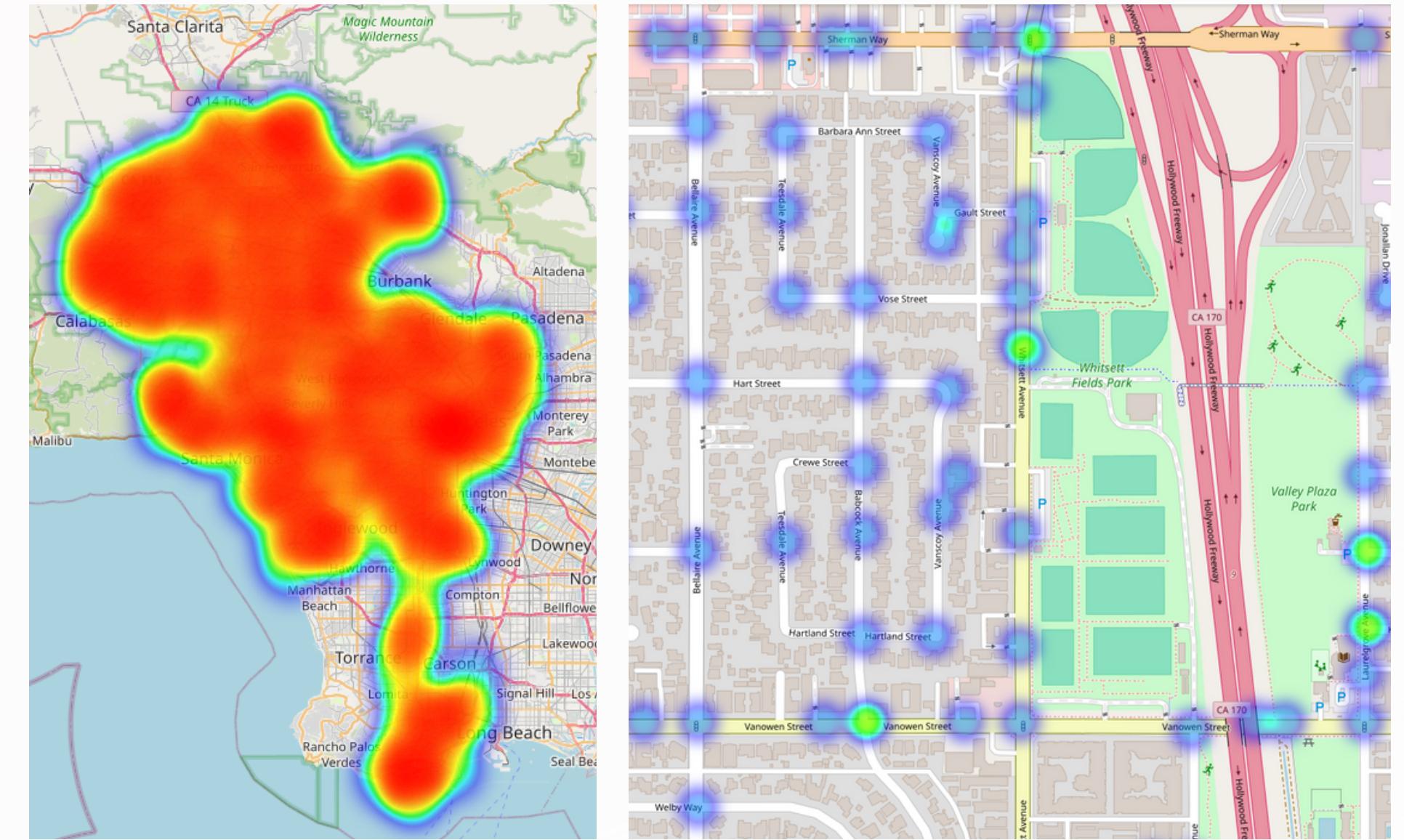
