Final Project Report

Mohit Saini: saini.mo@northeastern.edu

College of Professional Studies

Northeastern University

Boston, MA (Massachusetts)

Section – I

Introduction

Crime is a pressing societal concern that impacts communities worldwide, affecting public safety, law enforcement strategies, and policy decisions. In our endeavour to understand the complexities of crime dynamics, we have undertaken an analysis of a comprehensive dataset spanning the years 2010 to 2016. This dataset provides valuable insights into various aspects of criminal activity, including the types of offenses committed, their geographical distribution, and temporal trends.

The objective of this report is to explore and interpret the findings derived from the analysis of the crime dataset. By harnessing the power of data visualization tools such as Power BI, Tableau, and Shiny R, we aim to uncover underlying patterns, trends, and correlations within the data. Through visualization, we seek to transform raw data into actionable insights that can inform decision-making processes and facilitate evidence-based strategies for crime prevention and law enforcement.

The dataset under examination contains a diverse array of variables, ranging from the nature of offenses to demographic characteristics of offenders and victims. By dissecting and visualizing this data, we endeavour to shed light on the spatial and temporal dynamics of crime, identify hotspots of criminal activity, and discern patterns that may inform targeted interventions.

This report is structured to provide a comprehensive overview of our analytical approach, findings, and their implications. We begin with an overview of the crime dataset, followed by a detailed exploration of our visualization efforts using Power BI, Tableau, and Shiny R. Finally, we discuss the implications of our findings and their relevance for future research and policy development in the realm of crime prevention and law enforcement.

By undertaking this analysis, we aim to contribute to the collective understanding of crime dynamics and advance efforts to promote public safety and well-being in our communities.

Section - II

Overview of Crime Dataset (2010-2016)

A) Categorical Variable Columns:

- **AREA NAME:** Location where the crime occurred.
- **CRIME CODE:** Numerical code representing the type of crime.
- **Victims Sex:** Gender of the victim.
- **Victims Descent:** Ethnicity of the victim.
- **Premise Code:** Type of premise where the crime occurred.
- **Weapon Used Cd:** Type of weapon used in the crime.
- **Status:** Status of the case, such as Adult Arrest or Investigative Continuation.

B) Integer Columns:

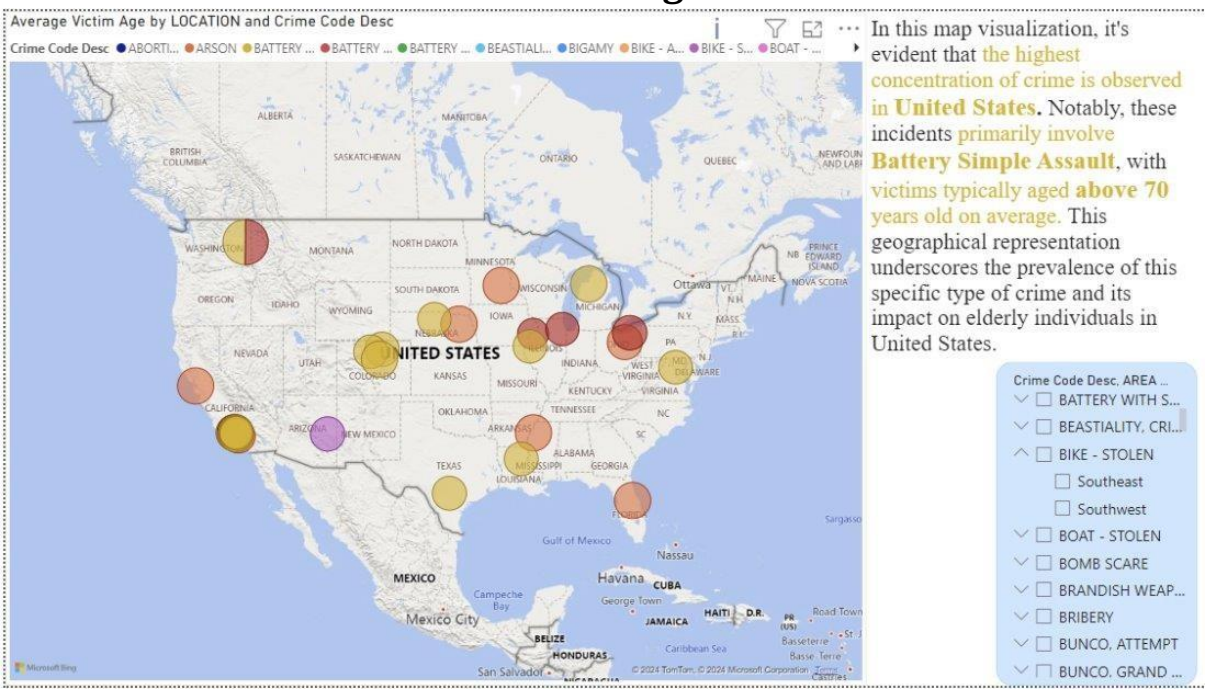
- **FILE_NO:** Unique identifier for each crime record.
- **Victims Age:** Age of the victim.
- **LAT:** Latitude coordinate of the crime location.
- **LON:** Longitude coordinate of the crime location.

C) String Columns:

- **DATE REPORTED:** Date when the crime was reported.
- **DATE OCCURRED:** Date when the crime occurred.
- **TIME OCCURRED:** Time when the crime occurred.
- **Crime Code Desc:** Description of the crime based on the crime code.
- **Premise Description:** Description of the premise where the crime occurred.
- **Weapon Description:** Description of the weapon used in the crime.
- **Status Description:** Description of the status of the case.
- **LOCATION:** Updated location description of the crime.

Section-III

Visualization using Power BI



[Fig 3.1 "Geospatial Insight: Battery Simple Assault Crimes and Elderly Victims Across the United States"]

