Exploring Spatiotemporal Patterns and Dynamics of New York City Crimes

Nicholas Tsioras

Abstract - An effective analysis of spatiotemporal patterns and dynamics of New York City crimes could be a crucial tool to allocate resources more effectively and on time. This could enhance public safety and reduce the number of crimes in New York City. This report focuses on the spatial and temporal dynamics of crimes across the five main boroughs of New York City. More than 1 million records of crimes in Bronx, Brooklyn, Manhattan, Queens and Staten Island were analysed during the years of 2014 and 2015. This paper aims to identify trends across space and time by applying geospatial analysis, time series analysis and clustering algorithms.

1 Problem Statement

In recent years, New York City has faced several challenges related to crime, across different neighbourhoods of the city. According to [7], New York is the 25th state with the highest total of crimes and had the 7th highest total of homicides in 2022 [8]. By understanding the insights of the spatiotemporal distribution of crimes, this report aims to provide valuable information that can contribute to the development of crime prevention strategies to improve the quality of life of New York City Residents. This study, aims to address the following questions:

- How do crime rates differ across the five boroughs of New York City over space and time? Which boroughs are more dangerous?
- How does the frequency of crimes during the day and during the week vary across the different boroughs of New York City? Which hours of the day are the most dangerous?
- Are there specific weeks or periods where crime spikes or declines? Is this the same across all boroughs of NYC?
- Which are the most common premises of crimes and how do they differ across the neighbourhoods?
- Which types of crime show high correlation over space and time?

The data was found in Kaggle and extracted from the NYC Open Data repository [9]. It contains crime records for 2014 and 2015 and the most important features include geographical attributes, temporal information, types of crimes, completion rate and description of premises.

2 State of the Art

Nasir et al. (2016) [6], in the paper 'Spatiotemporal Analysis of Crimes, A Case Study of Mardan City' study the spatial and temporal distribution of crimes while also exploring the effect that the weather has on the prevalence of crimes in Mardan city with the use of Geographical Information System (GIS). In this paper, the data was collected mainly from the three main police stations in Mardan include attributes such as crime type, location, and time of occurrence. The main questions that were addressed correlate to the seasonal distribution of crimes, the distribution of crimes on the bases of the weather and the identification of hotspots across the city. The main approaches that were applied in this paper include the mapping of Geographical Information System alongside the creation of a geo-database to provide a geographical context to the data for the further spatial analysis, the creation of a point map which shows the spatial distribution of the different crime types and the creation of thematic maps that answer the above questions under the parameters of the police station wise distribution of crimes and the distribution of crimes in relation to the weather.

Feng et al. (2016) [3], examine the spatial distribution of crime in Beijing, using advanced GIS techniques. The research focuses on Beijing and uses crime data from 2005 that include seven types of crime and seeks to understand the spatial and temporal patterns by creating quantitative representations of the geographical characteristics of crime. Although the approaches of this paper are like ours, the research focuses only on seven specific crimes in comparison to our project where we cover a larger range of crime types, which may require a more complex analysis to uncover patterns and trends. Also, the research can identify which socio-economic factors influence crime such as population density and economic conditions. In our case we don't have such data, so our analysis is quite limited in comparison to this paper.

Vivek and Prathap (2023) [5], provide an innovative approach to spatial and temporal crime analysis using social media data. The main questions addressed relate to the classification, visualization and time series forecasting of tweet counts related to crime in India using crime-related keywords from twitter posts. Although both studies involve spatio-temporal visualization of crime data, our project does not involve prediction, whereas this paper includes forecasting using different machine learning algorithms. The main focus of our project is to use a variety of visualization techniques that can help us uncover patterns and trends but in future work we could apply and validate machine learning algorithms for time-series predictions. The main visualization techniques used include heatmaps, scatterplots, choropleth maps and time series analysis line charts which are all applicable to our study. The main difference though with our study is the nature of the data. This study uses social media which might not always capture the complete picture of crimes in comparison to our study in which we use open-source data.

3 PROPERTIES OF THE DATA

- **Data Collection and Structure**

The dataset is considered a 'NYPD Complaint Data' which refers to a dataset that contains complaints and reports that have been made in the New York City police department in the years of 2014 and 2015. So, each row in the dataset represents a single crime that has been reported to the NYC police department. Each incident includes temporal information about the date and time of occurrence of the crime, geographical reference of the location that the crime was attended, the type of incident and the level of offense, a description of the premises, information about whether the crime was successfully committed and the jurisdiction responsible for the incident.