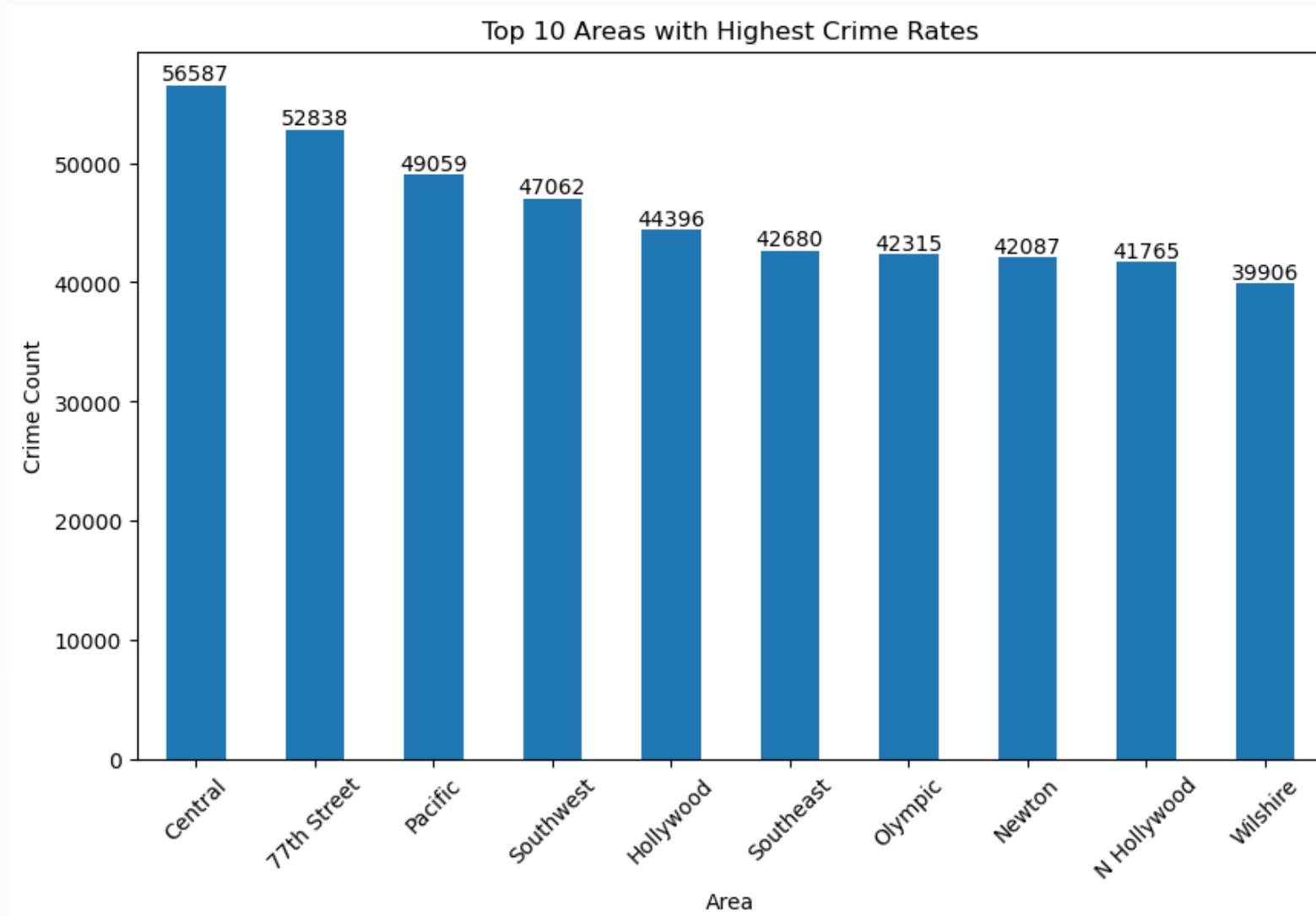
```



# GEOGRAPHICAL INSIGHTS