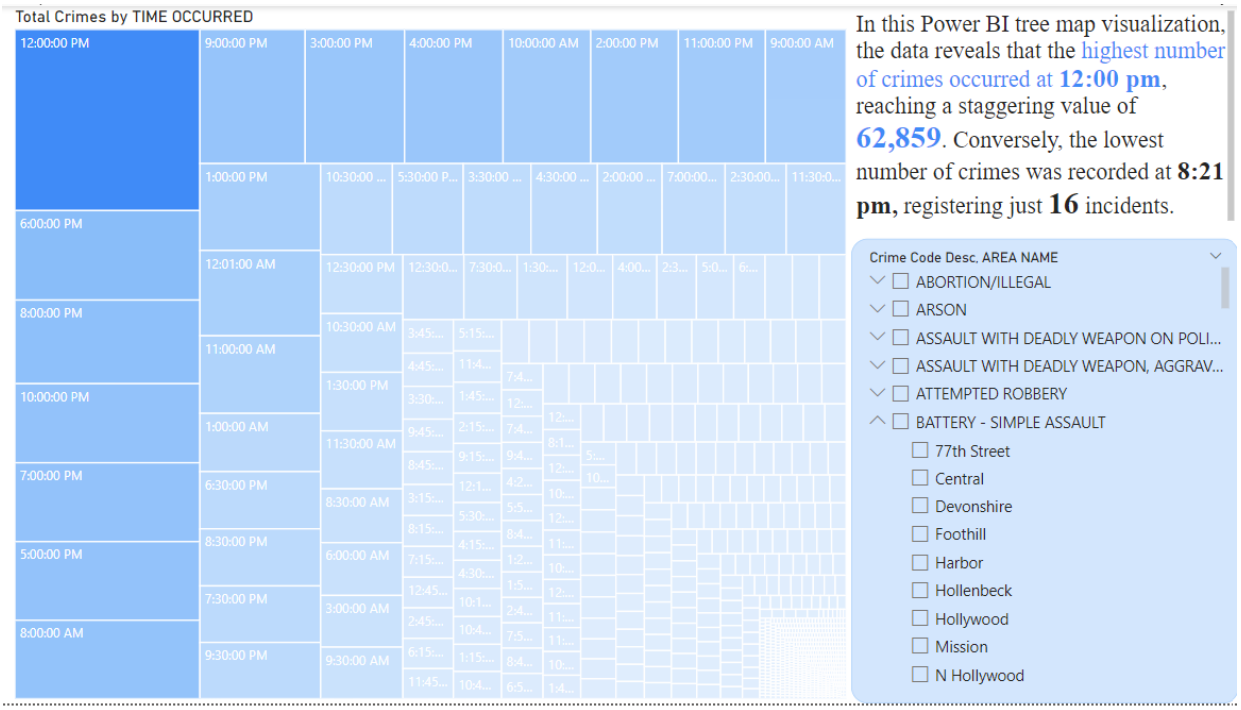
Description:

This map visualization highlights the geographical distribution of Battery Simple Assault crimes involving elderly victims across the United States. The image showcases the highest concentration of such crimes, revealing notable hotspots where incidents are prevalent. Additionally, it emphasizes the predominant crime type, which is Battery Simple Assault, and provides insights into the average age of victims, indicating that the majority are above 70 years old. This visualization offers valuable geospatial insights into the nature and scope of crimes against elderly individuals, informing targeted interventions and law enforcement strategies to safeguard vulnerable populations.

Use of Visualization:

This geospatial visualization serves as a powerful tool for identifying areas across the United States where Battery Simple Assault crimes involving elderly victims are most prevalent. By pinpointing these hotspots, law enforcement agencies and policymakers can allocate resources more effectively to combat crime and enhance public safety. Furthermore, this visualization enables authorities to develop targeted interventions and support services tailored to the needs of elderly victims in these high-risk areas. Understanding the spatial distribution of

such crimes is essential for devising proactive strategies aimed at reducing victimization rates and fostering safer communities for vulnerable populations.



[Fig 3.2 "Crime Density Analysis: Unveiling Peak and Off-Peak Hours with Power BI Tree Map Visualization"]

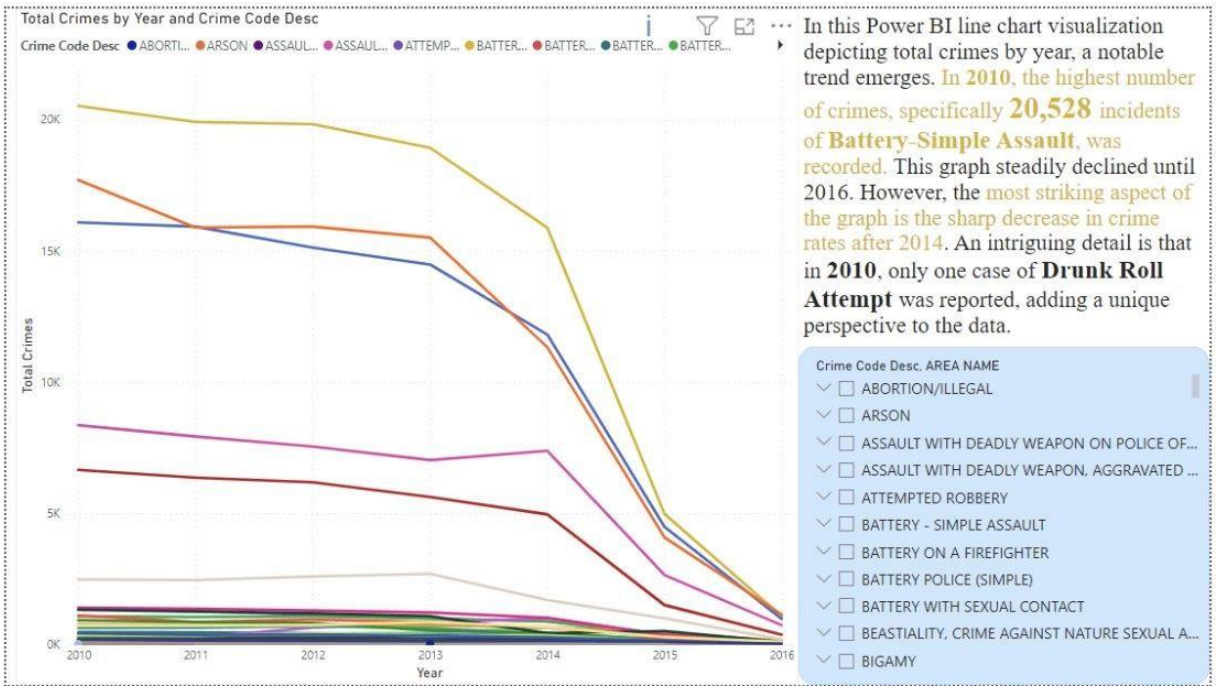
Description:

This Power BI Tree Map visualization provides insights into crime density by unveiling peak and off-peak hours of criminal activity. The image reveals that the peak crime hour occurs at 12:00 pm, with a staggering number of 62,859 crimes reported during this time period. In contrast, the off-peak crime hour is observed at 8:21 pm, with a significantly lower number of 16 reported crimes. This visualization offers valuable insights into the temporal patterns of criminal activity, aiding law enforcement agencies and policymakers in allocating resources and implementing targeted interventions to address peak crime hours effectively.

Use of Visualization:

Understanding these temporal patterns is crucial for law enforcement agencies and policymakers to allocate resources effectively and implement targeted interventions during peak crime hours. By deploying strategies such as increased patrols or community outreach programs during peak hours, authorities can enhance public safety and reduce the incidence of criminal activity in vulnerable areas. Additionally, this analysis can inform the development of crime prevention

initiatives aimed at reducing overall crime rates and improving the quality of life in affected communities.



