Boston Crime Data Analysis

Data Science Lab on Smart Cities

Induni Sandapiumi Nawarathna Pitiyage – 906451
Sara Campolattano - 906453

CONTENTS OUTLINE

INTRODUCTION & MOTIVATION

This presentation aims to present a crime analysis of Boston, Massachusetts, from 2015 to 2018. The objective is to assess criminal patterns in urban areas to provide actionable insights to government for targeted law enforcement interventions, enhancing public safety in the city.
This is done by focusing on:

- *Crime Type*
- *Temporal Dimensions*
- *Geographic Dimensions*

INDICATORS

To analyze criminal activities, three main categories of indicators were used to identify trends and patterns:

- **Crime Type Indicator:** focusing on specific crimes such as burglary, larceny, vandalism, arson, and theft.
- **Temporal Indicators:** such as Day of the Week and Time of the Day.
- **Geographic Indicators:**
  - **Crime Hotspots**: which identify areas with high concentrations of criminal activity, such as neighborhoods or specific locations like crosswalks;
  - **Proximity to Key Locations**: which analyzes crime near schools, parks, and transit hubs to determine the impact of these facilities.
  - **Street-Level Crime Analysis**: which pinpoints specific streets with high crime rates for targeted law enforcement efforts.

DATA PREPROCESSING

The dataset is available on the Kaggle platform, and it was provided by the Boston Police Department (BPD). Originally, it comprised **327,820 records** about crimes reported from June 14, 2015, to September 3, 2018. The analysis in this presentation focuses on crimes against property.

The data pre-processing step allowed to detect missing records. Specifically, a high percentage of missing values was detected for:

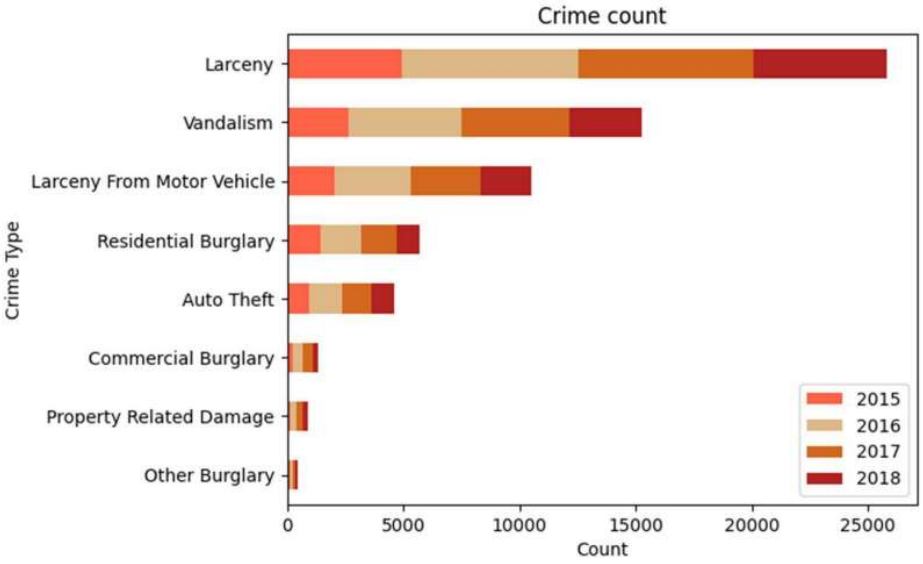
- **Shooting** attribute (99.9%) – this feature was removed.
- **District, Latitude, Longitude** – records with missing values or misinformation were discarded.