Column	Description	Type of Data
CMPLNT_NUM	Randomly generated persistent ID for each complaint	Numerical
CMPLNT_FR_DT	Exact date of occurrence for the reported event (or starting date of occurrence, if CMPLNT_TO_DT exists)	Date Object
CMPLNT_FR_TM	Exact time of occurrence for the reported event (or starting time of occurrence, if CMPLNT_TO_TM exists)	Time Object
CMPLNT_TO_DT	Ending date of occurrence for the reported event, if exact time of occurrence is unknown	Date Object
CMPLNT_TO_TM	Ending time of occurrence for the reported event, if exact time of occurrence is unknown	Time Object
RPT_DT	Date event was reported to police	Date Object
KY_CD	Three digit offense classification code	Categorical - Number
OFNS_DESC	Description of offense corresponding with key code	Categorical - String
PD_CD	Three digit internal classification code (more granular than Key Code)	Categorical - Number
PD_DESC	Description of internal classification corresponding with PD code (more granular than Offense Description)	Categorical - String
CRM_ATPT_CPTD_CD	Indicator of whether crime was successfully completed or attempted, but failed or was interrupted prematurely	Categorical - String
LAW_CAT_CD	Level of offense: felony, misdemeanor, violation	Ordinal - String
JURIS_DESC	Jurisdiction responsible for incident. Either internal, like Police, Transit, and Housing; or external, like Correction, Port Authority, etc.	Categorical - String
BORO_NM	The name of the borough in which the incident occurred	Categorical - String
ADDR_PCT_CD	The precinct in which the incident occurred	Categorical - Number
LOC_OF_OCCUR_DESC	Specific location of occurrence in or around the premises; inside, opposite of, front of, rear of	Categorical - String
PREM_TYP_DESC	Specific description of premises; grocery store, residence, street, etc.	Categorical - String
PARKS_NM	Name of NYC park, playground or greenspace of occurrence, if applicable (state parks are not included)	Categorical - String
HADEVELOPT	Name of NYCHA housing development of occurrence, if applicable	Categorical - String
X_COORD_CD	X-coordinate for New York State Plane Coordinate System, Long Island Zone, NAD 83, units feet (FIPS 3104)	Geographical - Numerical
Y_COORD_CD	Y-coordinate for New York State Plane Coordinate System, Long Island Zone, NAD 83, units feet (FIPS 3104)	Geographical - Numerical
Latitude	Latitude coordinate for Global Coordinate System, WGS 1984, decimal degrees (EPSG 4326)	Geographical - Numerical
Longitude	Longitude coordinate for Global Coordinate System, WGS 1984, decimal degrees (EPSG 4326)	Geographical - Numerical

Table 1. Description of Variables in the NYC Crime Dataset

- **Data Coverage, Precision and Resolution**

The dataset covers reported crimes in the years of 2014 and 2015. There are no reported incidents that extend over this period. The dataset consists of both spatial variables that have different levels of precision. It includes the X and Y coordinates from the New York State Plane coordinate system, which indicates the precise location of crimes measured in feet. It also has the longitude and latitude coordinates which offer a higher spatial precision and are measured by the Global Coordinate System (WGS 1984). The crimes are also categorized based on the boroughs of New York City and the NYPD precincts. These variables give a broader level of granularity to the spatial information.
The temporal information of the dataset includes the exact date and time of occurrence. This gives us the opportunity to visualize the hourly, daily, monthly, and even yearly distribution of crimes.

- **Data Quality and Accounting for Data Properties and Problems**

To assess the data quality and identify issues and problems in the variables the following aspects were considered:

❖ **Checking for outliers**

We examined the distribution of the longitude and latitude values using a scatterplot to visualize the spatial distribution of crimes to identify errors in the dataset. The outliers were also estimated with the use of the z-score. The figure below demonstrates a visual representation of the five main boroughs from where the incidents were reported. All the incidents fall within the boundaries of New York City and therefore there were no outliers found.

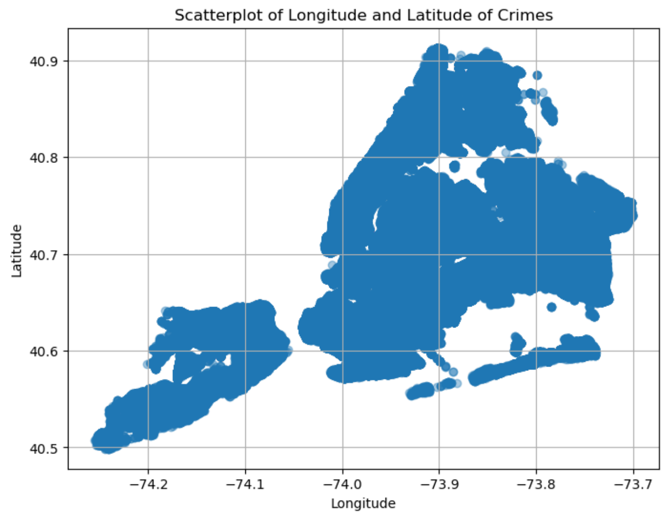


Figure 1. Scatterplot of Longitude and Latitude of Crime Incidents

❖ **Checking for Missing Values**

To check the missing values by column the isnull function was used in Python. The following figure shows the number of null values for each column of the dataset:

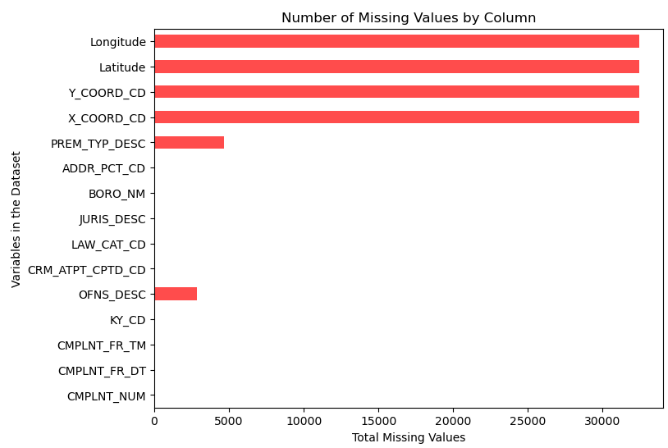


Figure 2. Number of Missing Values for Each Feature

The problem was addressed by dropping the rows that contain at least one null value rather than the actual columns. Although in some columns more than 30000 missing values seem a lot, if we take into consideration that there are more than 1 million rows, we would be dropping around 3-4% of the rows. So only a very small amount of data was dropped which should not affect the outcomes.