**Objective:** Generated a heatmap using latitude and longitude data to visualize geographical crime hotspots.

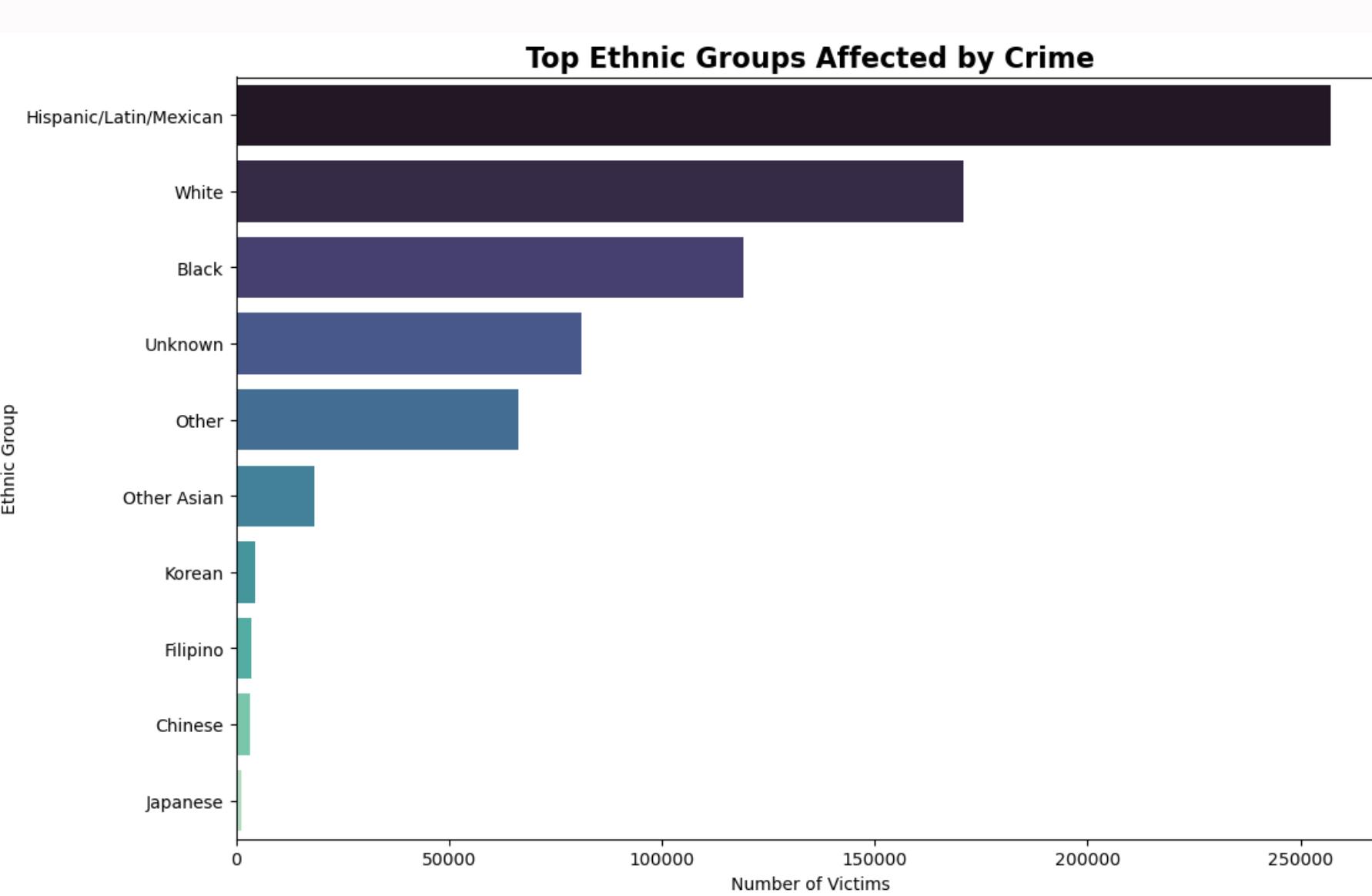
**Key Findings:** Certain areas, including Central, 77th Street, and Pacific, experience higher crime rates.



# DEMOGRAPHIC ANALYSIS

This analysis aims to explore the demography of crime victims focusing on the average age of victims, their ethnicity, and sex.

**Key Findings:** Males and the Hispanic/Latin/Mexican demographic group are more frequently victimized. The age group most affected by crime is 30-40 years.



