Discovering Spatio-Temporal Pattern of City Crime – Visual Analysis on Felony Crime in New York

Lau Hiu Ying, Hillie

Abstract— Pattern recognition has long been regarded as key role for crime prevention and reduction. Crime analysts and policy makers can formulate effective strategies and allocate resources with reference to spatial and temporal pattern of crime. In order the combat and prevent severe crime in New York City (NYC), this study analyzed Felony Crime data of NYC in previous 5 years (2015-2020) and discovered criminal hotspots pattern and temporal patterns with open criminal complaint data provided by New York Police Department (NYPD).

This study adapt a human-computer interactive appraoch to draw patterns from crime data, whereas computations and visualization are performed by Python libraries, and human to inform the decision of visualization methods, computational parameters and direction of this exploratory analysis. Density-based clustering algorithms, Grid Thematic Mapping and Density Heatmap are displayed to identify hotspots and demonstrates their associations with spatial features. Timeline analysis on moments of crime occurance demonstrates seasonality where crimes are mostly commited, while aoristic analysis showed hours of day when crime is mostly committed considering their timespan. Lastly, 3D visualization improved recognition of the displacement of hotspot over time, and suggested long-term hotspots in NYC in 3-D visualization. This inform strategic plans for police deployment.

--- ◆ ---

1 PROBLEM STATEMENT

City crime in New York has long been an issue since 1970s. Appreciation on the spatial and temporal characteristics of crime pattern is believed to contribute to crime reduction and prevention. In this study, the spatial and temporal pattern are analysed to provide insights for police forces. Felony criminal cases, which sentenced for more than one year in jail, are examined.

A hotspot is a geographical area with higher-than-average crime, as areas of crime concentration across the city [1][2]. Hotspots analysis identifies areas where criminal cases used to take place in, which helped allocate police resources.

In addition, it is crucial to understand crime may not stay in same space but displace from place to place over time. Offenders may change their activities in one area and find new targets in response to police activity [1]. Therefore, it is important to identify hotspots with consideration on temporal change.

Research questions are developed to investigate spatial and temporal pattern of felony crime in past five years.
(1) How do numbers of felony crimes changes along timeline?
(2) Is there a periodicity, or time cycles observed?
(3) What are the areas where felony crime used to take place in?
(4) How do the spatial patterns of felony crime displace/ remains over time?

Criminal complaint dataset provided by New York Police Department (NYPD) covers all valid felony cases reported since 2006, with details in accurate and precise locations and start/ end time of occurrence, as well as offense categories and victim profiles.

2 STATE OF THE ART

Clustering is common to identify hotspots, where spots of reported crime are partitioned into different groups based on similarities in distance between time/ locations. K-Mean clustering is common in traditional hotspots analysis [3][4][5]. Alkhaibari and Chung [4] performed clustering in NYC with same dataset to identify hotspots of Graffiti Crime, and to investigate reasons for stop and search by police in each cluster. Authors picked 2015 and time between 8p.m. and 8a.m. to compare different clustering algorithms including K-Means, agglomerative clustering. Evaluation is based on average silhouette score and optimal number of clusters. Visual analytics approach is only adapted to observe the spatial profile of clusters, for instance, whether clusters overlapped, and what boroughs covered by clusters.

However, above research did not analyse temporal pattern of NYC crime. Furthermore, in some clusters, internal variations within clusters are quite large, and may not deliver meaningful clusters that represent spatial autocorrelation. For instance, one cluster cover New Jersey, Staten Island and Brooklyn, where 3 regions are separated by sea. Moreover, some spots are closer to spots of another cluster. Density-based clustering is more preferred in my study.

Aryal and Wang [6] proposed SNN+ density-based clustering on spatial-temporal clustering on top of traditional Shared Nearest Neighbour clustering (SNN) by adding distance metric of semantic attributes in clustering, to find clusters of different sizes, shapes, and densities in noisy data. A weight factor is given to each attribute and add up all distances for final distance. Authors gave equal weighting to spatial and temporal data, and evaluate clusters by silhouette score, standard deviation and mean of cluster. Although spatial distance is calculated by Euclidean distance, they also

advised metric such as Haversine. Algorithm is reasoned by clustering crime data in Marland and taxi data in NYC.