❖ **Validating the integrity of the geographical and temporal data.**

After checking the consistency of the temporal and categorical data, some rows had inconsistent values. The main problem was that there were a few crimes that were reported outside the period of 2014-2015. To address this problem, the data was cleaned, and those incidents were removed from the dataset. Finally, some feature engineering was required since the date and time columns were not initially datetime objects. To address this problem, the columns were transformed to date and time objects and the hour, day, month, and year were extracted and added to new, separate columns to make the further analysis easier.

4 Analysis

4.1 Approach

Python will be used for visualizations, as it is a flexible and effective tool with many versatile options to create a wide range of visualizations with human interpretation. Python will be used for clustering to identify seasonal patterns in crimes and specific time periods that are considered hotspots. Also, the clustering algorithm will be used alongside a Multi-Dimensional Scaling visualization to identify spatial and temporal similarities between different types of crimes.

➢ **Temporal Approach – Steps**

For the temporal analysis, the hour, day of the week, week, month, and year will be extracted from the original date and time columns and will be used to identify patterns for seasons, days of the weeks and hours of the day.

A. Create an area chart that shows the monthly crime rate for each of the five New York City boroughs.
B. Heatmap visualizations that show the hourly crime rates for each day of the week by borough.
C. Look at the top 10 premises of crimes and compare the hourly and daily distribution of crimes with the use of side-by-side heatmaps.
D. Use the appropriate colour scale for better visual clarity.
E. Evaluate and compare the dynamics and patterns with the use of probability density histograms with KDE lines.

➢ **Spatial Approach – Steps**

For the spatial analysis, the 2 main variables that will be utilized are the longitude and latitude coordinates. Since the number of observations in the dataset is quite large, heatmaps and density maps with high transparency will be used to display areas with high crime rates.

A. Visualize the crime incidents across New York with a heatmap to identify dense areas,
B. Apply transparency for an easier representation of the crime prone areas.
C. Look at the top 10 premises of crimes and compare the geographical spread of crimes.
D. Visualize the percentage of completed crimes by level of offense for each borough with the use of a bar chart.
E. Employ demographic knowledge to New York to help explain the spatial patterns and dynamics. This step requires human interpretation.

➢ **Spatial and Temporal Modelling Approach - Steps**

Partition-based clustering will be applied to aggregated weekly data and to aggregated data by types of crime alongside a multi-dimensional scaling scatterplot.

▪ **Clustering Weekly Crimes**

A. Aggregating and applying of partition-based clustering to the weekly counts of total crimes.
B. Evaluate the similarities of the clusters using an MDS projection of the cluster centroids to choose the optimal number of clusters.
C. Create a scatterplot of the weekly number of crimes with the appropriate colour scale.
D. Evaluate the results by visualizing the statistics of counts by clusters for each borough to address the cluster characteristics.

▪ **Clustering Types of Crimes**

A. Aggregating the types of crimes based on the average longitude and latitude values, average hour of the day that crimes occur.
B. Applying normalization to the data to have the same scale of values and choosing only crime types with more than 100 yearly occurrences for better balance.
C. Applying Partition-based clustering and evaluating the optimal number of clusters with the use of the elbow method [2].
D. Visualizing and evaluating the spatio-temporal similarities between different types of crimes with an MDS scatterplot to see which types of crimes tend to happen in the same area and at the same time.

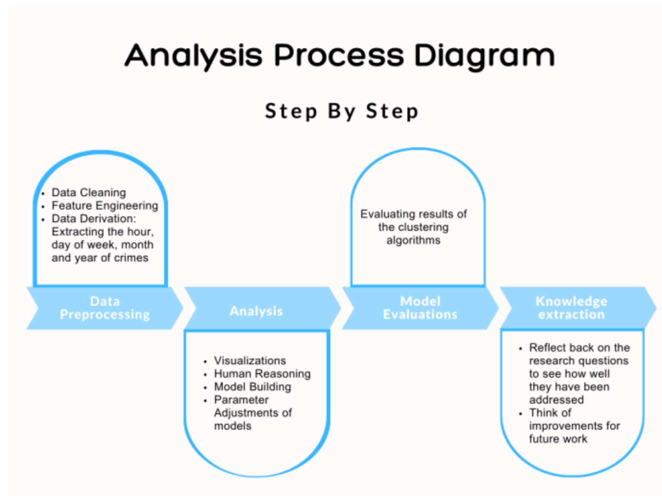


Figure 3. Diagram representation of computational and human interpretation steps applied

4.2 Process

4.2.1 Temporal Analysis

▪ **Data Preprocessing**

From the initial date and time columns, the hour, day of the week, number of the week since 01/01/2014 which is the starting date, the month and the year column were all extracted and added in separate columns. This way, we can analyse the temporal dynamics and patterns between the different boroughs of New York City. We want to find out which hours of the day the most crimes are committed across the boroughs, what is the monthly distribution of crimes for each borough and how similar or different the distributions are.