```
# Display statistical data on the ages of crime victims
print("\nStatistical Data on Ages of Crime Victims:")

# The average age of victims
average_victims_age = cleaned_data['Vict Age'].mean()
print(f"\nAverage Age of Victims: {average_victims_age:.2f} years")

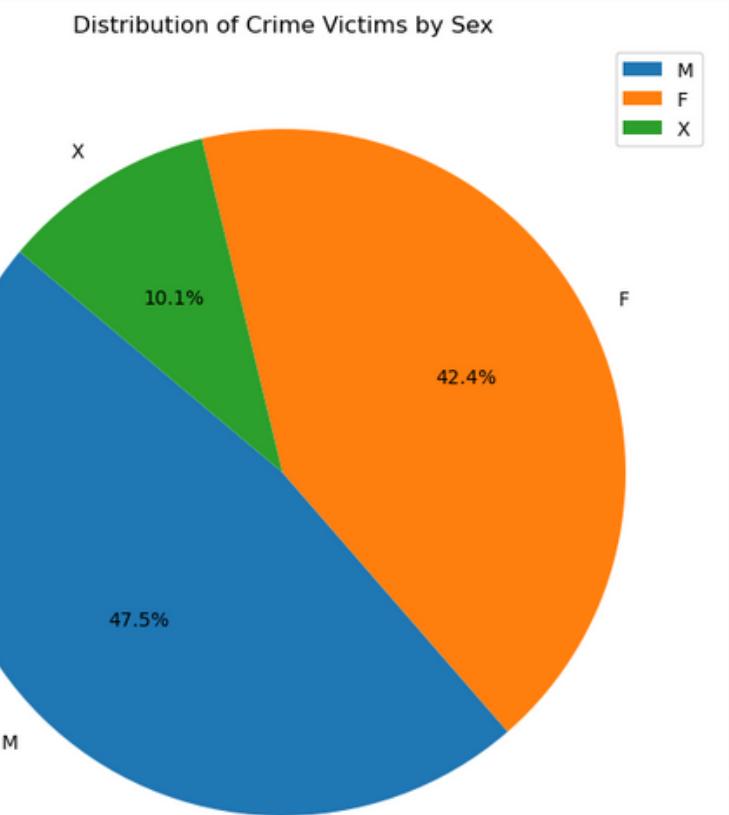
# The mode age of victims
mode_victims_age = cleaned_data['Vict Age'].mode()

# Check if there is more than one mode
if not mode_victims_age.empty:
    print(f"Mode Age(s) of Victims: {', '.join(map(str, mode_victims_age.values))} years")
else:
    print("No mode found for Victim Age.")

# The median age of victims
median_victims_age = cleaned_data['Vict Age'].median()
print(f"\nMedian Age of Victims: {median_victims_age:.2f} years")

Statistical Data on Ages of Crime Victims:

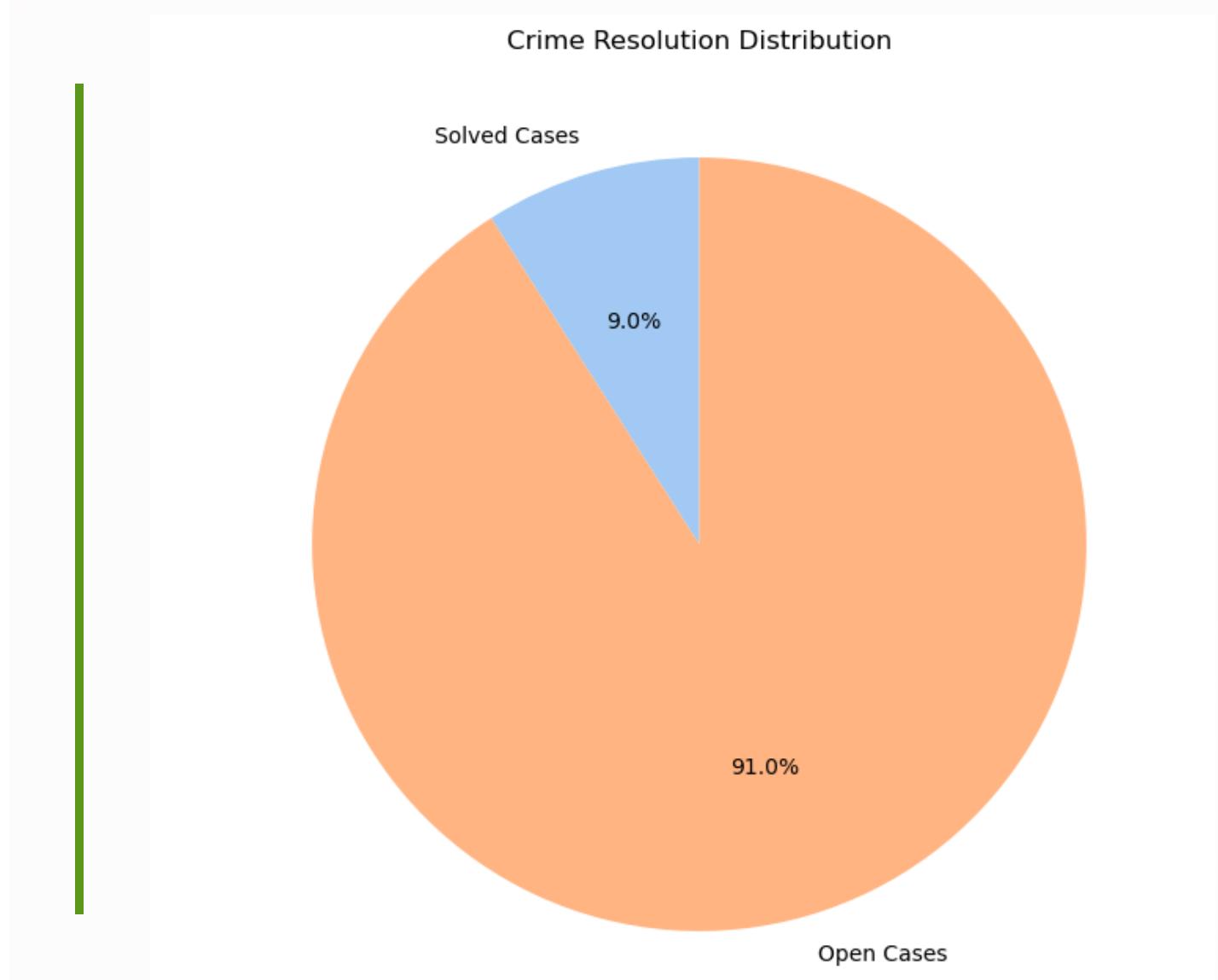
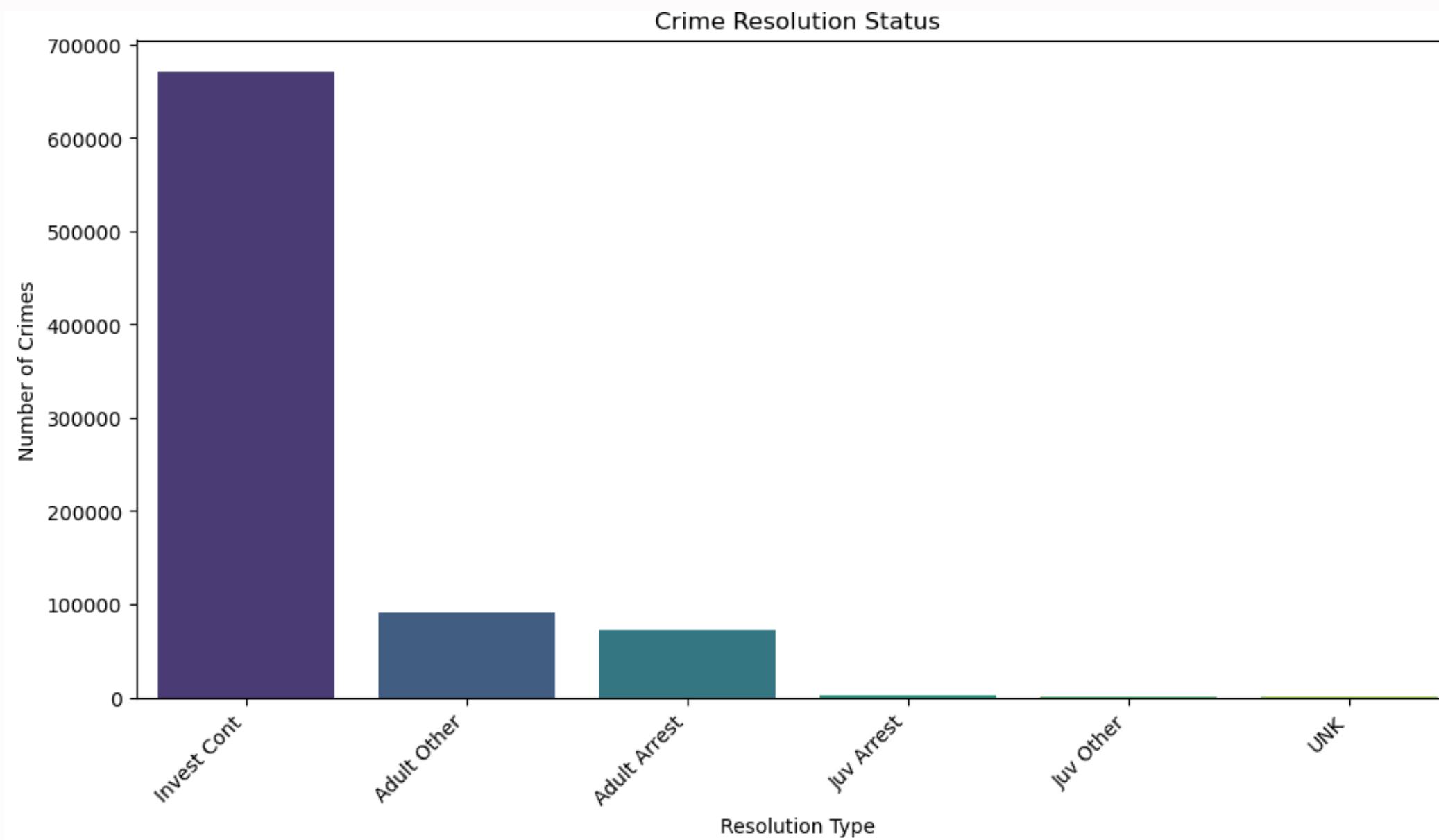
Average Age of Victims: 39.58 years
Mode Age(s) of Victims: 30.0 years
