Data Visualization

Final projects

Group

- Andrea Morando - 4604844
- Davide Miggiano – 4840761

Repo and links

- Repo: https://github.com/Emulars/la-crime
- Web page: https://emulars.github.io/la-crime/

Abstract

Our project proposal concerns the enhancement of a past project on crime-related data cleaning and analysis regarding the city of Los Angeles. For the project, the dataset used was for the period 2010 to 2024, but as the amount of data is around 4GB, it may be possible that during the development we will reduce the years under analysis to the range 2020-2024. The dataset in question is available at the following link; this data is derived from transcripts of reports compiled directly by the police, and is therefore subject to errors of various kinds.

Dataset Overview

The dataset used in this project focuses on crime incidents in Los Angeles.

It contains detailed information about various crimes, including their time of occurrence, location, and type, as well as demographic data related to the victims. Each record in the dataset represents a single crime event and includes the following key attributes:

- DateTime: The exact date and time when the crime occurred, formatted as ISO 8601.

- District: The administrative district in Los Angeles where the crime took place.

- Crime subtype and type: A description of the specific nature of the crime, such as "VEHICLE - STOLEN" or broader categories like "VEHICLE STOLEN."

- Age, Gender, Ethnicity: Demographic details of the victim involved in the crime.

- Weapon and Weapon Type: Information about the weapon used, if any, in the crime.

- Status: Indicates the case's current status, such as "Active" or "Archived."

- Street and Coordinates: The street address and geographic coordinates where the incident occurred.

The dataset enables an in-depth exploration of crime patterns over time and across different regions in Los Angeles. It also provides insights into how crimes correlate with victim demographics and other factors, such as the use of weapons.

