

VISUALIZATION OF COMPLEX DATA_DATS_6401_12

CRIMES IN LOS ANGELES

Introduction

Our goal in this project is to investigate Los Angeles crime incidents between 2020 to 2024. In order to help with the creation of focused crime prevention strategies, our analysis looks for patterns and trends. The particular questions that we aim to address are:

1. How have crime rates changed over time?
2. Which areas in Los Angeles have the highest crime rates?
3. Did significant events (e.g., COVID-19 pandemic) impact crime rates and types?

Team Contribution

Sai Pavan Mekala: Led the data collection and initial data cleaning process.

Jagadeeshwar Kalyanapu: Focused on data analysis and Visualisations.

Kundana Chowdary Cherukuri: Developed visualizations and managed the project documentation.

Description of the Data

Using its API, we were able to retrieve the crime incident data for our project from the Los Angeles Open Data portal. We can consistently add the most recent data to our analysis because of to this method. Every record in the dataset contains detailed information about crime incidents, such as the type of crime committed and the date, time, and location of the incident. Furthermore, the data contains information about the perpetrators and victims whenever it is available. In order to comprehend and illustrate how crime dynamics change over time in Los Angeles, we can conduct a thorough analysis of crime trends, distribution across various regions, and temporal patterns through to this wealthy dataset.

Noteworthy Features

Comprehensiveness: A wide range of crime categories, including murder, vandalism, robbery, and burglary, are included in the dataset.

Detail: Every record includes comprehensive details about the incident, such as the precise time and place.

Analysis of Data Quality

Data Dictionary

1. **Date Rptd:** The date on which the crime was reported.
2. **Date Occ:** The date when the crime incident occurred.
3. **Time Occ:** The time when the crime incident occurred.
4. **Area:** Numeric code representing a specific geographical area or police precinct.
5. **Area Name:** Name of the area or police precinct.
6. **Rpt Dist No:** Reporting district number, specifies a smaller zone within the area.
7. **DR NO:** Unique identifier for each record.
8. **Crm Cd:** Crime code, numeric representation of the crime type.

9. **Crm Cd Desc:** Description of the crime.
10. **Premis Cd:** Premise code, numeric identifier of the location type where the crime occurred.
11. **Premis Desc:** Description of the premise where the crime occurred.
12. **Vict Age:** Age of the victim.
13. **Vict Descent:** Race or ethnicity of the victim.
14. **Vict Sex:** Gender of the victim.
15. **Part 1-2:** Indicates if the crime is Part 1 (more serious) or Part 2.
16. **Weapon Used Cd:** Code for the type of weapon used, if applicable.
17. **Weapon Desc:** Description of the weapon used.
18. **Status:** Current status of the case (e.g., Investigative Continuation).
19. **Status Desc:** Description of the case status.
20. **Location:** Street address or intersection where the incident occurred.
21. **Crm Cd 1:** Additional code related to the crime.

Qualitative Data

1. **Crm Cd & Crm Cd Desc:** Statistical information categorized by crime type and code. Not ranked.
2. **Area & Area Name:** Precincts or geographic areas represented by categorical data. Not ranked.
3. **Premis Cd & Premis Desc:** Statistical information categorized by kind of property used for criminal activity. Not ranked.
4. **Vict Descent:** Statistical information pertaining to the victim's racial or ethnic background. Not ranked.
5. **Vict Sex:** Statistical information pertaining to the victim's gender. Not ranked.
6. **Sections 1-2:** Classifying crimes into Part 1 (serious) or Part 2 using categorical data. This offers a severity-based ranking system.
7. **Status & Status Desc:** Categorical information outlining the crime case's present state. Not ranked.

Quantitative Data

1. **Time Occ:** Originally numerical (e.g., 2200 for 10 PM), but for easier interpretation, converted to readable time format (e.g., 22:00). It gauges the precise moment a crime is committed.
2. **Victim Age:** A quantitative measurement expressed in years for the victim's age.
3. **Weapon Used Cd:** If applicable, numerically coded quantitative data representing various weapon types used in crimes.

With different attributes like the geographic distribution of crimes, the timing of incidents, victim demographics, and information about the nature and status of the crime, each variable provides a rich dataset for analysis. These characteristics are essential for trend analysis, hotspot identification, demographic impact analysis, and evaluation of law enforcement response efficacy.

Main Analysis

The objective of our project was to examine trends, geographic distribution, and the effects of notable occurrences such as the COVID-19 pandemic in Los Angeles' crime data from 2020 to 2024. A well-suited dataset with over 932,000 records from the Los Angeles Open Data was used for this thorough analysis because of its relevance, quality, and structure.