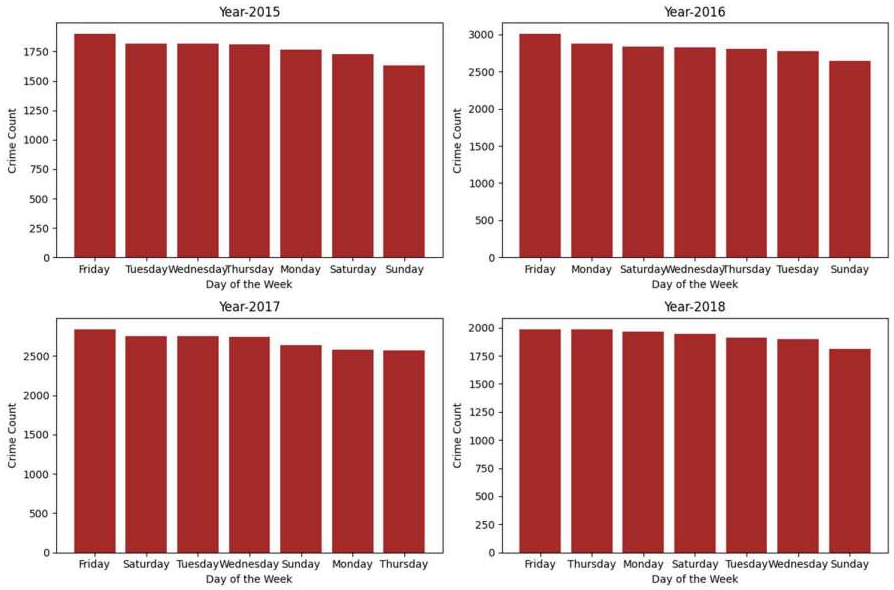
To provide useful insights, feature engineering was performed to create "**Date**" and "**Time**" features to analyze temporal trends. Moreover, the features **latitude** and **longitude** were employed to calculate proximity to amenities like schools, parks, and public transport for geographic insights.

The final dataset comprises **64,643 crime records** and **24 attributes**.

DATA ANALYSIS

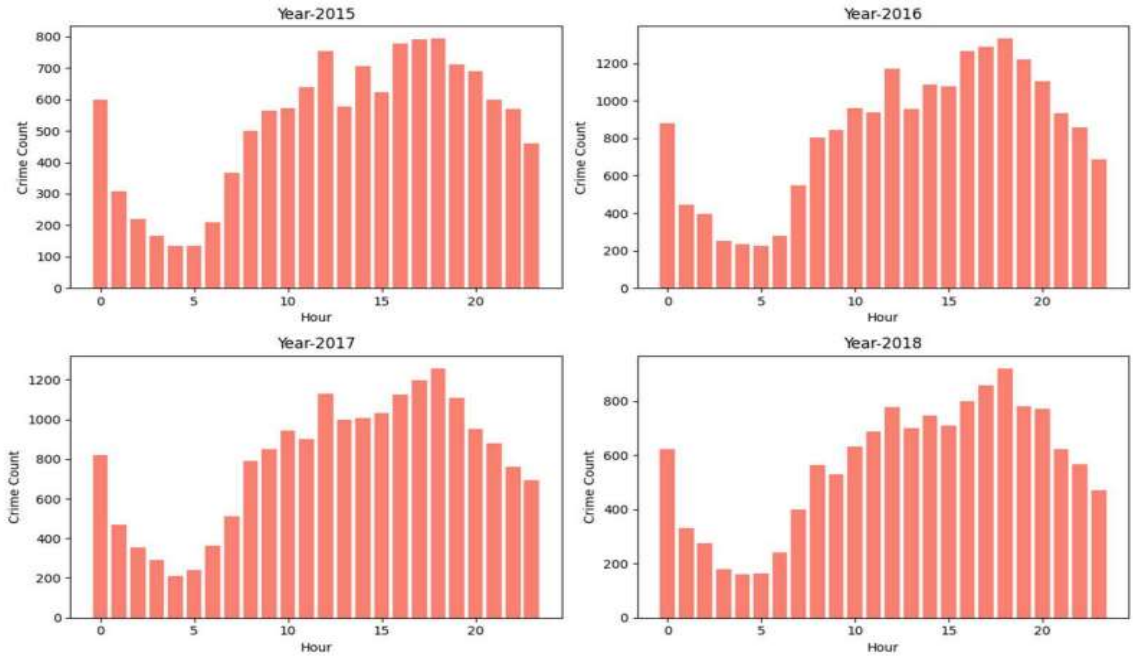
To better understand the specific crimes reported against property, in Boston, from 2015 to 2018, we can have a look at their distribution over said years and over the days of the week.