[Fig 3.3 "Analyzing Crime Trends: A Yearly Overview with Emphasis on Battery-Simple Assault and Drunk Roll Attempt in Power BI Line Chart"]

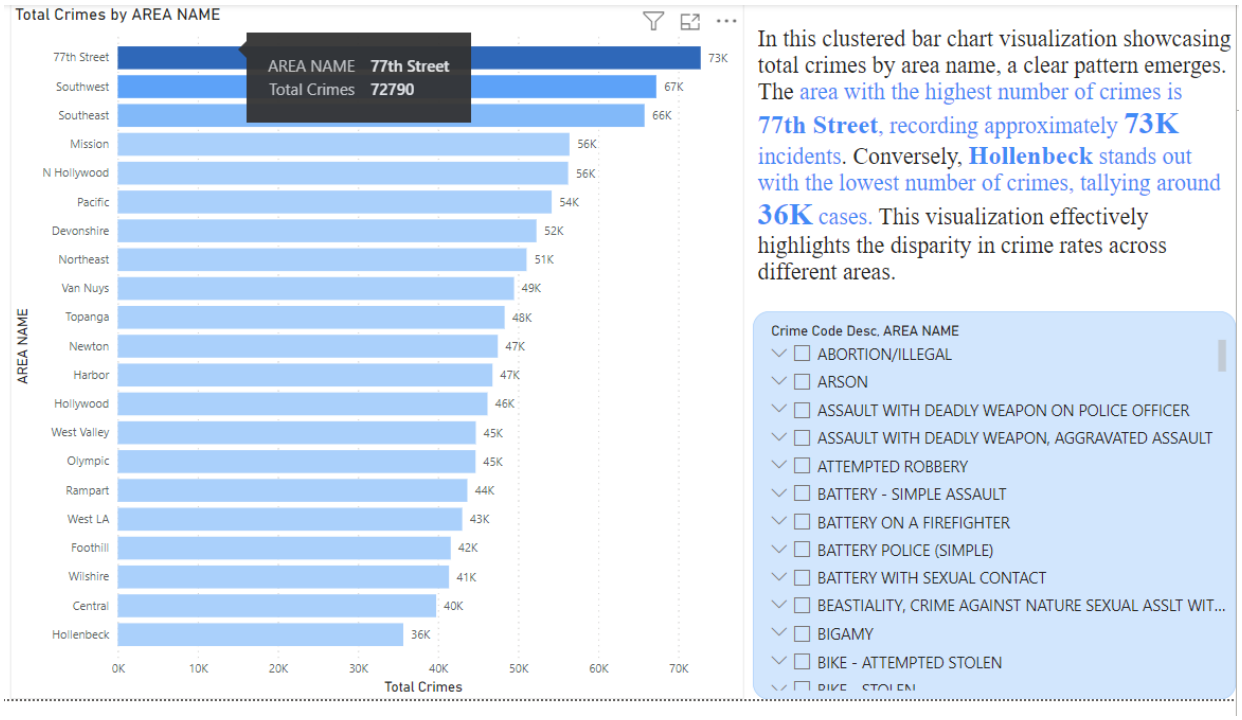
Description:

This Power BI Line Chart provides a comprehensive yearly overview of crime trends, with a specific focus on Battery-Simple Assault and Drunk Roll Attempt incidents. The chart highlights a notable trend of declining total crimes over the years, indicating a positive shift in public safety. In 2010, the highest number of crimes recorded was 20,528 incidents of Battery-Simple Assault, while only one case of Drunk Roll Attempt was reported. The analysis reveals a steady decrease in crime rates until 2016, with a sharp decline noticeable after 2014, suggesting significant changes in crime rates during this period.

Use of Image:

This visualization serves as a powerful analytical tool for law enforcement agencies, policymakers, and researchers to understand long-term crime trends and patterns. By identifying fluctuations in crime rates over the years and highlighting specific crime types, such as Battery-Simple Assault and Drunk Roll Attempt, authorities can develop targeted interventions and allocate resources effectively to address emerging challenges and enhance public safety. Additionally, this

analysis enables authorities to monitor the effectiveness of crime prevention initiatives and track progress towards reducing crime rates, ultimately contributing to the creation of safer communities.



[Fig 3.4 Comparing Crime Rates Across Areas: Insights from Clustered Bar Chart Visualization"]

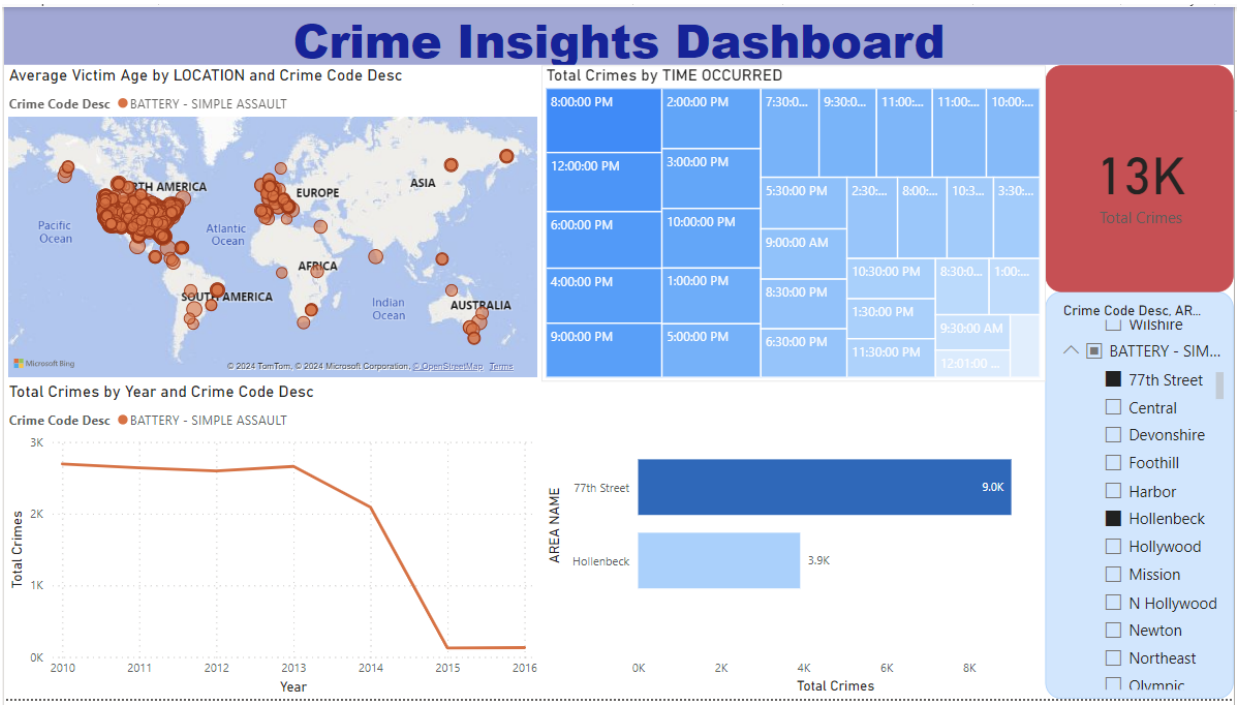
Description:

This clustered bar chart visualization provides insights into the comparison of crime rates across different areas, focusing on total crimes by Area Name. The analysis highlights the highest and lowest crime areas, with 77th Street recording approximately 73,000 incidents, making it the highest crime area, while Hollenbeck reports around 36,000 incidents, indicating the lowest crime area. The visualization effectively communicates the varying levels of crime prevalence across different regions, enabling authorities to identify areas requiring targeted interventions and allocate resources accordingly.

