Mapping Crime: Graph Data Science for Crime Investigation

Sidra Aamir, Areeb Kazmi

Problem Statement

Conventional crime investigation struggles with massive and intricate datasets, leading to time-consuming analyses and incomplete outcomes. Reliance on Relational Database Management Systems (RDBMS) further complicates matters by potentially missing crucial connections between entities involved in criminal activities. Hence, there is an urgent need for innovative approaches, employing Graph data modeling and Machine Learning to streamline investigations and unveil hidden connections essential for effective crime resolution.

Objectives

The project aims to use Graph Data Science to improve crime investigation. By analyzing the connections within the graph, we aim to achieve the following objectives:

- Identifying potential criminal linkages:
- Analyzing suspect movement patterns:
- Identifying potential witnesses:
- Linking stolen objects to crimes and people:
- Identifying crime hotspots:

Results

Our analysis reveals critical insights into crime patterns and trends, enabling law enforcement agencies to make informed decisions and implement targeted interventions. By identifying districts with the highest reported cases, the most common crimes, and the time of highest crime occurrence, we can optimize resource allocation and enhance proactive policing strategies. Additionally, understanding the demographic groups most affected by crimes, prevalent weapons used, and specific locations and premises where crimes occur provides valuable intelligence for tailored enforcement and community engagement initiatives. Overall, our analysis empowers law enforcement to address crime effectively and improve public safety outcomes.

Methodology

1. **Data Acquisition and Preprocessing:**
   - The post-crime deals with victims, the case they reported, place, and other details.
2. **Graph Schema Definition:**
   - Modelled the graph as separate node labels for each category and assign meaningful labels.
   - Loaded using the Neo4j's Cypher query language and potential data quality checks will be in place.
3. **Network Analysis and Exploration**:
   - Explored connections within the graph based on the objectives.
4. **Graph Enhanced ML:**
   - Applied Machine Learning for Node Classification and Link Prediction.

Conclusion

Our project showcases the transformative impact of Graph Data Model and Science on crime investigation. Using graph databases and advanced analytics, we've established a robust framework for analyzing crime data comprehensively. Our methodology unveils hidden connections and patterns, augmenting the efficiency of investigations. Moving forward, our work promises to advance crime investigation practices, benefiting law enforcement agencies and communities.

Future Scope

- Implementing a good frontend and larger graph
- Applying link prediction and node classification
- Generating additional data to have more labels
- Better use of GDS queries and use of centrality and PageRank queries

Limitations

- Data Limitation, post-crime data with no details about suspects, criminals, investigators.
- Not enough large data incorporated into Graph DB
- Graph Data Science queries could have been better
- Issue with Node classification and PC resources
- Could not implement/train Machine Learning model