This research delivers clear clustering for hotspot analysis and spatio-temporal pattern of city crime. Although my study does not consider other numeric attributes, I propose to adapt SNN algorithm for hotspot analysis and hoping to identify spatio-temporal pattern of NYC crime. Besides, as Haversine is suggested to provide quality clustering than Euclidean in larger areas [5] and NYC cover 5 large boroughs, my study uses Haversine to calculate spatial distance.

Leong, Chan and Ng [2] adapted an analytical framework from Gonzales, Schofield and Hart [1] to discover the displacement and similarities of crime hotspots pattern over time in Hong Kong. Grid Thematic Mapping Technique subdivides an area into regions and is used to determine whether spatial autocorrelation exists and to conclude potential disparate spatial relationship of event. Aoristic coefficient (0.0-1.0), which calculates probability of an event occurs within a given period within total time span is applied in temporal analysis to determine actual offense time. Displacement and similarities are quantified by weighted displacement quotients (WDQ) and area of correspondence (Ca). This study will adapt similar analytical framework and techniques to understand spatial autocorrelation and perform aoristic analysis to better observe periodicity. Nevertheless, I will use a visual approach to identify pattern.

3 PROPERTIES OF THE DATA

3.1 Overview

Dataset is opened to public by NYPD [8][9] where a total of 2,239,003 felony criminal complaints are recorded since 2006 and 707,090 complaints were recorded from 1 October 2015 to 30 September 2020 for this study. Criminal complaint refers to criminal case reported to NYPD and includes both completed and attempted crime. I assume all complaints were committed instead of false-reported crime in this study.

Longitude (-74.254939-73.700568), latitude (40.499025-40.912723), X and Y-coordinate of Coordinate System of exact crime locations are given. Data is also classified into 5 NYC boroughs and premises of location for supplementary analysis. This helped density-partitioned clustering by measuring exact distance and examined relationship with borough.

Detailed up to minutes (24-hour scale) of starting and ending date and time of occurrence, and exact date reported to NYPD are covered, which provides rooms for analysis of occurrence periodicity, crime timespan and trend over time.

3 attributes on victim profile (5 Age groups, 4 Gender, 6 Race) and 16 categories of criminal offense (offense < 1000 cases over 5 years are re-categorized as others) (Fig. 1) are given for examination. This help examining crime hotspots for each felony crime and vulnerable population. No numeric attribute is given; hence this study uses complaint count for pattern recognition.
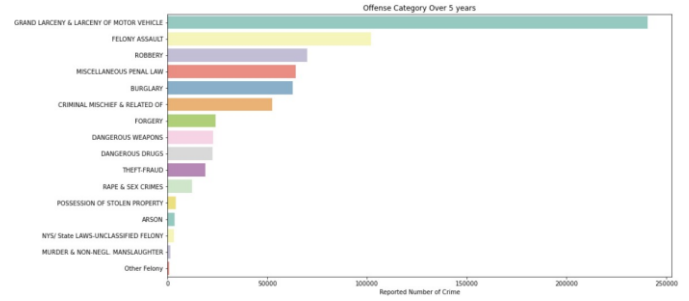

Fig. 1 Number of Felony Crime by Offense in NYC (Oct 2015 – Sep 2020)

3.2 Error and Missing Values

Missing values and input error are found in attributes. They are either filled by mapping unique ID (if any), classified to 'UNKNOWN' category (if already exists), or excluded from relevant analysis if missing count is relatively low (table 1).