Use of Image:

The clustered bar chart offers a visual representation of crime rates across different areas, facilitating easy comparison and identification of high and low crime areas. Law enforcement agencies can use this visualization to prioritize resource allocation and deploy targeted crime prevention strategies in high crime areas such as 77th Street. Additionally, policymakers can leverage this analysis to understand regional variations in crime prevalence and develop tailored

interventions aimed at improving public safety and reducing crime rates in vulnerable communities.



[Fig 3.5 "Power BI Dashboard: Insights into Crime Trends and Distribution"]

Description:

This Power BI Dashboard provides a comprehensive analysis of crime trends and distribution, featuring four key visualizations and a slicer control for enhanced data filtering.

First Visualization:

- The first visualization focuses on identifying areas with the maximum crime prevalence in North America.

Second Graph:

- The second graph highlights the peak occurrence of crimes at 8:00 pm, offering insights into temporal patterns of criminal activity.

Third Graph:

- The third graph depicts the decreasing trend of crime, particularly Battery Simple Assault, from 2010 to 2016, facilitating a longitudinal analysis of crime trends.

Fourth Graph:

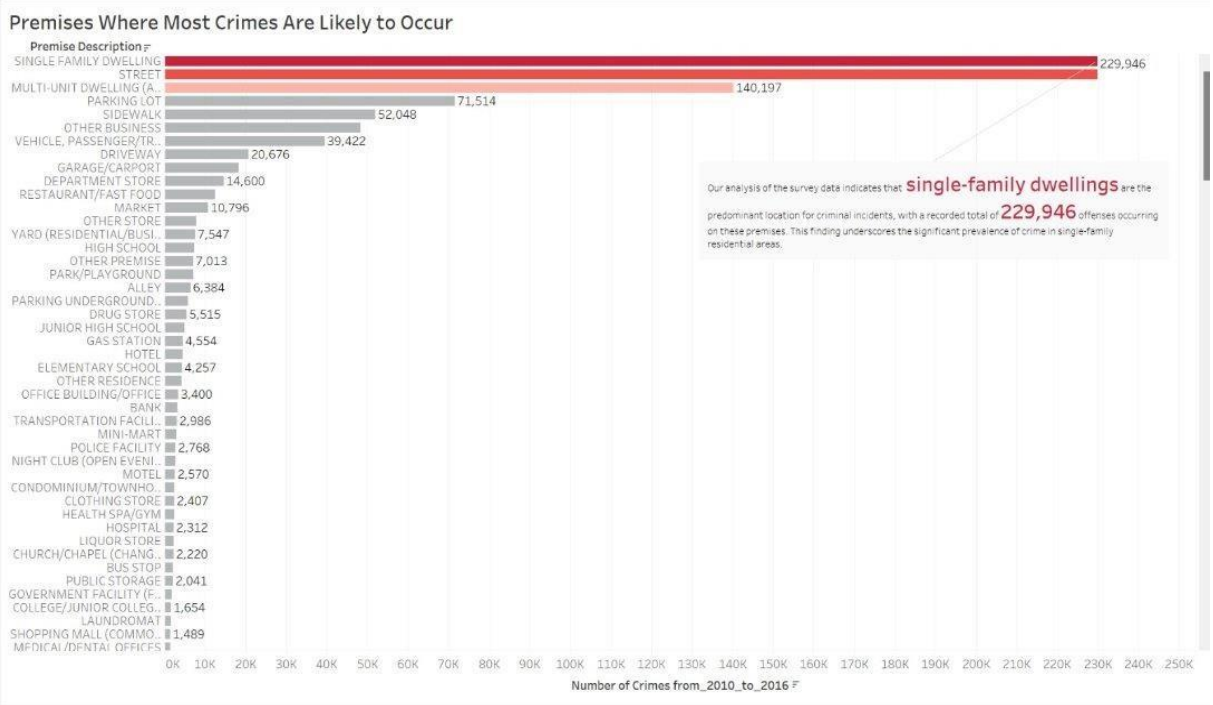
- The fourth graph delves into the distribution of crimes by area, enabling authorities to identify geographical hotspots of criminal activity.

Slicer Control:

- The slicer control feature empowers users to filter and control the data displayed in all graphs, enhancing interactive data exploration and analysis.

Section – IV

Graphs Using Tableau



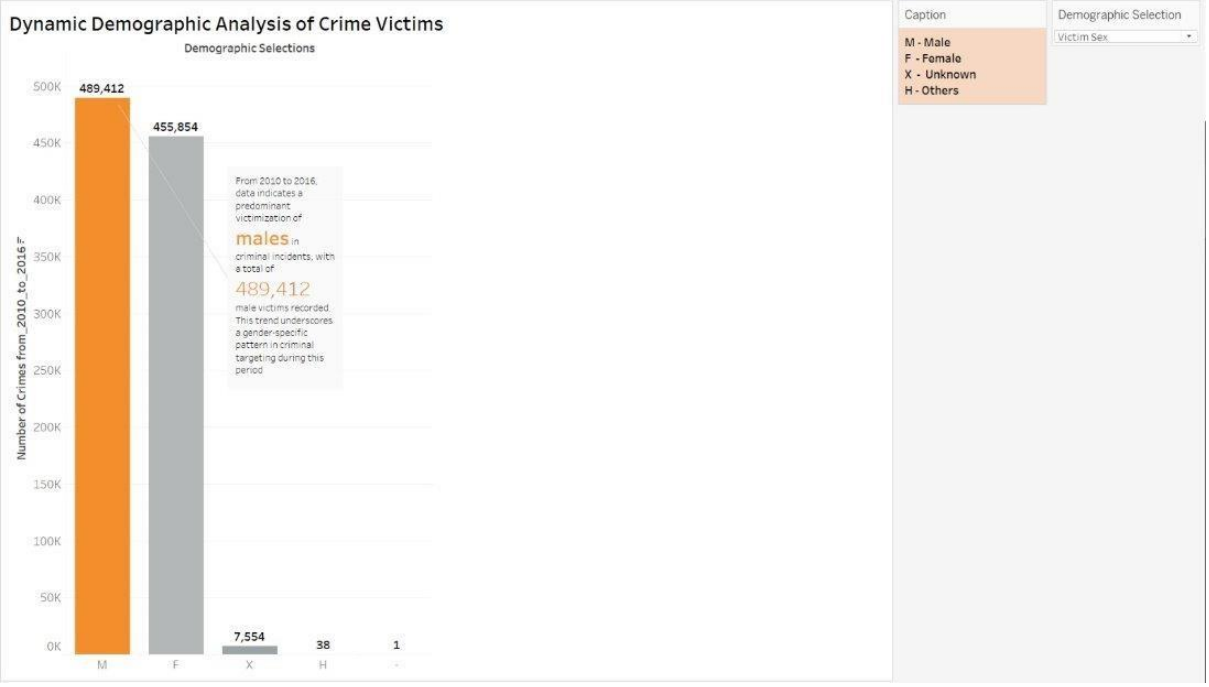
[Fig 4.1 "Premises where most Crimes are likely to occur"]

Description:

This visualization focuses on identifying the primary locations for criminal incidents, with a specific emphasis on single-family dwellings. The analysis reveals that single-family homes are the predominant locations for criminal incidents, with a total of 229,946 offenses recorded in these premises. This data highlights the significant prevalence of crime in single-family residential areas, underscoring the importance of addressing security measures and crime prevention strategies in these locations to enhance community safety.

