Exploratory Analysis and Spatial Representation of Chicago Crime Dataset Using Tableau

Rhythm
Department of Data Science and Analytics
Central University of Rajasthan, Ajmer, Rajasthan, India
rhythmtcurious@gmail.com

***Abstract*—The aim of this project is to gain important insights from Chicago Crime Dataset by Data Analysis and visualize this data on a map using Tableau.**

I. Introduction

Chicago Crime Dataset is a very famous dataset that provides a variety of results. The dataset is dynamic and diverse, which is a rich source for investigating crime sources within the city. Using this project, the primary goal is to see the visualization of different crimes on the map. The main tool for visualization is Tableau.

II. Literature Survey

A. Spatial Analysis of Crime Data

Numerous studies underscore the importance of spatial analysis in understanding crime patterns. Braga et al. (2008) demonstrated the efficacy of spatial analysis in crime mapping, emphasizing its utility in identifying crime hotspots and aiding law enforcement in resource allocation.

B. Tableau in Crime Data Visualization

In a study by Wang et al. (2015), Tableau was employed to visualize crime data in San Francisco, showcasing its intuitive interface and capability to communicate insights effectively.

C. Integrating Time and Space in Crime Analysis

The temporal dimension adds another layer of complexity to crime data analysis. Time-series analysis has been explored by Andresen (2016) in the context of crime patterns, highlighting the importance of considering temporal factors in addition to spatial components.

III. Data and Processing

A. Distribution Analysis

To gain an initial understanding of the dataset, a distribution analysis of crime incidents across different addresses was conducted. A pie chart was generated to visualize the proportion of crimes occurring at distinct addresses. The chart, depicted in Figure 1, provides insights into the concentration of criminal activities at specific locations.

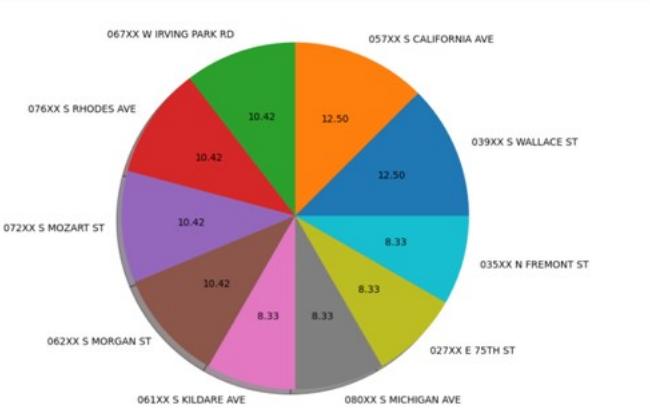

Fig. 1. Address Distribution

B. Temporal Analysis

The 'date' column was first converted to DateTime64 format to facilitate temporal analysis. Monthly crime trends were then examined by resampling the data on a monthly basis and visualizing the results using a line plot. The plot, as illustrated in Figure 2, highlights fluctuations in crime incidents over time.

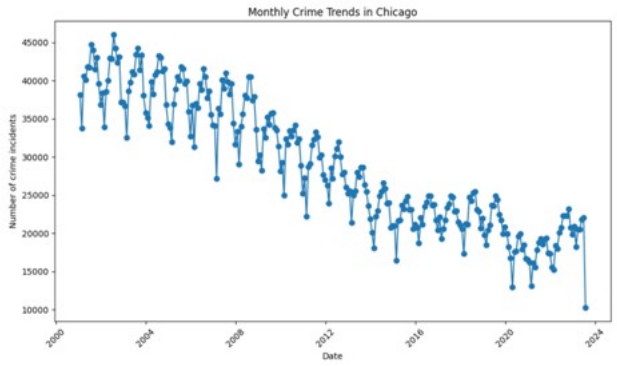

Fig. 2. Temporal Analysis with time.

C. Primary Type Analysis

Further exploration focused on identifying the top 10 primary types of crimes in Chicago. A pie chart was generated

to visually represent the distribution of these primary crime types.

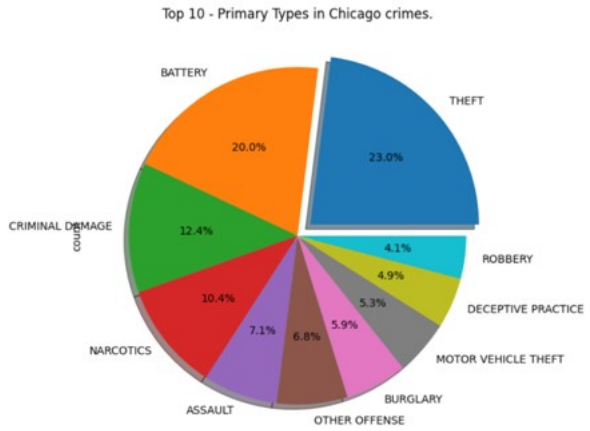

Fig. 3. Top 10 primary types of crimes in Chicago.

IV. Tableau Implementation

This section outline the step-by-step Tableau implementation used to gain various insights.

A. Data Import

The first step involves importing the Chicago Crime Dataset into Tableau. By opening Tableau and connecting to the dataset, we ensure a smooth transition from data preprocessing to visual exploration.

B. Initial Mapping

With the dataset loaded into Tableau, the next crucial step is to map the geographical coordinates appropriately. By dragging the 'Latitude' and 'Longitude' columns to the Rows and Columns shelves, respectively, a fundamental crime map is generated. This initial mapping sets the stage for a spatial representation of crime incidents across the city.

