Crime Trends in Chicago, USA from 2020 to 2023

Team members:

1. Fatima Aghapourasl
2. Kareem Albeetar
3. Chun Lam Cheung
4. Penelope De Freitas
5. Eltigani Hamadelniel Elfatih Hamadelniel
6. Yi Liu
7. Jeffrey Ching Lok Ng
8. Ivy Tsai

Table of Contents

1.0 Introduction

Chicago Analytica successfully concluded the latest phase of our project collaboration with the Chicago Police, where we conducted advanced analyses and visualizations on crime trends spanning the years 2020 to 2023. Our approach involved employing sophisticated techniques such as Pareto analysis, clustering, comparisons, and correlations to unveil valuable insights aimed at enhancing decision-making regarding crimes in Chicago. Behind the scenes, we leveraged Python algorithms, as well as Tableau parameters and calculated fields, to ensure a robust and comprehensive analytical foundation.

Furthermore, we employed Dashboards and Scorecards to concisely showcase summaries of crucial information gleaned from our findings. Moreover, we conducted comprehensive root cause analyses to explore the underlying issues, putting forth potential solutions to address the identified problems. We anticipate that our stakeholders will transform this report into actionable recommendations for effective implementation.

2.0 Assumptions about the Data

Attribute	Description	Data Type
Case Number	Assigned Number for a case	String
IUCR	Illinois Uniform Crime Reporting (IUCR) Codes for the City of Chicago	String
Crime Type	Type of crime	String
Date	The date of the crime report	Date/Time
Arrested	Indicates whether or not an arrest was made for the reported crime	True/False
Domestic	Indicates whether or not a crime occurred in someone's house	True/False
Community Area	Indicates the community area where the incident occurred	Int
Community Name	Associated name of Community Area.	String
Location Description	Description of the location of where the incident occurred.	String
X-coordinate	Corresponds to longitude (Geodata point)	Number
Y-coordinate	Corresponds to latitude (Geodata point)	Number
Crime Report Rate	Number of Each Crime Type's Arrest / Total of All Arrest. Calculated from attribute: Arrested.	Percentage
Crime Arrest Rate	Number of 'True' Arrests / Total of Arrest. Calculated from attribute: Arrested.	Percentage
Crime Rate Q3	Number of Case Number during Q3 2023 / The population in Chicago 2023(KPI) Calculated from attribute: Case Number	Percentage
Crime Rate Q2	Number of Case Number during Q2 2023 / The population in Chicago 2023 Calculated from attribute: Case Number	Percentage

Attribute	Description	Data Type
Arrest Rate Q3	Number of 'True' Arrest during Q3 2023 / The population in Chicago 2023(KPI) Calculated from attribute: Arrested	Percentage
Arrest Rate Q2	Number of 'True' Arrest during Q2 2023 / The population in Chicago 2023 Calculated from attribute: Arrested	Percentage
Arrest Ratio Q3	Arrest Rate Q3 / Crime Rate Q3(KPI)	Percentage
Total Population Of Chicago 2020/21/22/23	The Parameter of the population number get from the government datasets from 2020-2023	Number

3.0 Data Discoveries

3.1 Pareto Analysis

3.1.1 *Details on the 80% Reports from Top 20% Crime Types*

A text table was utilized to display comprehensive information on crime reports for *Year 2023 Week 46*. The reports exclusively feature the top 20% crime types from the Pareto Chart, namely:

1. Theft
2. Battery
3. Criminal Damage
4. Motor Vehicle Theft
5. Assault

Each row in the text table corresponds to a report case and includes details such as case number, crime type, location description, community name, arrest and report date.

We have also created a line and bar chart to visually see how the top 5 crime types (the top 20% of crime types) represented more than 80% of the crimes. This can provide us valuable information on how to better distribute police resources, and where to focus our efforts on tackling crimes to lower the crime rate.

3.1.2 *Crime Report and Arrest Rate*

A further analysis was conducted for the top 20% Pareto Chart crime types. Comparisons were made between the current week (*Year 2023 Week 46*) and the previous week (*Year 2023 Week 45*) for each crime types' report and arrest rates. The results showed that, although only the report rates for the crimes of 'Battery' and 'Motor Vehicle Theft' increased, 'Criminal Damage' was the only crime that exhibited a decrease in the arrest rate. A decrease in the report rate does not necessarily imply a lower arrest rate and vice versa. Therefore, future analyses are required to determine the correlation between these two ratios.