- **Monthly and Yearly Distributions**

In figure 4, we can see that there seems to be a seasonal pattern of crimes across all boroughs of NYC. Crime peaks tend to be during the summer months and reach their lowest during the Winter with a year-over-year consistency. This could be mainly because of the weather temperatures and their correlation with crime incidents. According to a review from Corcoran and Zahnow (2022) [4], most studies report a positive, linear relationship between the temperature and the number of crimes. Brooklyn seems to be the borough with the highest number of crimes consistently while Staten Island has far fewer crimes. Looking at the population graph below in figure 4 and figure 5, although the population stats are not the most accurate (from 2010), there still seems to be a positive relationship between the population density and the number of crimes committed. Since we did not have the population of the boroughs for 2014-2015 in our dataset, we could not use the crimes per 1000 population metric.

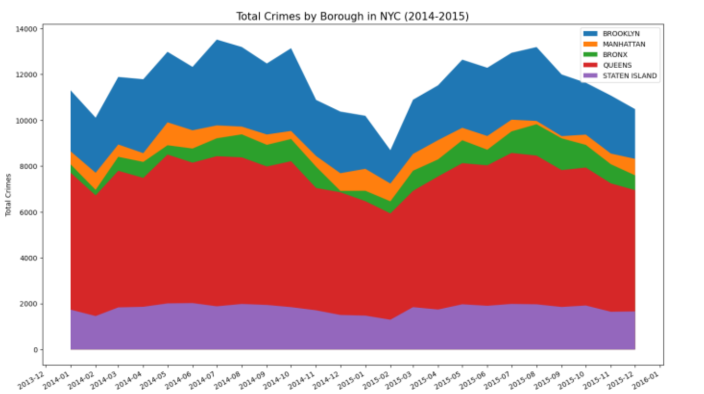


Figure 4. Monthly Distribution of Crimes across the boroughs of NYC

- **Hourly Distributions**

We then explored the hourly distribution of crimes for each of the five neighbourhoods to visualize potential patterns and similarities between them. The viridis colormap scale was used to distinct the hourly peaks of high and low crime rates. Across all boroughs, there seem to be some significant similarities. The crime rates tend to be low during the early morning hours with an increase from the late morning to the evening. Also, during the weekend there seem to be more crimes in the early hours of the morning. The patterns suggest that the most frequent crimes happen during hours where people are more active and outside of their homes rather than when people are most likely indoors.

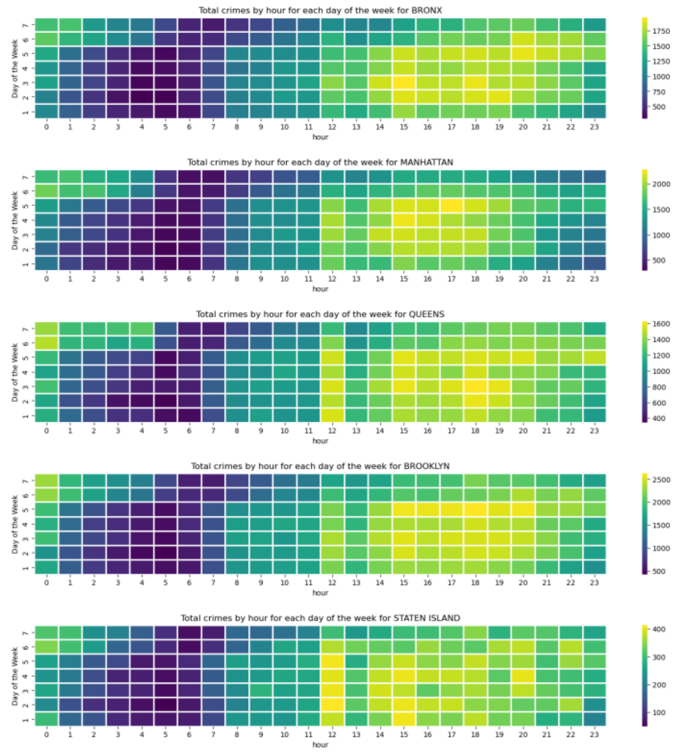


Figure 5. Hourly distribution of crimes across the boroughs of NYC

- **Probability Density Histograms**

From figure 4 and figure 5, we expect the KDE line to show peaks corresponding to the graphs.

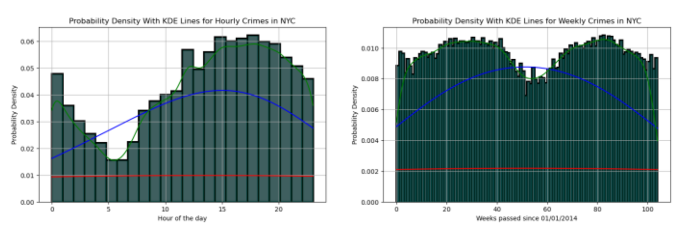


Figure 6. Probability density charts with KDE lines

As expected, the KDE lines show a higher peak during the late hours of the afternoon in the first graph and decrease during the first months of 2015.

4.2.2 Spatial Analysis

- **Spatial Distribution of Crimes in NYC**

After checking for outliers in the longitude and latitude coordinates with the use of a scatterplot and the z-score (figure 1) to make sure that there are no errors in the data and all crimes fall within the boundaries of NYC, the folium library in Python was used to create a density heatmap with the blue, orange and red colour scale being used to highlight the dense areas.