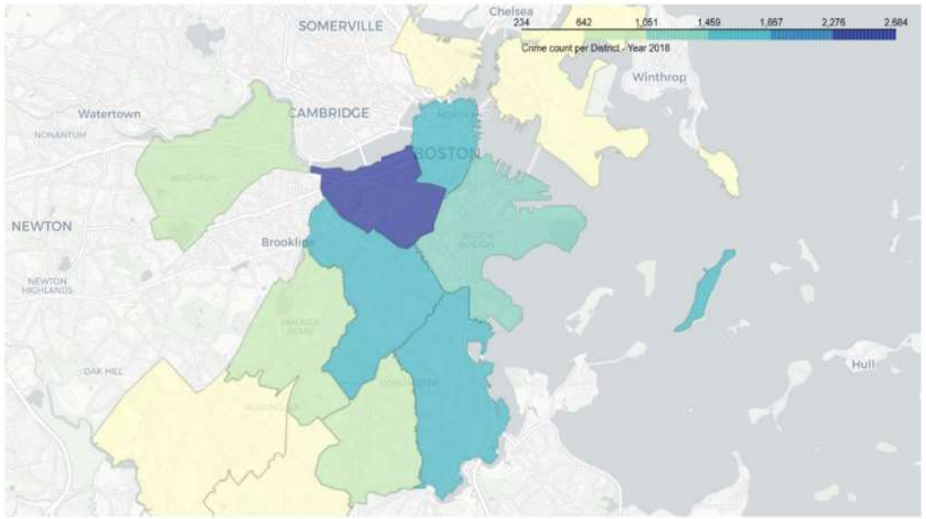
DATA ANALYSIS

To gain insights into potential trends, the crimes distribution over the years during specific hours of the day were observed.

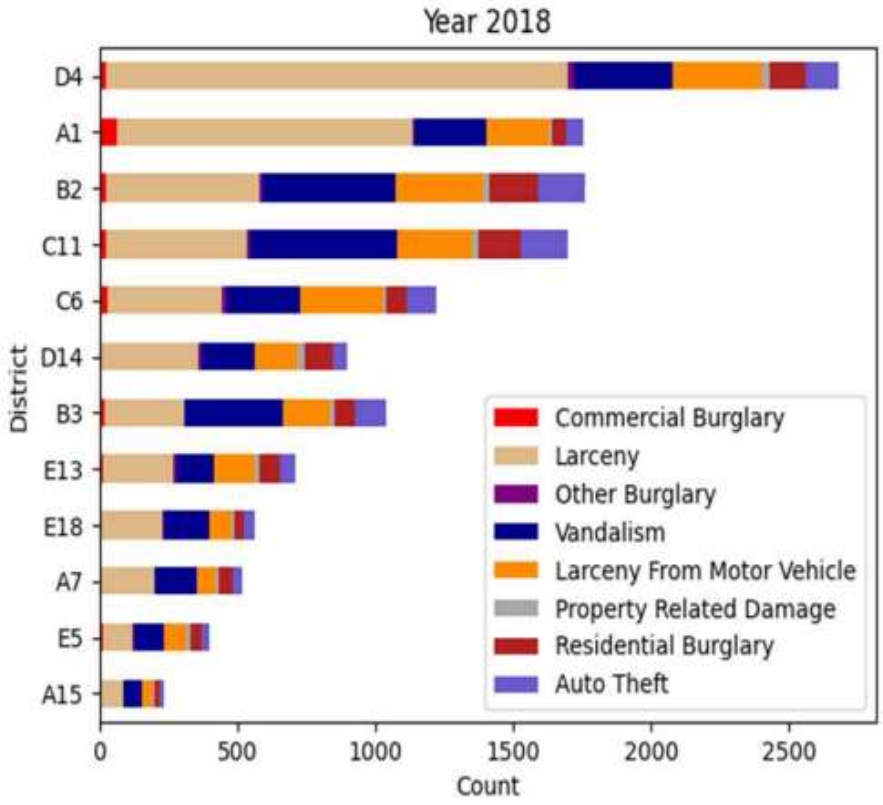


DATA ANALYSIS

Focusing on 2018, we can have a look at the crime counts across different districts and the specific crime types reported