Use of Visualization:

By highlighting single-family dwellings as the primary locations for criminal incidents, this visualization enables law enforcement agencies and community authorities to prioritize resources and interventions effectively. Understanding the prevalence of crime in residential areas empowers policymakers to implement targeted strategies aimed at improving security measures and fostering safer neighbourhoods. Additionally, this analysis facilitates informed decision-making regarding resource allocation for crime prevention efforts, ultimately contributing to the creation of more secure and resilient communities.



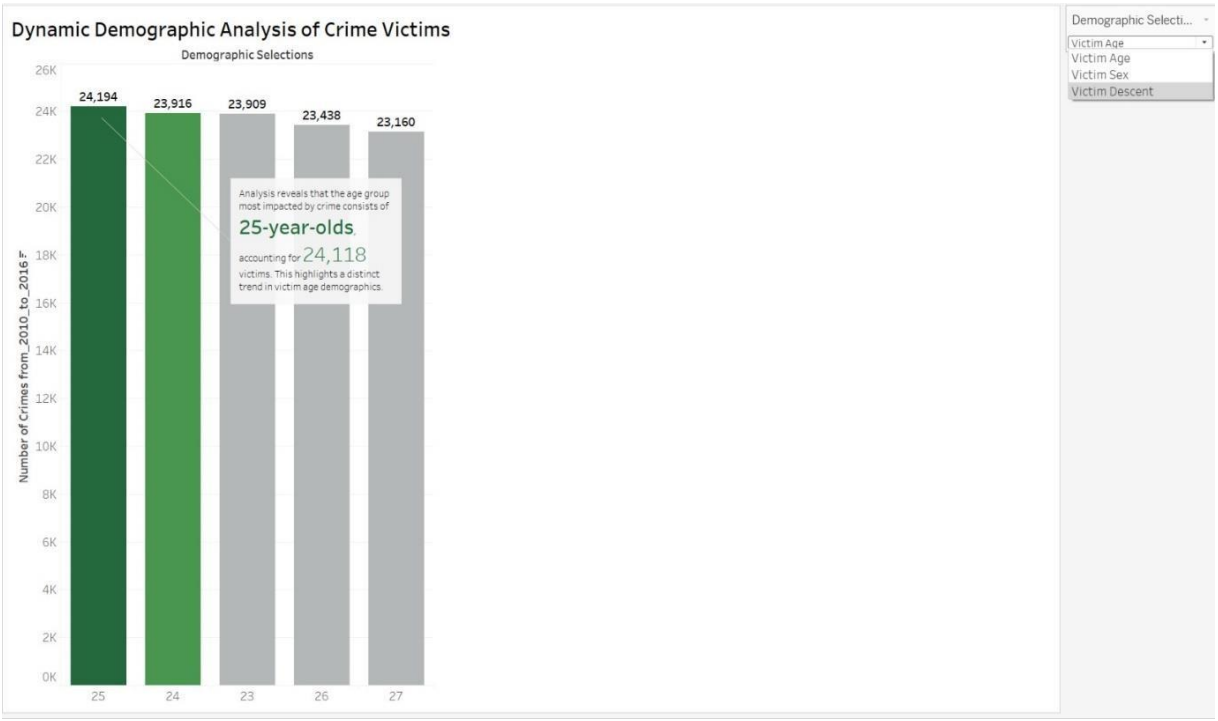
[Fig 4.2 "Dynamic Demographic Analysis of Crime Victims"]

Description:

This analysis provides a dynamic demographic overview of crime victims, covering the period from 2010 to 2016. The data reveals a predominant victimization of males in criminal incidents, with a total of 489,412 male victims recorded during this period. This gender-specific pattern in criminal targeting underscores the need for understanding and addressing the underlying factors contributing to such victimization trends over time.

Use of Visualization:

By illustrating the demographic composition of crime victims over a seven-year period, this visualization offers valuable insights into patterns of victimization and vulnerability within the population. Law enforcement agencies and policymakers can utilize this information to develop targeted intervention strategies aimed at reducing victimization rates among specific demographic groups, such as males. Additionally, this analysis facilitates a deeper understanding of the dynamics of crime victimization, enabling authorities to implement preventive measures and support services tailored to the needs of affected communities.



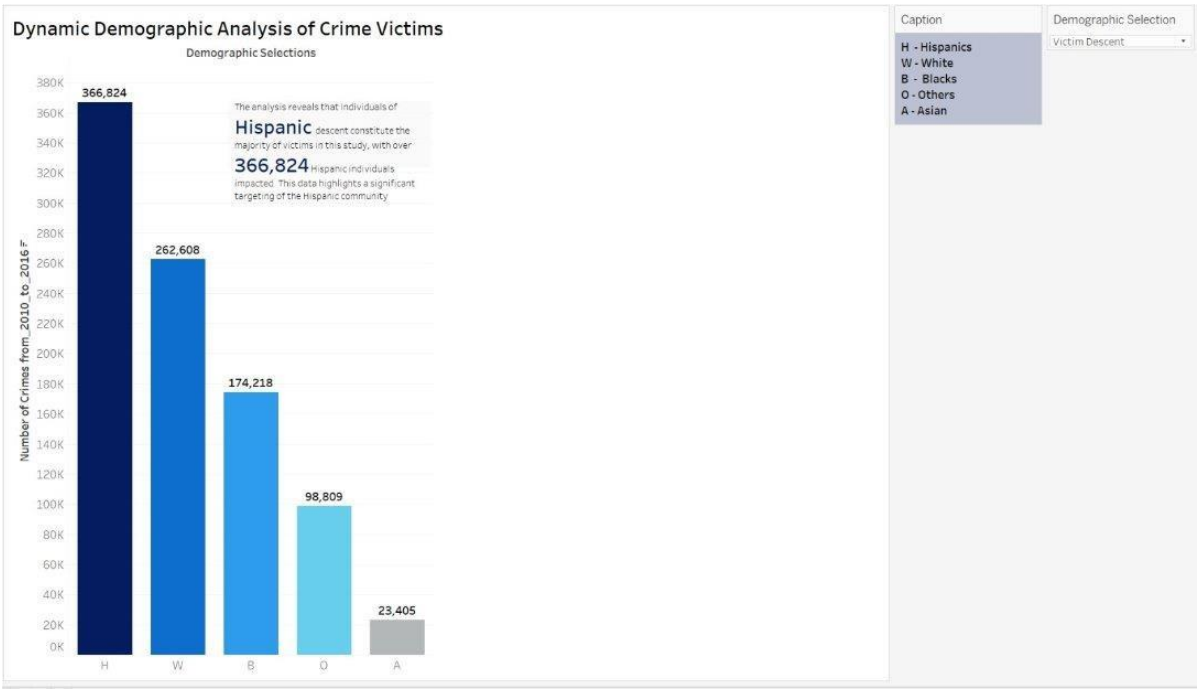
[Fig 4.3 "Demographic Impact of Crime: Focus on Age Groups"]

Description:

This analysis focuses on identifying the age group most impacted by crime, revealing that 25-year-olds are the most affected demographic. A total of 24,118 victims belong to the 25-year-old age group, indicating a significant impact on individuals within this age range. This data highlights a distinct trend in the distribution of crime victims across different age demographics, shedding light on the specific vulnerabilities and risk factors associated with particular age groups.