3.2 Clustering

Cluster analysis of the crimes reported in Chicago was conducted to segment various crime types and their respective locations based on a calculated arrest ratio. The arrest ratio was derived as the occurrences of arrests over the number of reported crimes. This segmentation focused on the *previous quarter of the year 2023 (Q3)*. Based on this methodology we were able to observe the formation of three distinct clusters for all the data points. It is important to note that these clusters were then extracted as sets and further analysis on them was performed.

Based on this analysis we were able to observe that the crimes reported within the first cluster, heavily feature crimes with no risk to life with the most occurring crime types in this cluster being "Deceptive Practice" and "Other Offenses". Additionally, the crimes often occurred in College or University residence halls. This analysis suggests that the offenders of these crimes are likely to be young college students.

Furthermore, within the second cluster and third, we can observe more serious crimes that incur threats and harm to life. The most frequently occurring crime types in these clusters are: "Battery", "Theft", "Assault" and "Criminal Damage". It is important to highlight where the two clusters differed were the locations in which these crimes were committed. The second cluster consisted of serious domestic crimes whereas crimes reported in the third cluster were committed on public transport platforms.

Finally, during this analysis, we observed that the less serious crimes reported were present within *Cluster 1* comparatively had a significantly higher arrest ratio ranging from *0.75 to 1* followed by *Cluster 3* with arrest ratios ranging from *0.75 to 0.3* and *Cluster 2* with arrest ratios ranging from *0.3 to 0*. This analysis suggests that crimes reported in residential areas have a significantly lower chance of resulting in arrests.

3.3 Correlations

3.3.1 With District as a Dimension

Examinations were conducted to ascertain the correlation between *domestic and non-domestic crimes* across districts, using a Scatter plot and trend line. Over the entire span from *2020 to 2023*, a discernible pattern emerged, that is, an uptick in domestic crimes corresponded with a concurrent increase in non-domestic crimes, by following a linear trend. Moreover, the calculated *p-value* of *0.0152844* (obtained from the Trend Line model description), falling below the conventional significance threshold of 0.05, underscores the statistical significance of this observed relationship. This finding substantiates the robustness and credibility of the identified correlation.

Similarly, assessments were undertaken to compare *true and false arrests per district*, revealing a pattern akin to that observed between domestic and non-domestic crimes. The calculated *p-value* of *0.0017583* (obtained from the Trend Line model description) further substantiates the strength of this relationship, indicating its statistical significance.

3.3.1 With District, Crime Type and Correlation as a Dimension

The dimension list was extended to include district, crime type, and correlation coefficient. Notably, Python's NumPy library was used with the support of TabPy to compute the correlation coefficient. In particular, the correlation coefficient was calculated with respect to crime type.

For both the *domestic and non-domestic crimes* across districts and *true and false arrests per district*, the resulting p-value of the trend line was < 0.0001. Here is the script that was used for *domestic and non-domestic crimes*:

```
SCRIPT_REAL("import numpy as np
return np.corrcoef(_arg1,_arg2)[0,1]",
[ _COUNTFalseArrests],[ _COUNTTrueArrests])
```

In addition, here is the script that was used for *true and false arrests per district*:

```
SCRIPT_REAL("import numpy as np
return np.corrcoef(_arg1,_arg2)[0,1]",
[ _COUNTDomestic], [ _COUNTNonDomestic])
```

3.4 Comparisons

A *horizontal bar chart* was developed based on the *crime type and the presented police districts* in *Q3 over the span of time 2020-2023*. This aimed to highlight the most happened types of crimes in the time frame to assist the police force in distributing the manpower and patrolling effort within each police district.

- The top six crime types that happened the most were Theft, Battery, Criminal Damage, Assault, Motor Vehicle Theft, and Deceptive practice among all police districts.
- The most common crime type (Theft) comprised 22.09% of the total crimes.