ATTRIBUTE	NO. MISSING VALUE	NO. ERROR VALUE	DATA PROCESSING
VICTIM PROFILE	1	15	Classify as unknown and excluded from profile analysis
END TIME	87,546	0	Excluded from timespan analysis
COORDINATE	62	0	Excluded from spatial analysis
BOROUGH	1625	0	Mapped by coordinates
OFFENSE	103	0	Mapped by Offense ID
PREMISES	3993	0	Classify as unknown and excluded from examination in premises

Table 1: Data processing for missing values/ input errors

3.3 Outliners

A trend is plotted to show temporal outliners. The distribution showed daily number of felony crime approaches normal and extreme values are mostly larger than 600. These values are thus highlighted in temporal trend (Figure 2).

Abnormality is observed in 1-2 Jun 2020 (Mon-Tue) as number soared to 880 and suggested a short-term incident may occur these two days, which resulted in the increase. 'Burglary' offense is the major contributor. After checking news, 3 changes were found within the season: COVID-19 pandemic, state's bail reform law since Jan 2020, and mass protests opposed to police brutality on killing of George Floyd. Protests in NYC took place since previous weekend, and curfew was enforced on 1 Jun night. Lootings occurred overnight in Manhattan [7] and may cause the sudden rise of cases. Compared with last year pattern, cases are also concentrated in Lower and Upper Manhattan, yet are more concentrated (fig. not shown).

Outliners in spatial/ spatial-temporal analysis will be handled by density-based clustering adapted in this study. However, for pure temporal analysis, outliners detected in Jun 2020 will be imputed by interpolation.
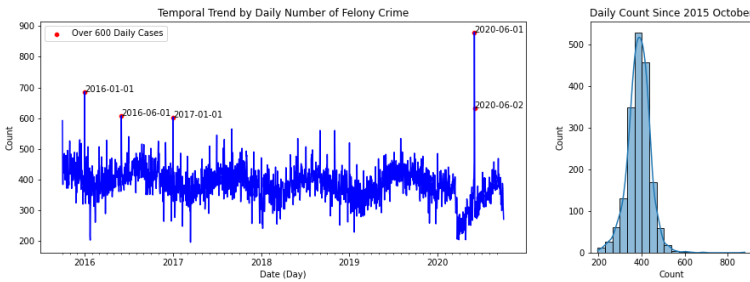

Fig. 2. Daily Number of Felony Crime in NYC (Oct 2015 – Sep 2020)
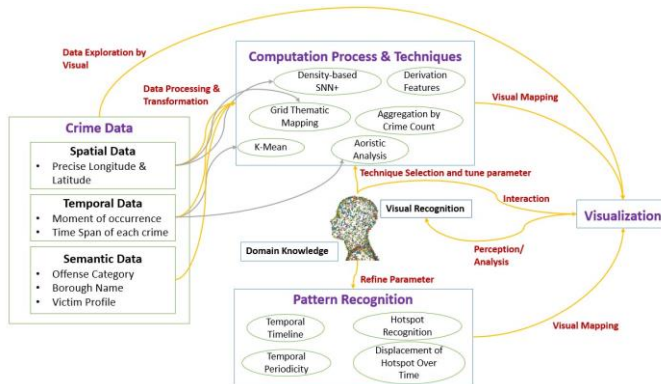
4 ANALYSIS

4.1 Approach

Fig. 3. Analytical approach on NYC Felony Crime Analysis

This study adapts visual analytics approach on recognizing spatio-temporal pattern, where human reasoning is involved in deciding visualizations and computational parameters such as clustering and aggregation. Python libraries are used for data processing, aggregating, clustering and visualizations, including Seaborn, Folium, Scikit-Learn, Pandas.

Crime pattern is recognized based on aggregation of crime number at different times and locations, while aggregation method is decided based on visualization effects (i.e., variance) among count/ maximum/ minimum/ mean/ median.

On top of the overview of pattern, offense categories may be partitioned and visualized if differences are found among them. Categories not sensitive on location/ time (e.g., Forgery, drugs, possession of stolen goods, etc.) are excluded.