Use of Visualization:

By examining the demographic impact of crime across age groups, this visualization provides valuable insights into the vulnerability of different segments of the population. Law enforcement agencies and policymakers can use this information to develop targeted intervention strategies aimed at addressing the unique needs and challenges faced by specific age groups, such as 25-year-olds. Additionally, this analysis facilitates a deeper understanding of the dynamics of crime victimization, enabling authorities to implement preventive measures and support services tailored to the needs of affected age demographics.



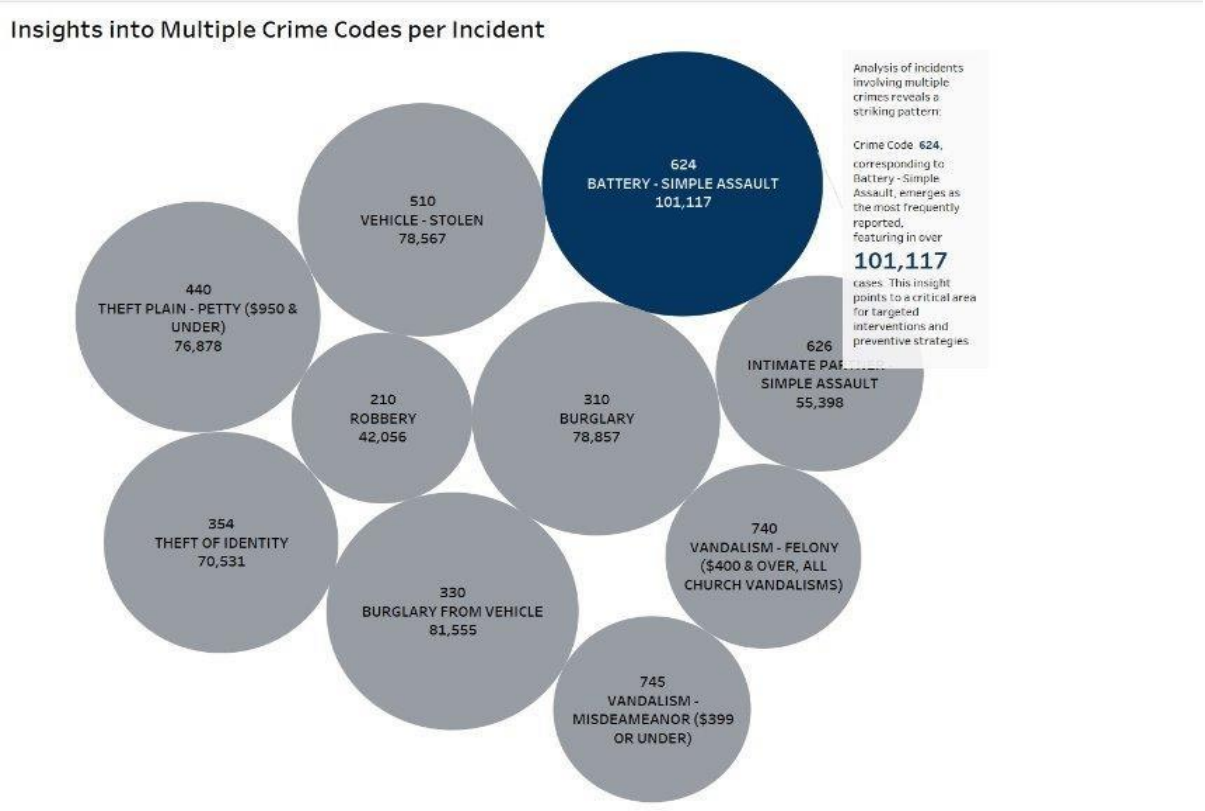
[Fig 4.4 "Dynamic Demographic Analysis of Crime Victims"]

Description:

This dynamic demographic analysis of crime victims reveals that individuals of Hispanic descent constitute the majority of victims in the study. The analysis indicates a significant targeting of the Hispanic community in criminal incidents, with over 366,824 Hispanic individuals impacted by crime according to the data. This underscores the importance of understanding and addressing the specific vulnerabilities and challenges faced by the Hispanic community in the context of crime victimization.

Use of Visualization:

By highlighting the dominant victim demographic and the significant targeting of the Hispanic community in criminal incidents, this visualization provides crucial insights for law enforcement agencies and policymakers. Understanding the demographic composition of crime victims enables authorities to develop targeted intervention strategies aimed at addressing the unique needs and challenges faced by specific communities, such as the Hispanic population. Additionally, this analysis facilitates efforts to enhance community engagement and trust-building initiatives within the Hispanic community, ultimately contributing to more effective crime prevention and victim support efforts.



[Fig 4. 5 "Insights into Multiple Crime Codes per Incident"]

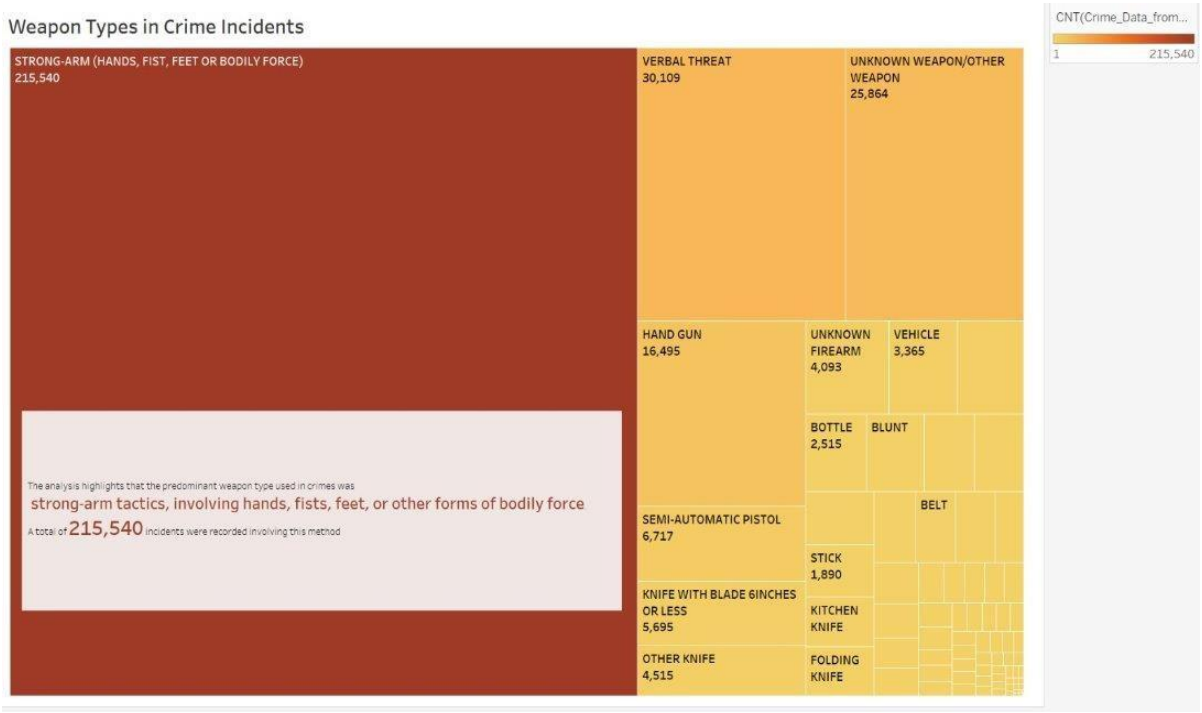
Description:

