Carjacking Analysis in Chicago: Visualizing Trends for Policy Implications

PPHA 30538 Final Project Group 28 - Surya Hardiansyah and Astari Raihanah

February 11, 2025

Introduction

Background

- Carjacking is a pressing urban issue, with significant impacts on public safety, insurance policies, and law enforcement resource allocation.
- According to **WTTW News** in 2023, Chicago is reporting more incidents than any other city in the United States.
- Carjacking incidents in Chicago vary widely across neighborhoods and time, requiring data-driven interventions.

Research Questions

- How do temporal and spatial patterns of carjacking incidents in Chicago inform actionable policy and solutions?

Methods in Data Analysis (1)

Data Sources

- Chicago Data Portal: Carjacking Incidents (2001-2024) and Chicago Community Area Boundaries (GeoJSON format)

Key Data Attributes:

- Incident date, time, and coordinates
- Neighborhood area boundaries

Data Preparation:

1. API-based data retrieval (pagination used to overcome 1,000-row limit)
2. Spatial Joins and aggregations by year, month, and time of day

Methods in Data Analysis (2)

Visualization and Shiny Dashboard

- **Static choropleth maps** and **line charts** to explore trends.
- Shiny app for dynamic filtering by neighborhood and date range.

Challenges:

1. **API Limitations**: The Chicago Data Portal limits API downloads to 1,000 rows per request, hence we used pagination to retrieve all 22,192 records.
2. **Data gaps**: Missing or incomplete records, such as missing coordinates, were excluded from the analysis (144 data rows per November 30th, 2024).

Natural Language Processing Analysis and Limitations

SerpAPI Data Collection

- Applied Natural Language Processing (NLP) to analyze Polarity (Positive/Negative Sentiment) and Subjectivity (Degree of opinion/bias) with 1000 observations (2021-2024)
- Data sources: Google search results, news websites, and public repositories.
- Queries used: “Chicago car insurance policy”, “auto insurance Chicago”, “carjacking auto insurance Chicago”, “auto insurance Chicago car theft”, “Chicago carjacking”, and “Chicago car theft”

Limitations and Biases:

- Most of the insurance keyword extracts from insurance company website.
- Search results may prioritize recent high-profile cases, overlooking older or less-publicized incidents.
- Overrepresentation of specific neighborhoods based on media priorities.
- The news sources focus more on recent events, historical trends might be incomplete.

Spatial Patterns

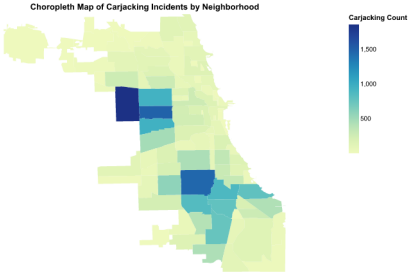

Figure 1: Choropleth Map of Carjackings

Findings

- Hotspot Neighborhoods: **Austin, West and East Garfield Park, Englewood**.
- These neighborhoods economic challenges, high-crime level, and limited infrastructure.
- Policy implication: Localized interventions (e.g., police patrols, infrastructure upgrades)

Temporal Trends

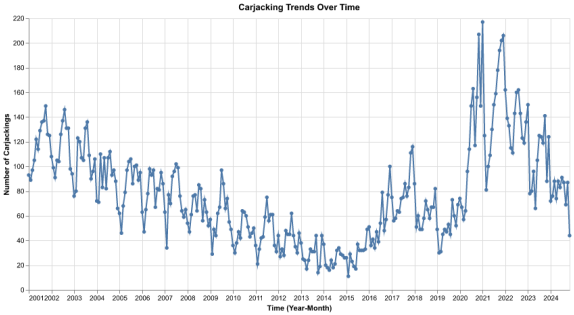

Figure 2: Carjacking Trends Over Time

What might explain the 2020-2021 spike?

- **Pandemic disruptions** - reduced police presence in the neighborhood.
- **Economic uncertainty** - promote more people into crime.
- Police resource reallocation during social unrest.

Shiny Interactive Dashboard

- Dynamic choropleth map for spatial analysis with date and neighborhood filters
- Dynamic line charts for time series analysis with date and neighborhood filters
- Empowers policymakers and stakeholders to explore patterns interactively

NLP Analysis: Polarity and Subjectivity

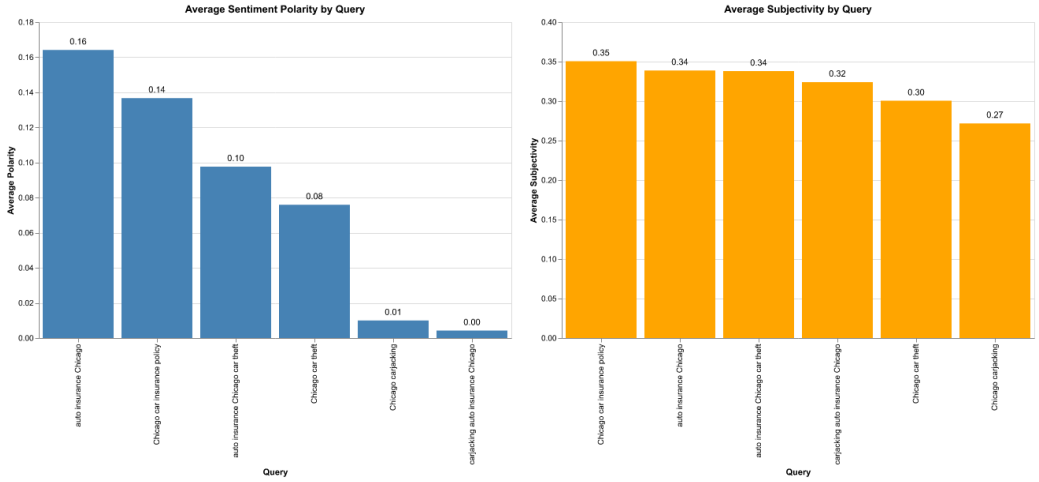

Figure 3: Average Polarity and Subjectivity by Query

NLP Key Findings

- Insurance-related discussions (e.g. “auto insurance Chicago”) are more positive and subjective, likely due coverage of financial protections and consumer advice.
- Crime-related discussions are neutral to slightly negative (e.g. “Chicago carjacking”, and “Chicago car theft”) - lower polarity and less subjective framing (more factual and data-driven rather than opinions).

Policy Implications: Data-Driven adjustments

- Hotspot-based Policing: More patrols in Austin, West/East Garfield Park, and Englewood.
- Infrastructure improvement: lighting, surveillance cameras, community safety programs.
- Time-Based Resource Allocation: Deploy more law enforcement at night and during high-crime periods (e.g. 2020-2021 spikes).
- Proactive crime mitigation during external shocks (e.g. economic crises).
- Insurance premiums should be risk-adjusted based on carjacking incidence rates (e.g. higher premiums in high-risk areas like Austin and lower premiums in Hyde Park).

Conclusions

Key Takeaways:

- Combining spatial, temporal, and sentiment analysis provides a holistic view of carjacking trends and public perceptions.

Future Directions:

1. Integrate additional datasets (e.g. traffic patterns, socioeconomic factors)
2. Enhancing the dashboard's interactivity and performance by optimizing data structures will improve user experience and scalability.