Median Age of Victims: 37.00 years
```



# RESPONSE TIMES

The aim of this analysis was to comprehend the distribution of crime resolution statuses and to assess the average response times of law enforcement to various incident types in Los Angeles County.

**Key Findings:** Vary significantly across crime types, with faster responses to more severe crimes like arson and assaults on police officers.



## DISCUSSION

### Limitations:

**Data Completeness:** The analysis relies on reported crimes, potentially omitting unreported incidents, which could skew the findings.

**Response Time Accuracy:** The calculated response times are based on reported times, which may be affected by reporting delays or data inaccuracies.

**Ethnicity and Sex Data:** The significant portion of data categorized as 'Unknown' in victim's ethnicity and sex indicates gaps in data collection and reporting.

**External Factors:** The analysis does not account for external socio-economic factors or policy changes that could impact crime trends.

**Predictive Limitations:** The report primarily focuses on descriptive statistics and does not offer predictive insights or causal analysis.

### Recommendations for Future Research:

**Socio-Economic Integration:** Incorporating socio-economic data could provide a more comprehensive understanding of the underlying factors influencing crime trends.

**Predictive Modeling:** Developing predictive models could aid in forecasting future crime trends and planning proactive crime prevention strategies.

**Improved Data Collection:** Enhancing the accuracy and completeness of data collection, especially in victim demographics, is crucial for more informed analyses.

# Thank You!



[github.com/apatino16/LA\\_Crime\\_Data\\_Analysis\\_Project](https://github.com/apatino16/LA_Crime_Data_Analysis_Project)