This analysis focuses on incidents involving multiple crimes and sheds light on the most frequently reported crime code within these incidents. Crime Code 624, corresponding to Battery - Simple Assault, emerges as the most frequently reported, with over 101,117 cases featuring this code. This insight underscores the critical need for targeted interventions and preventive strategies in addressing

Crime Code 624 incidents, emphasizing the importance of addressing this specific crime type to enhance public safety and reduce victimization rates.

Use of Visualization:

This visualization provides crucial insights for law enforcement agencies and policymakers by highlighting the prevalence of Crime Code 624 within incidents involving multiple crimes. Understanding the frequency of specific crime types within these incidents allows authorities to allocate resources effectively and develop tailored intervention strategies to address the underlying causes and risk factors associated with such crimes. By focusing on targeted interventions and preventive measures, authorities can work towards reducing the occurrence of Crime Code 624 incidents and improving overall community safety.



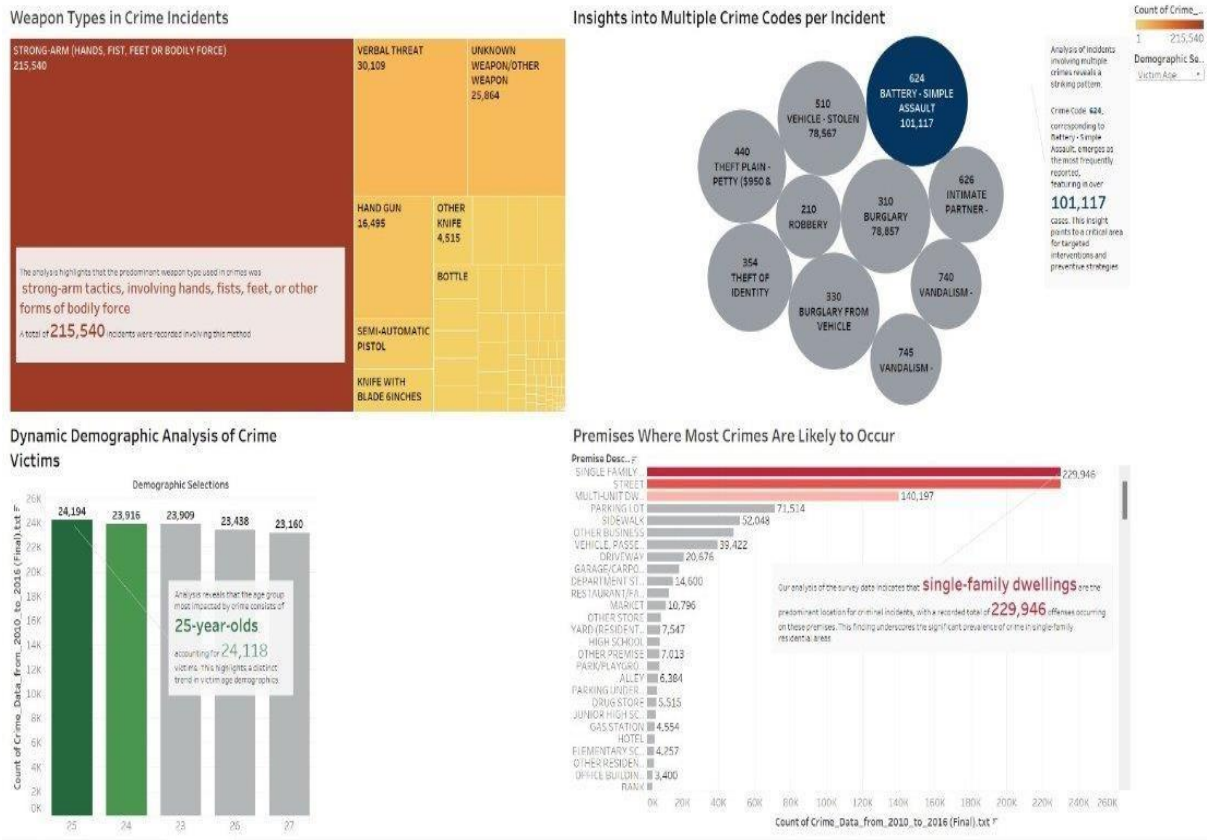
[Fig 4.6 "Weapon Types in Crime Incidents"]

Description:

This analysis focuses on weapon types used in crime incidents, highlighting strong-arm tactics as the predominant method. Strong-arm tactics involve the use of hands, fists, feet, or other forms of bodily force. A total of 215,540 incidents were recorded involving this method, indicating its prevalence in criminal incidents.

Use of Visualization:

By emphasizing the predominant use of strong-arm tactics in crime incidents, this visualization provides valuable insights for law enforcement agencies and policymakers. Understanding the prevalence of specific weapon types allows authorities to develop targeted intervention strategies aimed at addressing the underlying causes and risk factors associated with such crimes. Additionally, this analysis facilitates efforts to implement preventive measures and enhance community safety by addressing the use of strong-arm tactics in criminal incidents.



[Fig 4.7 "Dashboard in Tableau: Insights into Crime Trends"]

Description:

This Tableau dashboard provides a comprehensive overview of crime trends, featuring multiple visualizations focusing on different aspects of crime incidents.

Predominant Weapon Type Map:

- Strong-arm tactics emerge as the primary weapon used in crimes, involving bodily force. This visualization highlights the prevalence of physical violence in criminal incidents.

Crime Code Analysis Map:

- Crime Code 624 (Battery - Simple Assault) is identified as the most frequently reported crime code, necessitating targeted interventions. This analysis underscores the importance of addressing incidents related to simple assault to improve public safety.