Storytelling and Chart types

1. **Problem overview** – Bar chart and line chart about total crimes with difference with the average;
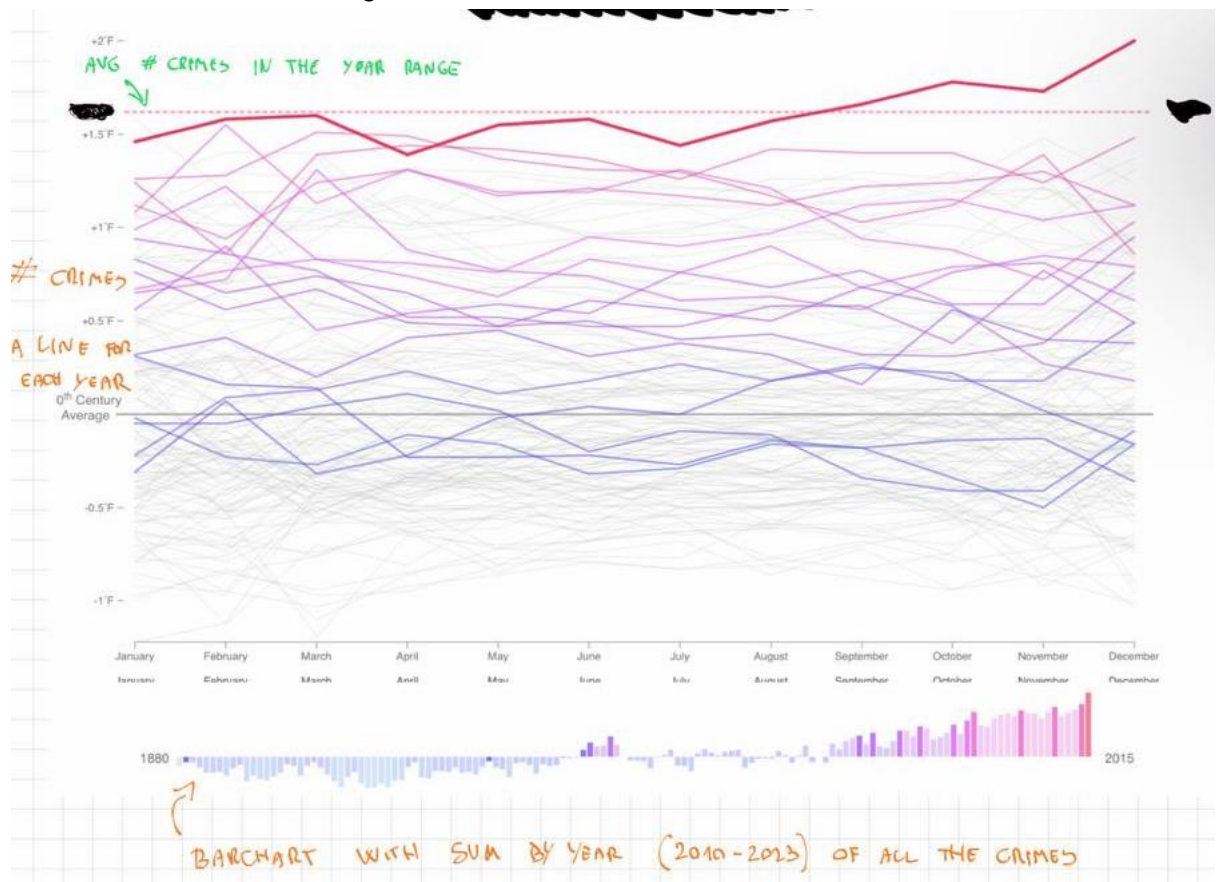
2. **Important events in LA –** highlight important event like covid to improve overview understanding;
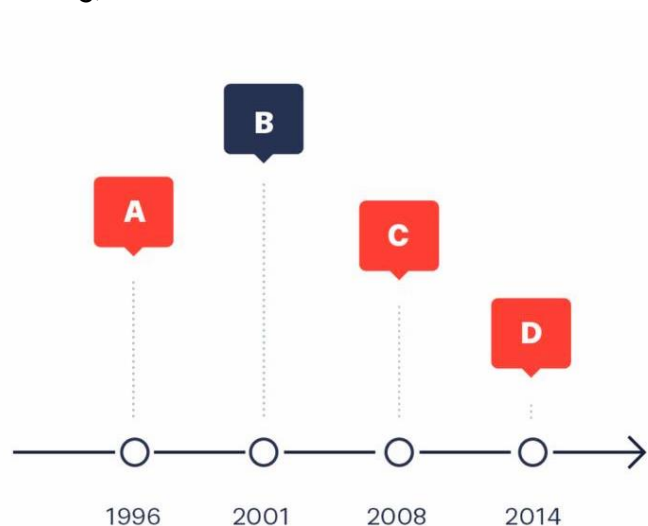
3. **Map, identification of critical areas** – the map shows the LA district with a color gradient based on a criminality index computed on parameter like weapon's types or other characteristics. There will be a tooltip with additional information like: # of crimes or most used weapon (if any). The map will be clickable to select a specific