Temporal Pattern (Q1 &2)
This study identifies temporal pattern based on two temporal categories: Moment of occurrence and time span of offense. The first refers to time when offenses start, while later refers to length of period that offenses may be committed within. Following steps are performed:

(1) Timeline analysis: Use line graph to observe overall trend of aggregated crime number by month/ week/ day and see if significant variances are present
(2) Seasonality analysis: Based on (1), further interpret and adjust temporal resolution on seasonality with heatmap
(3) Aoristic analysis on day/ hour resolution: Identify hours/ days when crime is mostly committed in with heatmap
(4) For (1)-(3), partition data by offense category to investigate on difference in pattern.
(5) Adjust appropriate aggregation methods, colour palette and temporal resolution by human reasoning from visualization throughout analysis.

Aoristic values, which estimates probability of offense being committed within temporal units, are calculated to identify temporal periodicity. This is because some offenses, such as murder, burglary, theft, reported only estimated time length of occurrence, e.g., across few days/ hours. Aoristic value takes consideration of probability and weight crime number within different temporal resolution.

Spatial Pattern (Q3)
Good cartographic design is important in hotspot mapping to identify areas suffer from crime, especially for large dataset and dense crime spots of this study. To focus on recent patterns, only data covering recent 2 years and chosen offense are visualized. 3 visualization methods are to be compared to show best visualization on map by latitude and longitude, as I adjust corresponding parameters (in "[]"). The purpose is to investigate detailed locations of crime.

(1) Density Heatmap [Radius and Blurriness Per spot], which dense area are identified
(2) Grid Thematic Mapping [Cell Size], which overcome the varying sizes and shapes of geographical boundary
(3) Density-based Clustering [Number of Neighbours, Radius of Cluster] based on Haversine Distance.

Partitioning by offense category will be performed and visualized to understand spatial variance between categories.

Human reasoning is necessary for tuning parameter by visualization, choosing appropriate visualizations and linking up NYC features with visualizations.

Spatial Displacement Over Time (Q4)
Density-based Clustering is performed to identify spatial displacement of hotspots over time, by calculating the Haversine Distance and day difference. Parameters are tuned on top of previous spatial clustering. This study filters noises and observe how clusters shifted over time.
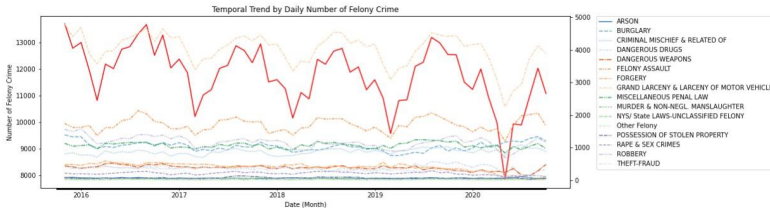
Temporal Pattern

Timeline Analysis

Fig. 4. Monthly Trend of Daily Number of Felony Crime by (Red Line), partitioned by offense (refer to Legend) (Enlarged figure in appendix)

Moment of Occurrence are depicted along a timeline in Fig. 4, and aggregated by day (Fig. 2), week and month to identify temporal pattern from Oct 2015 to Sep 2020. Periodic patterns are observed, yet considering stability of temporal trend, data is aggregated monthly. Monthly crime number fluctuates between 10-14 thousand complaints annually except in 2020. They used to peak before year end (around September and October), and dropped to the bottom in January/ February. Generally, monthly numbers of crime decrease slightly from 2015 to 2020, as peak and bottom are slightly lower than previous year.

2020 is an exception when compared by years, as crime number plummeted to 8000 in April and resumes slowly back to usual trend in July and August, presumably be related with Outbreak of COVID-19 since March.

Similar patterns are observed by offense categories. They are coloured by categorical palette as semantic attribute. Furthermore, crime numbers are stable, and no offense override others/ drop rapidly over time.
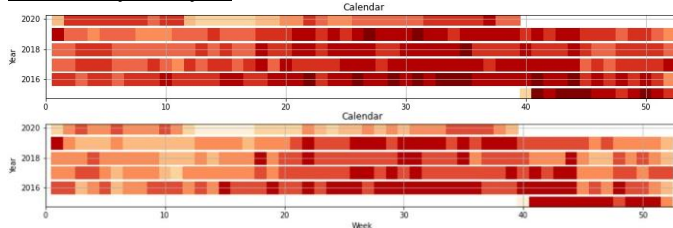
Seasonality Analysis

Fig. 5 Seasonality on Total Number of Crimes Per Week every year, colored by (1) Fig. 5a Value of Crime Numbers (2) Fig. 5b K-Mean Clusters of Normalized Crime Number (k=6). Darker colors indicate higher crime number.