Victim Demographics Map:

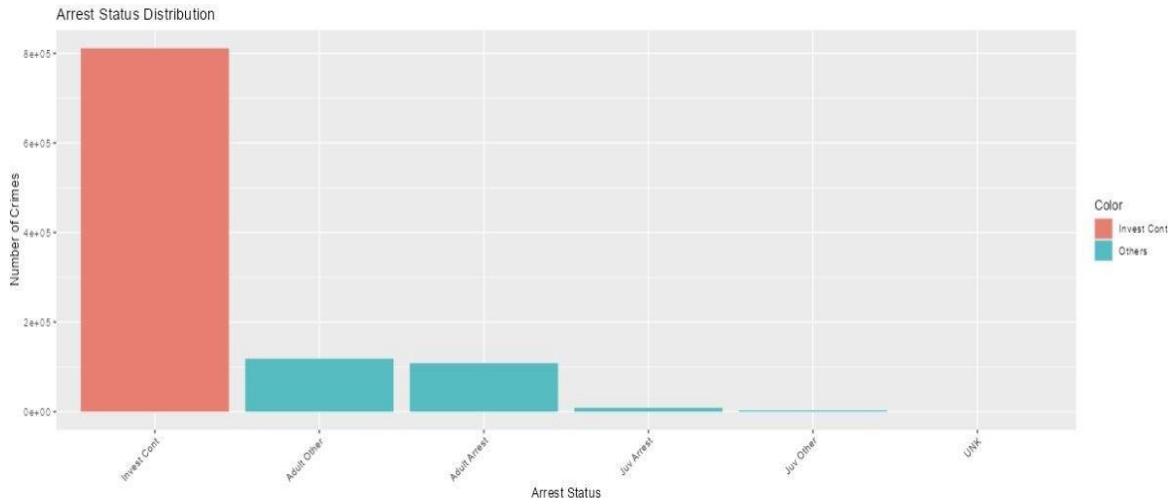
- Hispanic individuals are identified as the majority victims in crime incidents, indicating significant targeting in criminal activities. This visualization sheds light on demographic disparities in victimization rates.

Location of Incidents Map:

- Single-family dwellings are revealed as the predominant locations for criminal incidents, highlighting a substantial prevalence of crime in residential areas. This insight emphasizes the importance of addressing security measures in residential neighbourhoods to enhance community safety.

Section – V

Graphs using shiny R



[Fig 5.1 Arrest status description]

Description:

This analysis focuses on the status of arrests in criminal cases, highlighting the number of cases under investigation and those involving adult arrests. Approximately 800,000 cases are still under investigation, indicating ongoing legal processes for a significant number of cases. In contrast, adults make up around 100,000 arrests, suggesting resolved cases where individuals have been apprehended. The investigation status significantly outweighs the number of arrests, underscoring the ongoing legal processes for a large proportion of cases.

Use of Visualization:

By examining the status of arrests in criminal cases, this analysis provides valuable insights for law enforcement agencies and legal authorities. Understanding the distribution of cases under investigation and resolved cases allows stakeholders to prioritize resources and allocate manpower effectively to expedite legal processes and enhance case resolution rates. Additionally, this analysis facilitates efforts to streamline legal procedures and improve the efficiency of criminal investigations, ultimately contributing to the administration of justice and the maintenance of law and order in communities.


```
# A tibble: 10 × 2
   `Crime Code Desc`                                        Count
   <chr>                                                    <int>
 1 THEFT OF IDENTITY                                         7231
 2 BURGLARY FROM VEHICLE                                     1814
 3 VEHICLE - STOLEN                                          1478
 4 VANDALISM - FELONY ($400 & OVER, ALL CHURCH VANDALISMS)   1040
 5 BATTERY - SIMPLE ASSAULT                                  1010
 6 THEFT-GRAND ($950.01 & OVER)EXCPT,GUNS,FOWL,LIVESTK,PROD  1004
 7 THEFT PLAIN - PETTY ($950 & UNDER)                         955
 8 BURGLARY                                                   954
 9 VANDALISM - MISDEAMEANOR ($399 OR UNDER)                   947
10 THEFT FROM MOTOR VEHICLE - PETTY ($950 & UNDER)            780
```

[Fig 5.2 "Crime Count by Area and Top Crime by Time of Day"]

Description:

This analysis, conducted using Shiny R, provides insights into crime counts by area and the top crime observed by time of day.

Maximum and Minimum Crime Observations by Area:

- The maximum crime is observed in 77th Street, with approximately 70,000 incidents recorded.
- In contrast, the minimum crime is observed in Hollenbeck, with around 36,000 incidents reported.

Top Crime by Time of Day:

- During the time range of 12:00 PM to 5:00 PM, Theft of Identity is the most frequently observed crime, with 7,231 counts.
- Conversely, Theft from Motor Vehicle is the least observed crime during this time range, with 780 counts.

Use of Visualization:

By utilizing Shiny R, this analysis provides interactive visualizations that enable stakeholders to explore crime counts by area and understand the top crimes observed during specific time periods. Law enforcement agencies and policymakers can use this information to allocate resources effectively and implement targeted interventions to address prevalent crimes in high-risk areas and during peak times of criminal activity. Additionally, this analysis facilitates evidence-based decision-making and enhances situational awareness, ultimately contributing to more effective crime prevention and law enforcement efforts.

Section – VI

Conclusion

In conclusion, the comprehensive analysis presented in this report underscores the importance of leveraging data-driven approaches to address crime challenges and enhance public safety. By continuously monitoring crime trends, identifying patterns, and detecting anomalies, law enforcement agencies can proactively develop strategies to combat criminal activities effectively.

Moreover, the insights derived from this analysis facilitate optimized resource allocation, enabling authorities to focus efforts on crime hotspots, target

demographics, and prevalent crime types in different locations. This strategic allocation of resources not only improves law enforcement efficacy but also maximizes the impact of crime prevention initiatives.

Furthermore, the integration of data-driven insights into policymaking processes empowers local authorities, policymakers, and law enforcement agencies to make informed decisions. By presenting actionable insights derived from crime data, policymakers can develop evidence-based policies that address the underlying causes of crime and promote community safety.

Additionally, the utilization of predictive analytics offers a forward-thinking approach to crime prevention. By developing predictive models using historical crime data, authorities can forecast future crime rates, prioritize preventive measures, and allocate resources efficiently, thereby staying ahead of emerging crime trends.

Finally, the creation of informative dashboards and visualizations plays a crucial role in enhancing public awareness, fostering community engagement, and supporting crime prevention initiatives. By disseminating relevant information to the public, communities can actively participate in crime prevention efforts and collaborate with law enforcement agencies to create safer environments for all.

In essence, the insights generated from this analysis serve as a foundation for proactive law enforcement strategies, informed policy decisions, and collaborative efforts between authorities and communities to address crime effectively and build safer, more resilient societies.

References

- *17 Introduction to Shiny | Reproducible Research Techniques for Synthesis*. (n.d.). https://learning.nceas.ucsb.edu/2019-11-RRCourse/introduction-to-shiny.html
- *Build dashboards for accessibility*. (n.d.). Tableau. https://help.tableau.com/current/pro/desktop/en-us/accessibility_dashboards.htm
- *Crime Data from 2010 to 2019 | Los Angeles - Open Data Portal*. (2024, March 14). https://data.lacity.org/Public-Safety/Crime-Data-from-2010-to-2019/63jg-8b9z/about_data

- Mihart. (2023, December 26). *TreeMaps in Power BI - Power BI*. Microsoft Learn. https://learn.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-treemaps?tabs=powerbi-desktop