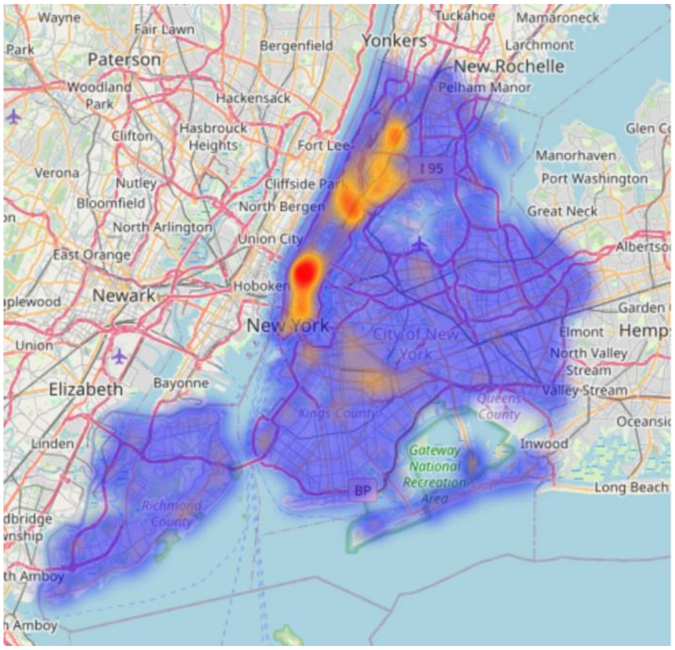
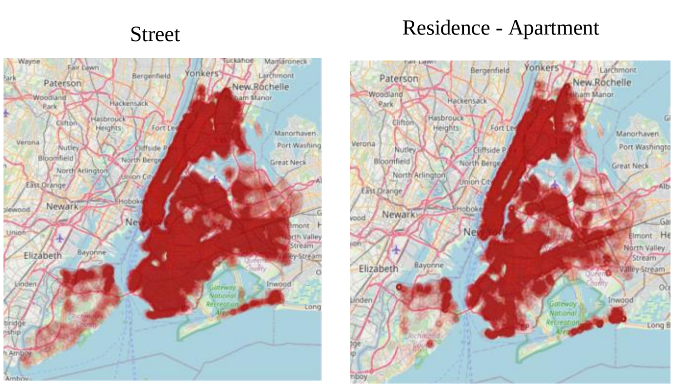


Figure 7. Spatial distribution of crimes in NYC

As expected, there are certain areas located mainly in Manhattan and Brooklyn that have high concentration of crimes and could be considered as hotspots. In terms of the spatial pattern that crimes follow, there seems to be a clear decrease in crime density as you move away into areas located far away from the city centre.

The heatmap showed that there is higher concentration of crimes in the central urban areas of the city which tend to have higher density of premises types that were identified as the top premises of crimes in figure 8. In terms of the correlation between human activity and crime rates, it seems that more crimes tend to happen in areas where there is higher human activity such as the streets and residents. The wide spread of crimes can be explained by the number of crimes in residential areas, which is balanced across all five boroughs.

To have a better understanding of which premises contribute mostly to the overall distribution of crimes in each area, the folium library was used to create density maps for some of the top premises. Transparency was applied to highlight the dense areas.



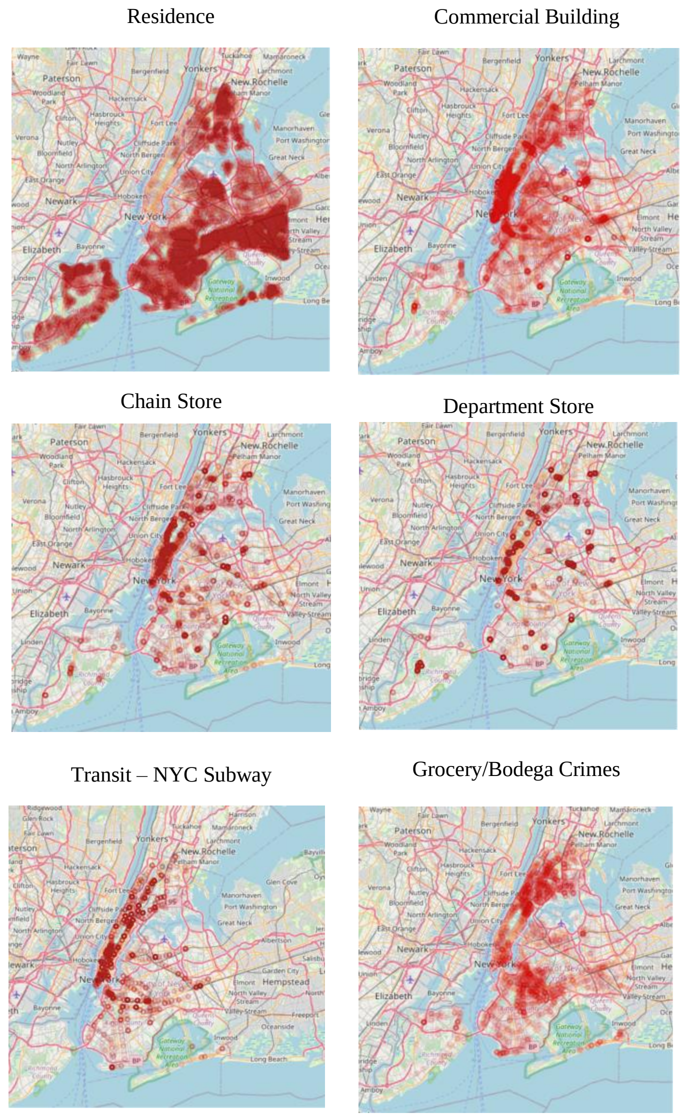


Figure 8. Spatial distribution of the top premises of crime

Certain types of locations such as commercial buildings, chain stores, department stores, grocery stores and the NYC Subway are not widely spread and show hotspots in certain areas mainly in Manhattan and Brooklyn, while crimes on the streets, at residential areas and at grocery stores tend to be more spread across the city, hence why there is a wide spread of crimes on the heatmap of figure 7.