district for the next plots. The map will be filtered by year with a slider for better visualization.
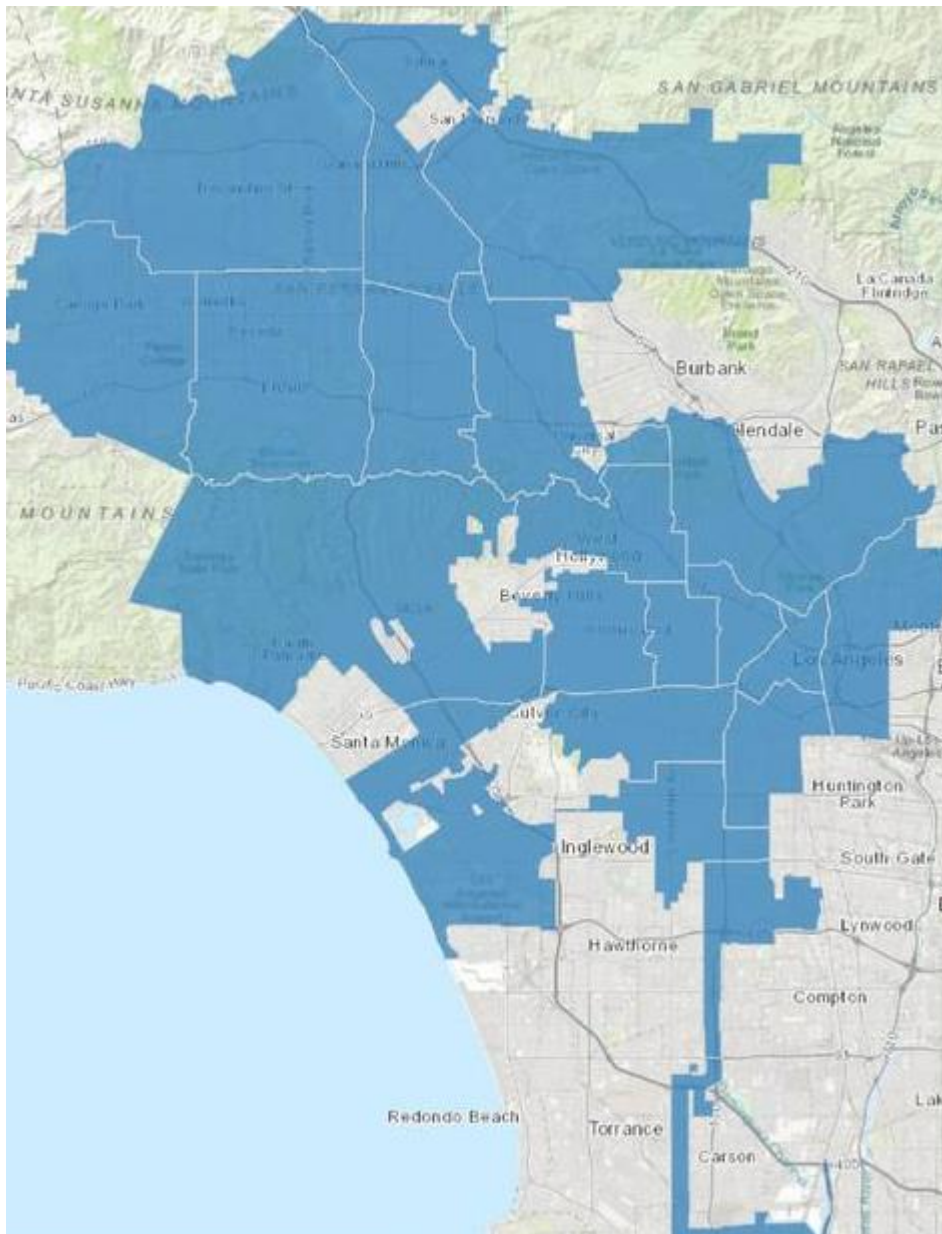
4. **District specific crime visualization during day/night time** – shows the average number of crimes at each hours of the day with the dot color gradient that depends on a crime computed index. In addition the dot will be surrounded by a circle if +50% of the crime used to compute the average is done with the use of weapon.
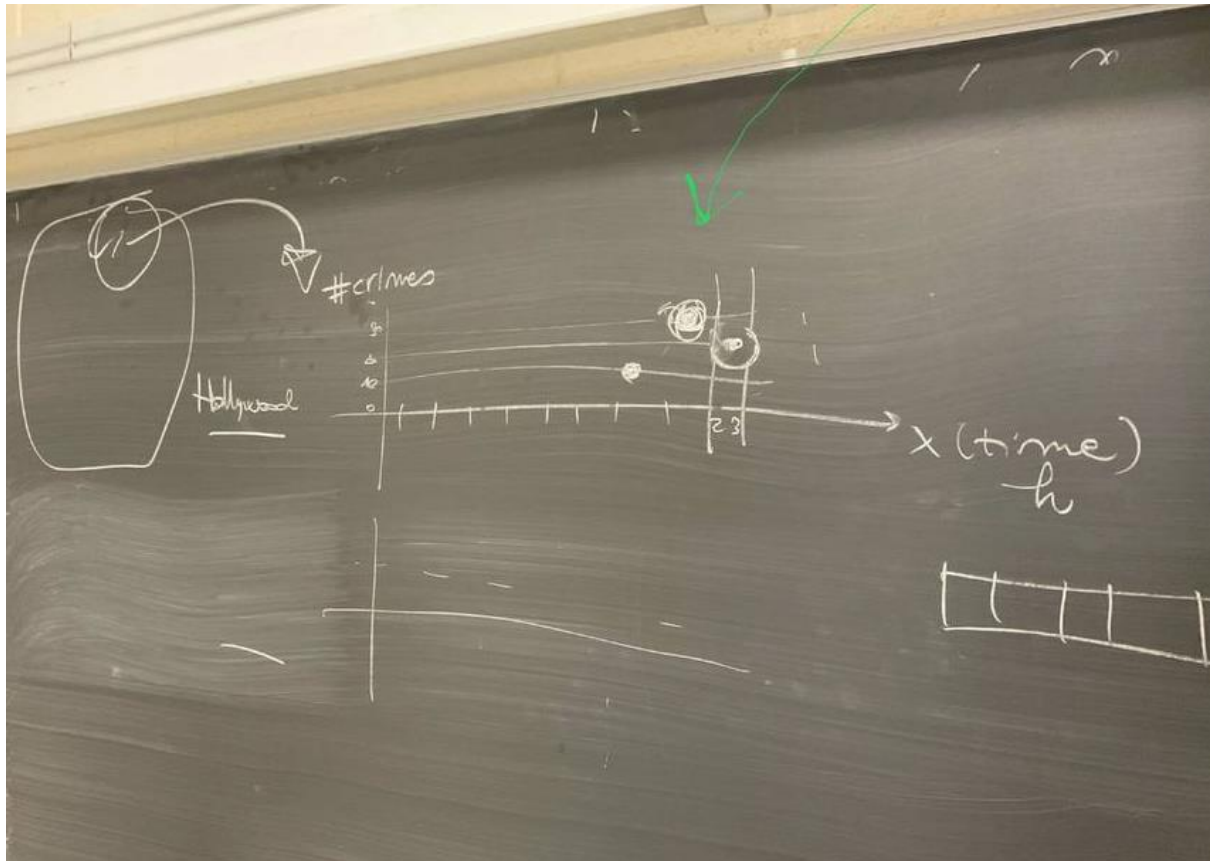
5. **Victims demographic groups** – alluvial plot to show how the different classes of victims (gender, ethnicity and age) are spread across the crimes.