Previous session indicates periodicity across seasons and month. To examine crime seasonality, heatmaps are used to compare seasonal differences over years. Both year vs. month and year vs. week are compared. As the later one gives more details on periodicity of crime, visualization is built by weekly variations. Moreover, as total count shows more variance in color (i.e., variance in number) than using

maximum/ minimum/ mean/ median, this statistical aggregation method is adapted.

Color of heatmaps is first decided by value of weekly crime number across years with sequential color palette (Fig. 5a). However, as deviations of weekly number are not as large between that of monthly number, pattern is not clear visually. K-Mean clustering is performed to accurately partition normalized numbers, while k=6 is determined by maximum number of sequential palettes, and similar count of crime number per cluster. Lastly sequential palette is assigned by order the value of cluster centroids.

High weekly numbers are observed between week 21 (4th week of May) to week 45 (November), while the opposites are found between week 1 to week 18 (4th week April), except first week in 2019. Seasonal variation that higher crime is found in summary and autumn, and lower crime in winter and spring.
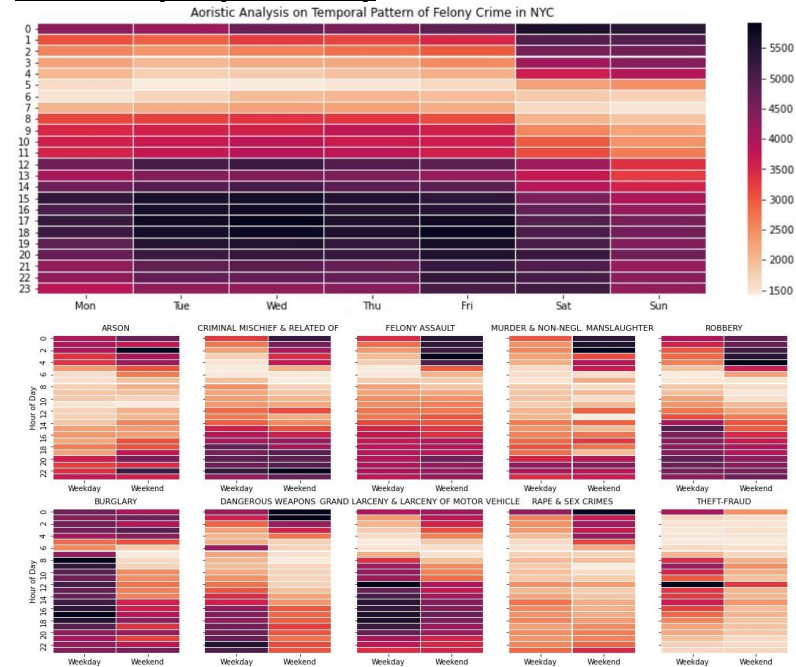
Aoristic Analysis by Hour of Day

Fig. 6 Aoristic Value by Hour of Day Per Weekdays
(1) Fig. 6a Overall Crime (2) Fig. 6b by Offense Category

Aoristic analysis takes consideration on timespan, for instance, if a burglary was reported between 10am on Monday to 9am on Wednesday (victim discovered crime), aoristic value on Monday is 14/ (14+24+9) hours*1 (complaint count). This provides a more accurate statistic for understanding hourly pattern. Count is hence used for aggregation.

Data is first visualized by comparing every weekday with hour of day in heatmap (Fig. 6a). Results suggest it is safest in early morning (5-7am) regardless of weekday, and most dangerous at midnight (11-12pm). In addition, weekday-weekend differences are obvious, particularly between 1-4am and in evening of days. Compared with other days, Sunday

night is the safest, followed by Saturday. This may be due to more people on street in weekend night.