Fig. 4. Initial mapping.

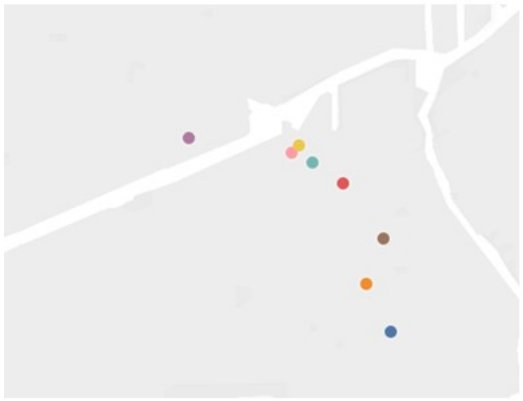

Fig. 5. Adjusting size and zoom.

C. Crime Representation

To enhance the visual appeal and interpretability, the 'Primary Type' field is incorporated. Utilizing color-coded representations, each crime type is distinctly highlighted on the map. Adjusting the size of data points further refines the visualization, offering a clearer depiction of crime prevalence in different areas.

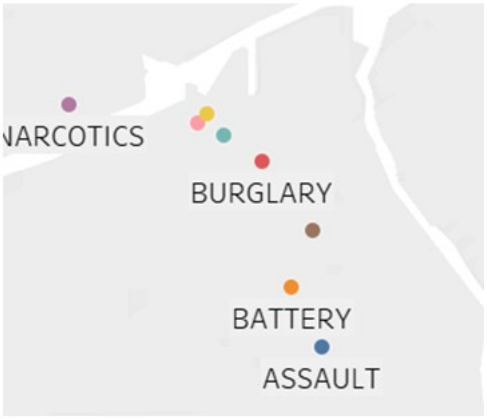

Fig. 6. Adjusing size.

D. Temporal Analysis

The temporal dimension of crime data is explored. The 'Date' column, previously converted to the 'Date' data type during pre-processing, is utilized to create a time series. By dragging the 'Date' field to the 'Columns' shelf, a dynamic representation of crime trends over time is generated. This temporal analysis provides valuable insights into increase and decrease in the flow of criminal activities throughout the dataset period.

V. Results and Discussion

The visualizations produced offer valuable insights into spatial distribution, temporal variations, and the prevalence of specific crime types. The following key findings were observed:

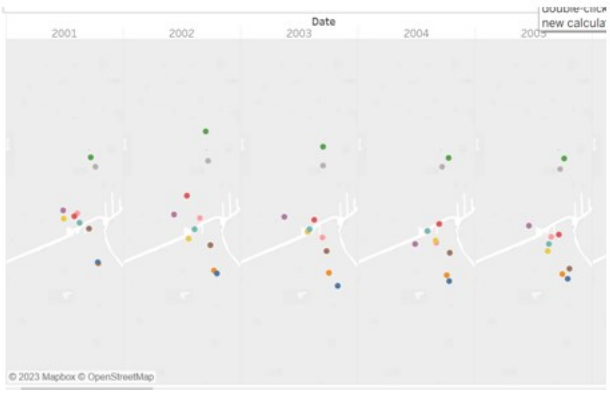

Fig. 7. Temporal Analysis with Time.

A. Spatial Distribution of Crime

The initial crime map revealed distinctive hotspots and clusters of criminal activities across Chicago. High-density areas were identified, providing a spatial understanding of crime concentration. Through color-coded representations based on the 'Primary Type' field, specific crime types were distinguished on the map, offering a nuanced view of the diverse criminal landscape within different neighborhoods.

B. Temporal Trends

The temporal analysis, presented in the form of a time series plot, exposed fluctuations in crime incidents over the dataset's timeline. This insight aids in recognizing patterns related to specific months or years, contributing to a more informed assessment of temporal trends.

C. Top 10 Crime Types

The pie chart representation of the top 10 crime types illuminated the prevalence of specific offenses within the dataset. By highlighting the distribution of these primary crime types, the visualization enables a quick and clear understanding of the most frequently occurring offenses in Chicago.

VI. Acknowledgements

I would like to express my sincere gratitude to Dr. Bhavna Saini for her guidance and support throughout this project. Her expertise and insights have been instrumental in shaping the research and ensuring its quality.

VII. References

References

[1] Andresen, M. A. (2016). Crime Measures and the Spatial Analysis of Criminal Activity. *Springer.*

[2] Braga, A. A., Hureau, D. M., & Papachristos, A. V. (2008). The Concentration and Stability of Gun Violence at Micro Places in Boston, 1980–2008. *Journal of Quantitative Criminology, 26*(1), 33–53.

[3] Chainey, S., Tompson, L., & Uhlig, S. (2008). The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime. *Security Journal, 21*(1-2), 4–28.

[4] Waller, L. A., Sams, E. A., Lippmann, S. J., Kaufman, J. D., & Burke, J. (2017). Evaluation of Land Use Regression Models for NO2 in Montreal, Canada. *Journal of Exposure Science and Environmental Epidemiology, 27*(2), 212–218.

[5] Wang, Y., Phillips, A., & Song, Y. (2015). Crime Data Visualization using Tableau Public: A Study of San Francisco Crime Data. *Proceedings of the Annual Conference of CAIS / Actes du congrès annuel de l'ACSI.*