IS525 Final Project Abstract

Junjie Lai Youdong Lu  Richie Li  Emily Chang

**Dataset Overview**

The Chicago Crime dataset from the City of Chicago Data Portal provides detailed information on crimes reported in Chicago, Illinois, from 2001 to the present. With over two decades of records, this dataset is a valuable resource for analyzing crime patterns, understanding community safety, and examining the effectiveness of law enforcement in various neighborhoods.

Using Tableau to analyze and visualize this data allows for dynamic, interactive exploration of complex patterns and relationships within the dataset. Tableau’s geographic mapping and time-series features are particularly advantageous for this dataset, as they enable viewers to track crime distributions across different neighborhoods, observe temporal trends, and identify areas where law enforcement interventions may be effective. Tableau’s ability to create interactive dashboards also empowers users to engage with the data meaningfully, filtering by crime type, time period, or location to uncover insights tailored to specific safety concerns or community interests.

**Variables**

* **Incident Identification:** ID, Case Number
* **Time Attributes:** Date, Year, Updated On
* **Location Attributes:** Block, Location Description, Community Area, Ward, Beat, District, Location (Latitude and Longitude)
* **Crime Classification:** Primary Type, Description, IUCR, FBI Code
* **Crime Details:** Arrest (True/False), Domestic (True/False)
* **Geospatial Coordinates:** Latitude, Longitude, X Coordinate, Y Coordinate

**Data Cleaning Process in Tableau Prep**

To prepare the data for an effective Tableau dashboard, the following cleaning steps in Tableau Prep are necessary:

**1. Handling Missing, Anomalous, and Outlier Values**:

* **Missing Values**: Address any missing or null values, particularly in key columns such as Primary Type, Arrest, and Location Description, as these are essential for accurate analysis.
* **Anomalous Values**: Review the Date field for any errors, such as future dates or improbable timestamps. This step ensures data accuracy.
* **Outliers**: Identify and handle outliers, especially in numerical or geographic data (e.g., Latitude and Longitude) that may indicate incorrect entries or misreported locations.

**2. Filtering for Recent 10 Years of Data**: Limiting the dataset to the most recent 10 years (e.g., 2013–present) will focus the analysis on recent crime patterns, making the dashboard more relevant and efficient. This also reduces data volume, allowing for smoother performance and more relevant visual insights.

**3. Removing Irrelevant Variables**: To streamline the dataset, remove variables that don’t directly contribute to the analysis or visualization, such as Case Number, IUCR, Updated One. This simplifies the data and keeps the dashboard focused on insights most useful for public safety analysis.

**Data Story Concept**

With this dataset, the Tableau dashboard will tell a compelling story about crime in Chicago, focusing on the following areas

**Detailed Yearly Crime Statistics**: To gain insights into the types of crimes occurring in Chicago, we will analyze and visualize yearly statistics for each crime category. By breaking down crimes by type and year, we can observe trends in specific crime categories, understand shifts in crime patterns over time, and identify any significant increases or decreases in particular types of crime (Youdong Lu) .

**Total Yearly Crime Counts**: Additionally, we will calculate and display the total number of crimes reported in Chicago each year. This metric will allow us to assess overall crime trends and track whether crime is increasing, decreasing, or remaining stable over time (Youdong Lu) .

**Crime Type and Location Analysis**: By examining the relationship between crime types and their typical locations, we aim to reveal patterns in where specific crimes occur most frequently. This analysis helps identify high-risk locations associated with particular crime types (Junjie Lai).

**Crime Analysis by Administrative Areas**: We will break down crime statistics by Chicago’s administrative areas, analyzing crime counts, felony rates, and clearance rates. This approach will highlight variations across districts, providing insights into the distribution and severity of crime in different regions (Junjie Lai).

**Crime Analysis by Time of Day and Weekday Patterns:** This analyzes the distribution of crimes based on the time of day and day of the week. It will reveal peak hours or days when certain crimes are more likely to occur. For example, comparing crime trends on weekdays versus weekends or analyzing specific crime categories during different times (e.g., robberies at night vs. during the day). (Richie Li)

**Arrest Crime Analysis:** We want to know the correlation between crimes resulting in arrests, and potentially revealing patterns in law enforcement responses to different types of incidents. A breakdown by crime category and the proportion of those leading to arrests could further enhance this analysis. (Richie Li)

**Crime trends among seasons:** Are there any seasonal trends in crime rates? For example, are certain crimes more common in summer than in winter? We could possibly analyze monthly or seasonal variations to highlight times of year when specific crime types are more prevalent. (Emily Chang)

**Comparison on Domestic and Public Crimes by Location:** We want to look at how domestic crimes (crimes that happen at home) are different from public crimes (crimes that happen in public places). We'll do this by looking at where these crimes happen most often. This will help us see if certain neighborhoods or areas have more of one type of crime than the other. By understanding this, we can figure out where we might need to add more community support or resources to help prevent these crimes. (Emily Chang)

**IUCR Codes and FBI Classifications of Frequent Crimes:** We will investigate the most common IUCR codes in the dataset to determine which types of crime are reported most frequently. By aligning these codes with FBI classifications, we can categorize crime types more effectively and assess how local crime patterns correlate with national standards. This alignment can also reveal any specific crime categories unique to Chicago or more prevalent than national averages. (Emily Chang)