This informs comparison between weekday and weekend on their average daily aoristic value. Since difference is visually clear, further examination is made by offense category (Fig. 6b) and borough. No special pattern is observed among boroughs (no figure), yet dangerous hours do vary with offense. Offense which leads to human injury/ killing such as rape, murder, arson, and assault are very likely present in 1-4am in weekend, while offense involving having others' possessions like Burglary, Larceny and Theft, likely to occur in weekday daytime. Particularly, most Theft happens in weekday noon, which may be due to most people go to work that time.
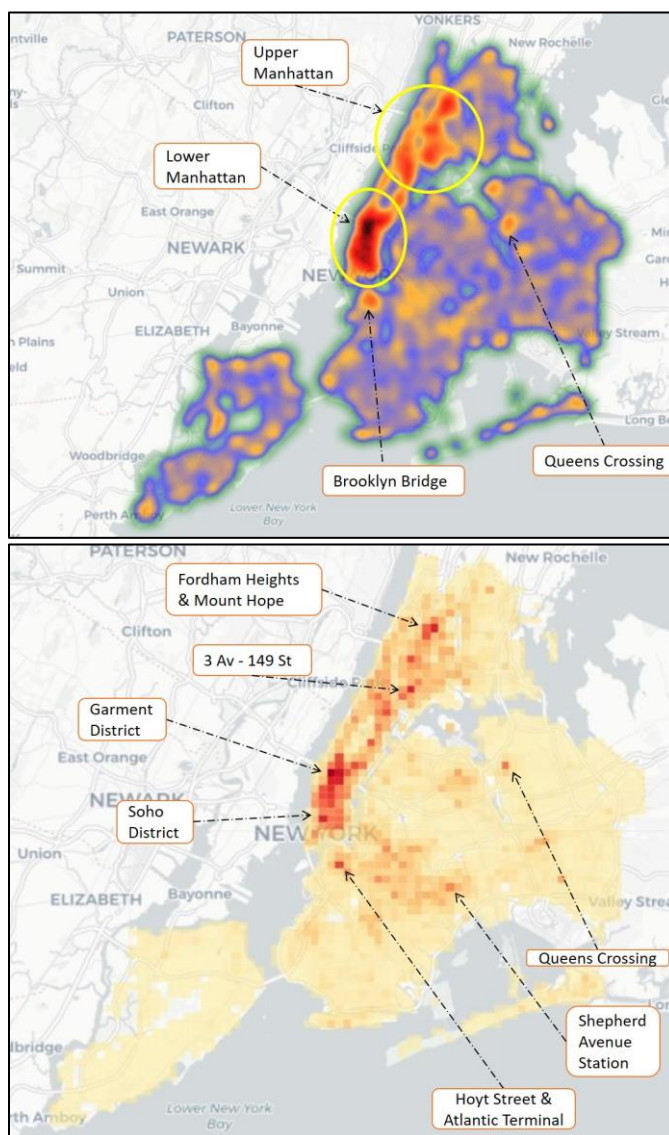
Spatial Pattern

Fig. 7 NYC Hotspot Mapping by Crime Number & Identification of Area by Visualization on (1) Fig. 7a Density Heatmap (spot radius = 6, blurriness =9), Black/ Red implies high density (2) Fig. 7b Grid Thematic Mapping (Cell Size = 0.62 km * 0.61km)

Hotspots are first mapped by plotting all valid spots on map, and adjust opacity, size, and blurriness of each spot, whereas dense spots in dark red or dark are hotspots (Fig.7a). Sequential-Multi-Hue colour palette (['Green', 'Blue', 'Red', 'Black']) provides additional hues for colour discrimination. In addition, due to dense but small area of Manhattan, both opacity and size are tuned to lower value to draw difference within Manhattan.