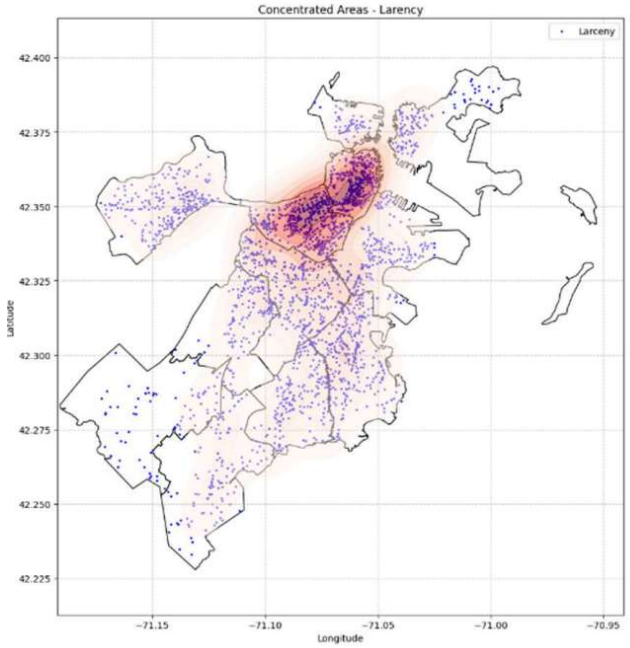


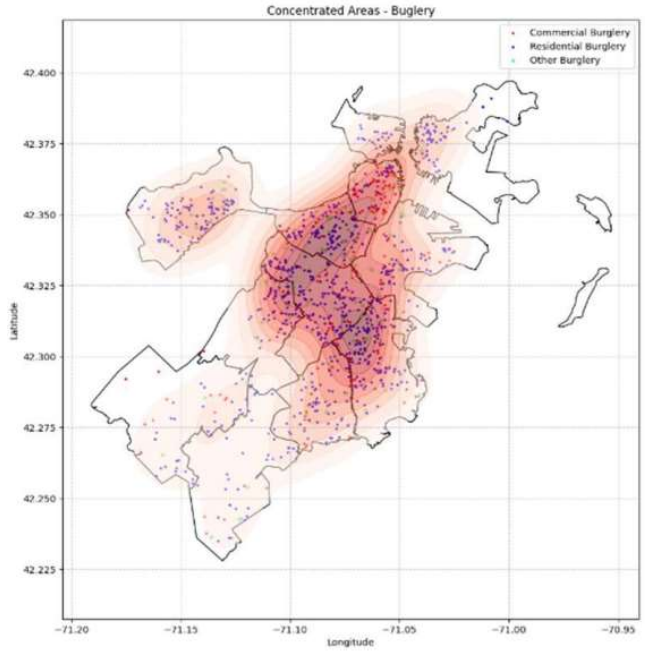
(a) Crime counts per District – Year 2018



DATA ANALYSIS

As larceny and burglary have been the most reported crimes in 2018, here we zoom in on the police districts D4, A1, C11 and B2 to have a more precise view of the density of crime locations.