Comparative Analysis of Monthly Crime Rate:

- In *January 2023*, there was a notable increase in Theft, Battery, and Assault, indicating potential post-holiday stressors impacting crime rates.
- *February 2022* saw a significant peak in Battery-related crimes, suggesting specific triggers during this period.
- *January of both 2021 and 2020* also experienced spikes in Battery and Theft, hinting at recurring issues or seasonal patterns influencing these crime types.
- Assault, Motor Vehicle Theft And Criminal Damage (Minor and moderate fluctuations): These are varying trends, with some months experiencing increases and others decreasing. This could reflect changing social conditions, the effectiveness of community policing, or even weather-related patterns that can influence crime rates.

3.5 KPIs - Crime rate, Arrest rate, and Arrest ratio in Q3 2023 versus the Previous Quarter

The *KPI analysis* focused on three crucial metrics: *crime rate, arrest rate*, and *arrest ratio*. The crime rate, which was calculated as the number of reported crimes per the population in Chicago for 2023, showed an increase from the previous quarter. Similarly, the arrest rate, defined as the ratio of true arrests to the population in Chicago for 2023, also exhibited a notable rise compared to the preceding quarter. However, despite the aforementioned increments, the arrest ratio, which measures the ratio of true arrest crimes to reported crimes, surprisingly showcased a decrease.

The crime and arrest rates' upward trend indicate an escalation in criminal activities and law enforcement activities, respectively. However, the declining arrest ratio signifies a disproportionate increase in reported crimes to actual arrests, suggesting potential challenges in law enforcement efficacy or shifts in reporting patterns, warranting further investigation and strategic adjustments in law enforcement methodologies or resource allocation.

4.0 Root Cause Analysis

4.1 Summary of Problems

Seasonal Factors as Root Causes:

- The consistent increase in certain crimes during January could be due to post-holiday economic stress, increased alcohol consumption during the holidays, or fewer daylight hours which can affect crime rates.
- The decrease in crime ratio e.g. comparing Q2 to Q3, represents an imbalance of more crimes occurring versus actual arrests, and signals a decline in law enforcement efficacy.

Battery Incidents:

- The recurrent peaks in Battery could be rooted in social and familial tensions that come to a head during or after the holiday season, or they could be influenced by other factors such as changes in local law enforcement practices or reporting. Furthermore, these peaks can be associated with the surge in public transport usage during holiday seasons.

Annual Cycles:

- If the spikes in crimes are tied to annual events or societal rhythms, the root causes might include cultural practices, the effectiveness of seasonal law enforcement strategies, or environmental conditions that are more conducive to crime during these periods.

4.2 Crime Prevention Challenges (Why Questioning)

- **Why was there a sharp increase in Theft, Battery, and Assault in January 2023?**
    - Possible answers could include lack of adequate policing during the New Year festivities, and increased opportunity for crime during a busy holiday season.

- **Why do certain types of crime increase at specific times of the year?**
    - Economic cycles, social events, or seasonal factors that affect people's behaviour.

- **Why might the same types of crime occur repeatedly over the years?**
    - The root causes are not being closely examined. Additionally, preventative measures may either be ineffective or nonexistent.

- **Why are some crimes more difficult to prevent or solve than others?**

- Crimes reported in residential areas tend to have slower response times as opposed to crimes reported in more populated and active areas such as public transport stations and platforms.

4.3 Recommendations

1. **Preventive Measures:**
    - Develop targeted crime prevention campaigns during the holiday season to reduce motor vehicle theft.
    - Increase patrols and implement community watch programs in high-risk areas particularly residential areas in: “Austin”, “Near North Side”, “Near West Side, “South Shore” and “North Lawndale” and public transport stations during times identified as peak crime periods.
    - Improve street lighting within communities to deter crimes from happening.

2. **Community Engagement:**
    - Encourage community reporting and engagement through awareness programs, ensuring that citizens are aware of crime trends and prevention strategies.

3. **Technology and Infrastructure:**
    - Invest in surveillance and monitoring technology to deter potential criminals.
    - Enhance data analytics capabilities for real-time crime tracking and response.

4. **Policy and Training:**
    - Update law enforcement training to focus on data-driven approaches to crime fighting.
    - Adjust policies to respond flexibly to the dynamic nature of crime trends, allowing for rapid shifts in focus and resources.