From Fig. 7a, hot zones are mostly in Manhattan (Upper & Lower), while two small hotspots are located near Brooklyn Bridge and area around Queens's Crossing (an Asian mall). This is because Manhattan is a Central Business District with dense population and busy streets, especially lower Manhattan is highly developed with popular districts.

It is hard to visualize details in dense area like Manhattan with density heatmap. In addition, colour/ density of map is highly correlated with number of spots in surrounding circle, which implies coastal area (mostly green) are depicted as low density. Grid Thematic Approach is adapted since it provides more details by grids and not impacted by nearby grids.

Initial cell size is set as around 1.5x1.5km, since scholars suggested started by dividing map coordinates by 50, which still resulted in coarse map. Yet if grids are too small, it may lose visual impact to identify hotspots (i.e., low values among all grids). Final grid size is 0.6x0.6km (Fig. 7b).

More hotspots are identified in Manhattan, including Soho, Garment District in Lower Manhattan, which are popular party/ fashions spots and next to popular spots such as Time Square, and Fordham Heights and 3 Avenue, 149 Street in Upper Manhattan. Hoyt Street in Brooklyn are closely surrounded by several subway stations, while Queens Crossing is one of largest malls in borough. Crime activities are associated with population flow.
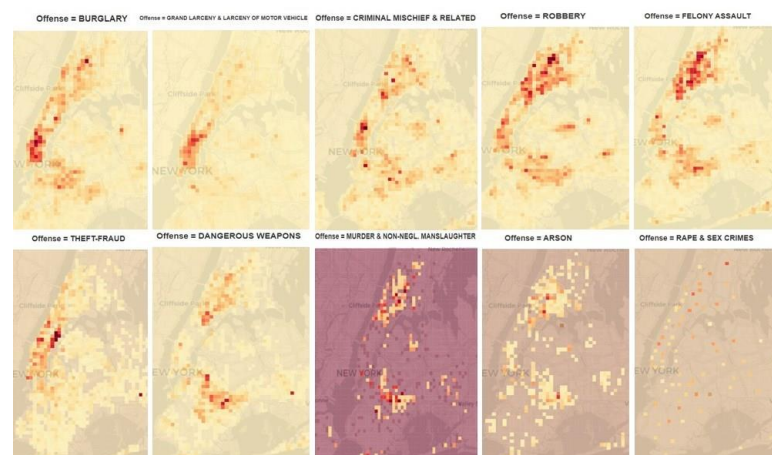

Fig. 8 NYC Hotspot Mapping by Crime Number with Grid Thematic Mapping (Cell Size = 0.62 km * 0.61km) by Offense Category (Enlarged figure in appendix)

To observe how crime activities associated with space, grid mappings are performed by offense (Fig.8). It is discovered that most Manhattan hotpots, especially Lower Manhattan, are contributed by possession of others' belongings such as Burglary, Theft, Larceny and Robbery. These crimes are mostly committed in Manhattan. On the other hand, crimes causing human death/ casualties, including Assault, Murder, Arson and Dangerous Weapons, are found mostly in Upper Manhattan and Brooklyn. This may be attributed to social issues such as homeless and poverty in two regions. Besides, no spatial pattern is found amongst Sex Crime.

Density-based clustering by Haversine Distance is adapted to further examine shape and locations of hotspots in Manhattan hot zones, while initial radius is set as 0.6km based on previous results, the radius should also remain low since Manhattan is small, high radius will only result in a large cluster. Due to high number of crimes, number of neighbours in cluster is also set high, or large number of clusters will be generated. Stopping criteria is set until street-level clusters are found.
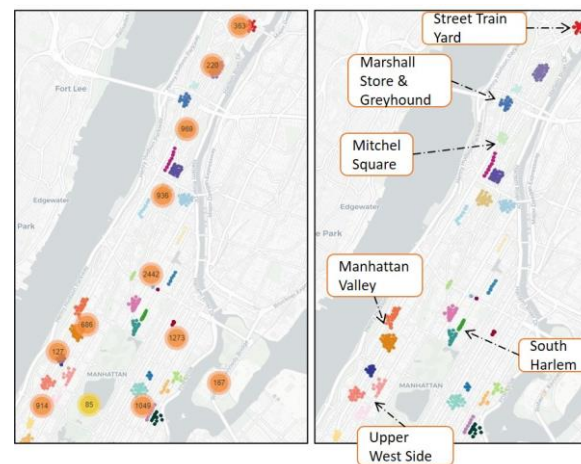