DATA ANALYSIS

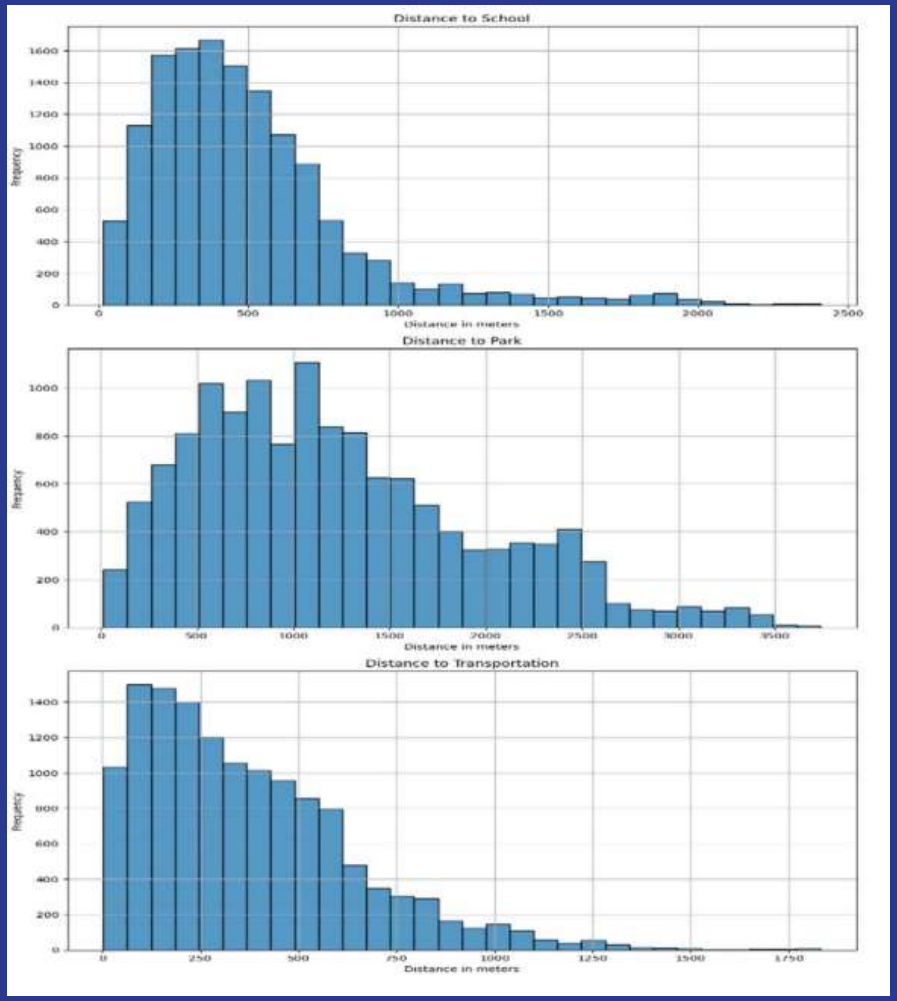
(d) D4 District

Using 2018 as a reference, it is interesting to see that police district D4, which had the highest number of reported crimes, experienced a concentration of repeated offenses, primarily along Boylston Street.

DATA ANALYSIS

To further gain insights into the neighborhood around the crime locations, proximity of the crimes to critical locations such as schools, public parks and transportation hubs were analyzed.

The analysis revealed that the majority of the crime locations are quite near to these public spaces (less than 1000m from crime location to public spaces).



CONCLUSIONS

Main Crime Types: larceny and burglary were the most frequent crimes across all police districts in Boston from 2015-2018.

Temporal Patterns: the crime counts peaked in 2017 and 2016, with Fridays reporting the highest crime rates (except in 2017). The busiest crime hours were between 14:00 and 16:00, while crime activity was the lowest from 1:00 to 7:00.

Geographic Trends: districts A1, B2, C11, and D4 were identified as crime hotspots. Crimes were frequently reported near public spaces like schools, parks, and transit hubs, indicating that many incidents occurred in active, non-isolated areas.

Actionable Insights: the analysis conducted provides valuable insights for law enforcement to optimize resource allocation and develop targeted interventions in high-risk areas.

Thank You

for Your Attention!