4.2.3 Spatial and Temporal Modelling

- **Clustering the weekly crime rates across the boroughs of NYC**

Having investigated spatial and temporal patterns and dynamics of the crime volumes in New York City, we decided to investigate further the temporal seasonality of crimes to have a better understanding of the trends of each borough. The dataset was aggregated by the weekly crimes in each borough. After normalizing the data with the MinMaxScaler function, the k means clustering algorithm was used to group the weeks into clusters based on the total crimes by boroughs. After applying the algorithm multiple times and visualizing the cluster centroid values with a multi-dimensional scaling projection to identify the optimal number of clusters, we decided to split the weekly counts of crimes into 3 clusters, since the centroids had the biggest distance in the MDS projection (figure 9). In the scatterplot below, we can see a visual representation of the clusters.

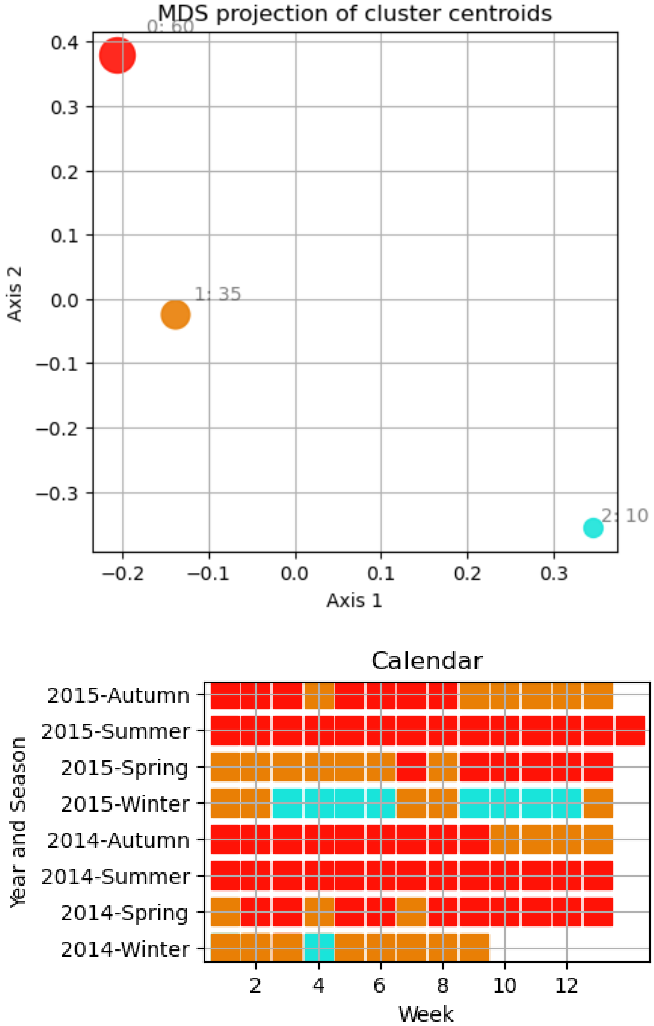


Figure 9. Clustering scatterplot of the weekly number of crimes based on the MDS projection of cluster centroids

Looking at this graph, we can identify some potential peaks of crime mainly during the summer and autumn weeks. What is interesting is that there are a few weeks during the 2015 winter with a distinctly lighter colour which could indicate a significant change of crime rates. There also does not seem to be a significant change in the yearly comparison of crimes rates.

- **Clustering the types of crime by location and time**

To have a better understanding of the spatio-temporal distribution of crimes in NYC, a further investigation was made in the spatial and temporal dynamics of different types of crimes. Because the dataset consists of 66 different types of crimes, it was difficult to visualize each type of crime separately using maps or time series plots. For that reason, we decided to cluster the crime types. The data was aggregated, and the average longitude and latitude values were used alongside the average hour of the most crimes committed. This way we can identify crimes that tend to happen in the same area and at the same time of the day. To make the analysis more balanced, only crimes with more than 100 incidents throughout the 2 years were selected. To find the optimal number of clusters, the elbow method was used just like in [2].

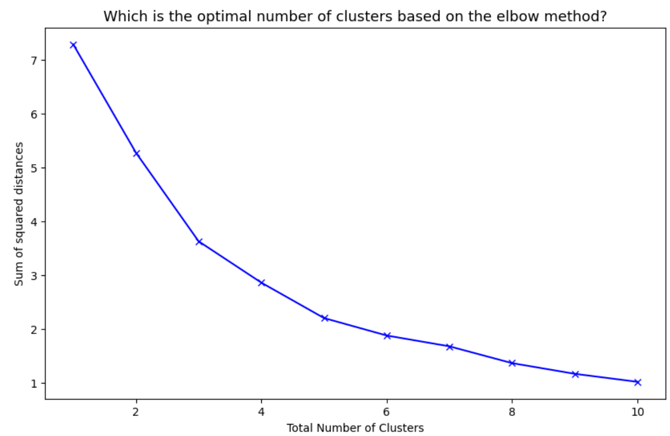


Figure 10. Line chart of the elbow method showing the sum of squared distances of cluster centroids

Looking at figure 10, the line chart seems to flatten out between the 5th and 6th cluster. We finally decided to split the crime types into 5 clusters. Finally, an MDS visualization scatterplot was created to see how the clusters are split and the similarities between the crime types. The scatterplot is shown in the next section.