Data Preparation and Cleaning

The dataset showed good uniformity and cleanliness at first, so not much cleaning was needed. The 'Time Occ' column was changed from numerical to a more comprehensible format (e.g., 2200 to 22:00), which was a noteworthy modification. This conversion was essential for simplifying the interpretation and analysis of the data.

Analytical Process and Challenges

The analysis was performed through various statistical and visualization techniques, including bar charts, tree maps, and temporal analyses. Making sure the geographic data was accurate was one of the main problems, as it was necessary to evaluate how crimes were distributed throughout Los Angeles.

Description of graphs

1. Crime Rates Over Time

The bar chart illustrates the significant increase in crime rates that our analysis showed would occur until 2024. Public safety concerns were aroused by this trend, which also prompted additional research into possible causes and mitigating measures.

2. Geographical Crime Distribution

We determined which areas of Los Angeles had the highest rates of crime by using a tree map. This graphic assisted in identifying regions in need of more targeted community intervention and law enforcement efforts.

3. Impact of COVID-19 on Crime Rates

Variations in crime rates and types both during and after the pandemic period demonstrated the pandemic's impact on crime patterns. Understanding how outside factors, such as a pandemic, can change the dynamics of crime, required an understanding of this analysis.

4. Crime Incident Distribution by Victim's Gender

Our investigation also looked at how crime incidents were distributed according to the gender of the victim, providing important new information about the groups that particular crimes most frequently targeted or impacted..

5. Age Trends Among Crime Victims

Investigating the age trends helped identify which age groups were more likely to become victims of crime, which in turn helped shape community support initiatives and preventative measures.

6. Weapon Usage in Crimes

An analysis of the frequency of various weapon types used in crimes revealed that knives and firearms were most often used. For law enforcement organizations to customize their approaches to crime prevention and response, this data is essential.

7. Temporal Patterns in Crime Occurrence

By determining the most common times of day or week for crimes to occur, police patrols and public safety initiatives could be more effectively planned.

8. Status of Crime Cases

In order to assess how well the local criminal justice system handles cases, we also looked at the status of crime cases.

9. Correlation Between Time of Day and Crime Types

Our research revealed that particular hours of the day were linked to particular kinds of crimes, which may be crucial information for law enforcement to use when developing time-sensitive crime prevention tactics.

10. Crime Rate Comparison by Year and Type in Los Angeles

The graphic offers a thorough summary of Los Angeles' crime trends, indicating both a rise in criminal activity and shifts in the kinds of crimes committed during the studied time.

Visualizations

In order to effectively communicate complex data insights, visual aids were essential. They simplified the understanding of Los Angeles' criminal landscape for a wide range of stakeholders, including the general public and policymakers.

Key Findings

Our analysis of Los Angeles's crime statistics from 2020 to 2024 revealed numerous crucial insights into the dynamics of the city's criminal activity, demonstrating the enormous impact of outside variables like social developments on crime rates. One important finding was the association between significant societal events, such as the COVID-19 pandemic, and variations in crime rates. The kinds of crimes that were committed during the pandemic clearly changed from pre-pandemic times, with crimes like burglaries and domestic abuse exhibiting distinct patterns.

The analysis also revealed that some Los Angeles neighborhoods had consistently higher crime rates. This recurring pattern points out potential areas for increased resource concentration by law enforcement. Police can better manage resources, concentrate community engagement efforts, and improve patrol routes by pinpointing these high-crime neighborhoods.

Another important factor that came out of our analysis was the time of day. There were specific times when certain types of crimes were more likely to happen. For instance, late nights and early mornings were the most common times for thefts and burglaries, whereas late evenings were typically the peak for assault incidents. Law enforcement organizations can better

coordinate patrols and interventions to deter similar crimes by having a better understanding of these temporal patterns.

Additionally, the information shed light on how different crime types and their frequency relate to one another as well as to public holidays and day of the week. Planning for preventive measures is made possible by the development of increasingly complex predictive models that are able to predict spikes or dips in crime.

Overall, the thorough analysis we conducted on the crime data in Los Angeles helps to plan proactive strategies for preventing and controlling crime while also providing insight into the current state of crime in the city. In the end, this can improve community safety and well-being by enabling a more focused approach to combating crime.

Conclusion

Limitations and Future Directions

1. **Accuracy of Spatial Data:** By incorporating more exact geographic information systems, future research may be able to rectify certain inaccuracies in location data.
2. **Longitudinal Studies:** Further investigation may be necessary to assess the long-term viability of crime prevention measures put in place in response to these discoveries.

Lessons Learned

1. The significance of thorough data preparation and cleaning in big datasets.
2. The efficiency with which visual aids can communicate complex data insights to a variety of audiences.