DME Integrative Assignment

Angelo Barisano; 508903

September 16th, 2022

1 Task 1: Plan & Explore: DEfine Questions for project and familiarize with data

1.1 Overall nature of the Data

Data drives security. The wide spread adoption of data information systems to recognize crime-hotspots patterns to increase policing efficiency and protect people. Thus, creating effective information systems for crime prevention is at the center of policy maker and executive branches of governments. This trend has led to the Chicago PD reaching out to this Data Management Team with the request of creating a Data Management System for crime-data in Chicago.

To this end, the Chicago PD and this Data Management Unit is interested on setting the foundation for a FAIR database; i.e. a Data Base ment to flexible and easy in use of slicing and dicing data to find answers and, contemporaneously, make them available.

1.2 research question

The overarching goal of this project is to set the foundations to funnel the existing crime-data into a data model to improve its readability (analytics), accessibility (logical structure & metric-interpretability), and further use in feeding systems for advanced predictive crime prevention.

In an iterative (agile) developmentcylce the following questions have been set up: Chicargo is known for its high homocide rate; this leads to policy makers focusing on this issue the most.

Upon conducting an initial Exploratory Data Analysis (EDA), four major categories could be identified in the existing data:

- Location

- Crime Type

- Arrest(Yes/No)

- Time

Based on the data types, the EDA, and the aforementioned goal of creating a flexible DB, the following research questions guide the creation of the database increaing in complexity:

Hypothesis 1 (H??) *How do certain crime types (homicide) patterns distibute by seasonality patterns, by location, by time of day?*

Hypothesis 2 (H??) *Do crime patterns persist? Do homicide patterns persist in eg a certain district; do certain crimtypes persist wrt. location over time ?*

Hypothesis 3 (H??) *Effective allocation of policing resources to prevent homicides: In order to prevent homocides; which "beats" are the most prevalent among homcides? During which time of day (for effective allocation of policing resources)? What type of crimes necessitate the repeated Loging of information? → these cases might need specialists*