4.3 Results

The most important findings that address the research questions are:

✓ <u>How do crime rates differ across the five boroughs of New York City over space and time? Which boroughs are more dangerous?</u>

Brooklyn and Manhattan experience the highest crime rates, with Manhattan having concentrated hotspots, while Staten Island experiences the lowest number of crimes. In terms of the seasonal trends, the five boroughs are quite similar, with the highest peak occurring during the summer and the lower peak during the winter months. The spatial distribution indicates that the main economic centres and the most populated areas, have higher crime rates, but we can be sure of that due to the absence of socioeconomic data.

✓ <u>How does the frequency of crimes during the day and during the week vary across the different boroughs of NYC? Which hours of the day are the most dangerous?</u>

All the boroughs seem to follow similar patterns in the hourly and daily distribution of crimes. Crimes increase during the early afternoon hours, and slowly start to decrease after midnight. As expected, during the weekend there are also high crime rates during the early hours of the morning.

✓ Are there specific weeks or periods where crime spikes or declines? Is this the same across all boroughs of NYC?

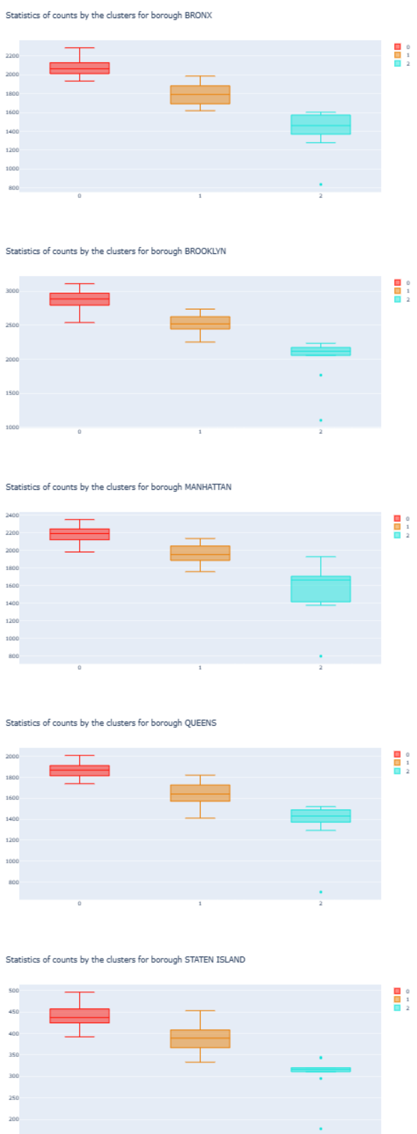


Figure 11. Statistics of total crimes by the clusters for each borough

- Crime significantly declines during the winter years of 2014 across all boroughs.
- There are more crimes during warmer months.
- Crimes tend to increase during the end of each season. The boxplots suggest that the intensity of these variations may differ within specific areas of each borough.

✓ Which are the most common premises of crimes and how do they differ across the neighbourhoods?

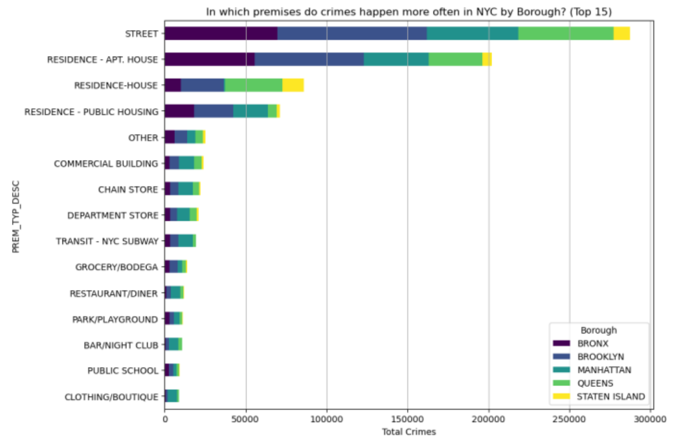


Figure 12. Bar chart showing the top premises of crimes

The streets and residential areas are common crime locations across all five boroughs. Crime rates in commercial and transit areas are higher in Manhattan. Brooklyn seems to have a broader spread of crimes, while Bronx shows a particular prevalence to residential crimes.

✓ Which types of crime show high correlation over space and time?

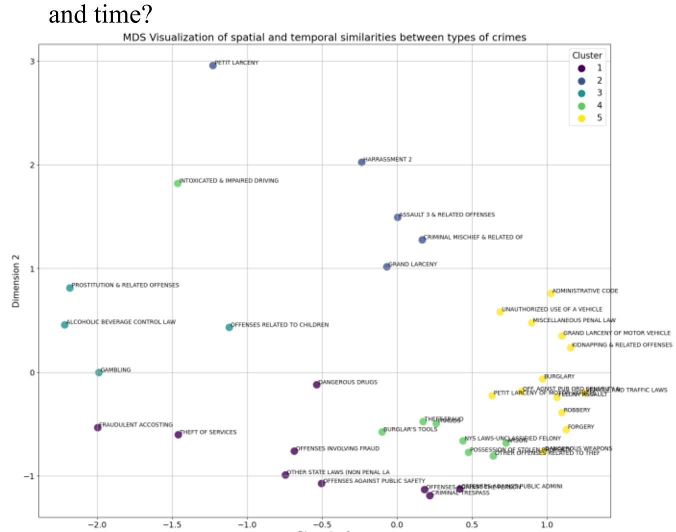


Figure 13. MDS visualization of spatial and temporal similarities between types of crimes

Cluster 5 mainly includes crimes that are property-related, while cluster 3 includes offenses that tend to happen during the late hours of the day. Clusters 1 and 2 is more widespread across the graph indicating that these crimes are more common and widespread. Crimes in cluster 3, are more likely to occur in public spaces.

5 Critical Reflection

As of August 2023, New York City is the 11th most populated urban city in the world [1]. As the population tends to increase, crime rates also tend to do so too, and for a large city like New York it is essential to understand the crime rates and how they differ across the boroughs of the city over space and time. In this paper, the goal is to investigate the frequency of crime rates across different periods of time, like for example which hours of the day tend to have higher crime rates. We also try to explain the variation or seasonality of crime rates over different weeks and periods. We also aim to find out which types of locations within the areas of the city tend to be more dangerous and how that explains the overall spatial and temporal distribution of crime rates.

Answering these questions is highly important for various reasons. Understanding how crime rates are different across various areas can lead to more effective policing strategies. Resources can target these specific areas and improve public safety. Insights into spatial and temporal crime patterns can also be very useful for urban planners and policy makers to invest more in certain areas with infrastructure investments. Knowing the most dangerous hours of the day and periods of time, can raise community awareness and encourage proactive measures. Finally, understanding the seasonal crime trends, alongside the knowledge of which type of crimes tend to happen in similar locations and in the same hours of the day can also be useful for further predictive policing efforts.

Having identified the main spatial and temporal patterns, in further work, acquiring data related to the sociodemographic characteristics of each borough such as the age, gender, level of income and education level, as well as economic data such as poverty rates, unemployment rates and economic growth and maybe some data related to the health conditions or the transportation or traffic patterns could all be very useful to further explain these patterns and understand them in detail. Then we would be able to explain which factors influence crime the most. Unfortunately, there was no such data in the datasets used so our main focus was to see which factors influence crime by clustering the different premises of crimes.

Table of word counts

Problem statement	237
State of the art	494
Properties of the data	543
Analysis: Approach	498
Analysis: Process	1154
Analysis: Results	256
Critical reflection	381

References

[1] B. Banerjee, "Ranked: The World's Largest Cities By Population," Visual Capitalist, 2023. [online]. Available: https://www.visualcapitalist.com/ranked-the-worlds-largest-cities-by-population/?fbclid=IwAR1MOZSVT4vJXfqGLEbByYDlG2_CmV5mp6SZ2jtngLqbtF2ioJUJFo9ANgQ . [Accessed on 2 January 2024].

[2] B. Saji, "Elbow Method for Finding the Optimal Number of Clusters in K-Means," Analytics Vidhya, 2023. [online]. Available: https://www.analyticsvidhya.com/blog/2021/01/in-depth-intuition-of-k-means-clustering-algorithm-in-machine-learning/?fbclid=IwAR0u9SqJQkDDZ16xy1ObnX0bx8QzQPI2Xo-iIupbu__6OFhByvAdZrVwEnw . [Accessed on 29 December 2023]

[3] J. Feng, Y. Dong, and L. Song, "A spatio-temporal analysis of urban crime in Beijing: Based on data for property crime," *Urban Studies*, vol. 53, no. 15, pp. 3223–3245, Nov. 2016, doi: 10.1177/0042098015612982. [Accessed on 3 January 2024]

[4] J. Corcoran and R. Zahnow, "Weather and crime: a systematic review of the empirical literature," *Crime Sci*, vol. 11, no. 1, p. 16, Dec. 2022, doi: 10.1186/s40163-022-00179-8. [Accessed on 3 January 2024]

[5] M. Vivek and B. R. Prathap, "Spatio-temporal Crime Analysis and Forecasting on Twitter Data Using Machine Learning Algorithms," *SN COMPUT. SCI.*, vol. 4, no. 4, p. 383, May 2023, doi: 10.1007/s42979-023-01816-y. [Accessed on 4 January 2024]

[6] M.J. Nasir, Pakhtunyar, and S. Mahmood, "Spatiotemporal Analysis of Crimes, A Case Study of Mardan City," in Pakistan Journal of Criminology, vol. 8, no. 1, Jan. 2016. [Accessed on 2 January 2024]

[7] S. Stebbins, "Crime in the US: All 50 states ranked by murder, violent crime rates," USA Today, 2020. [online]. Available: https://eu.usatoday.com/story/money/2020/01/13/most-dangerous-states-in-america-violent-crime-murder-rate/40968963/?fbclid=IwAR0t4xeIcrTrRlPBoyu5MCfBmVBswwBB9Om_qLxLhDE1-eOaPTTM14K8toc . [Accessed on 5 January 2024]

[8] "Total number of homicides in the United States in 2022, by state," Statista, [online]. Available: https://www.statista.com/statistics/195331/number-of-murders-in-the-us-by-state/?fbclid=IwAR3YGnRR6xgIG3rb_E9rZkYVCs2WWUYXhOvSJ6Y5m5xbEHgcEt8JsXumWbc . [Accessed on 5 January 2024]

Datasets

[9] "Crimes in New York City," Kaggle, [online]. Available: https://www.kaggle.com/datasets/adamschroeder/crimes-new-york-city [Accessed on 24 December 2023]