Fig. 9 Density-Based Clustering in Manhattan by Haversine Distance (1) Fig. 9a Overview of Manhattan clusters (2) Fig. 9b Zoom-in to Lower Manhattan (3) Fig. 9c Zoom-In to Upper Manhattan; Right: Clusters with spatial features labelled. Left: Number of crimes within the region on top of clusters, as many spots overlaps each other, hence, hotspots size do not indicate the hotspot is larger/ smaller than the rest; Cluster radius = 1km, neighbors = 50
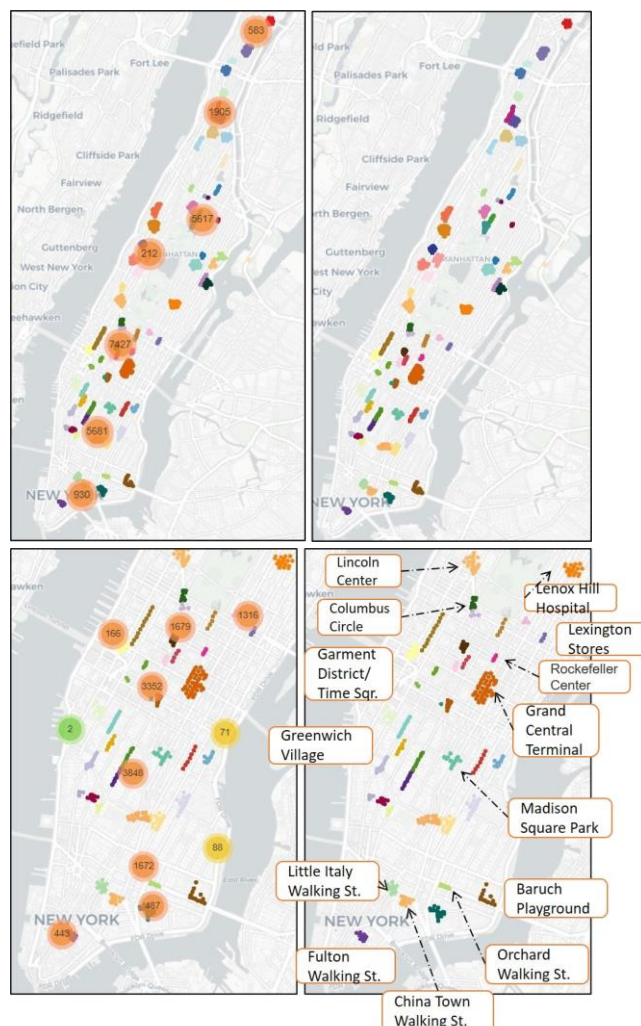
Fig. 9a suggested that most hotspots in Manhattan is in circle/ 2D shape which covers few street intersections, while only a few covers single linear street. Besides, hotspots are denser in Lower Manhattan as align with previous findings. On the contrary, hotspots in upper Manhattan, for instance, Harlem District, spread more across neighbourhood.

When looking into spatial features (Fig. 9b&c), several hotspots demonstrates their association with spatial features such as walking street, for instance, streets in Little Italy, Fulton and China Town share similar shape with hotspots. Hotspots are also associated with shopping centre and transport hub such as Lincoln Centre, Columbus District and Greyhound station in Marshall Stores.

On the other hand, some districts cover more hotspots, for instance, Greenwich Village, Garment District, Upper West Side and Harlem. It may be related to popularity of district, for example, the fashionable Garment, cultural and intellectual Upper West Wide and Greenwich. Harlem hotspots, in contrast, may be results of poverty and ghetto.
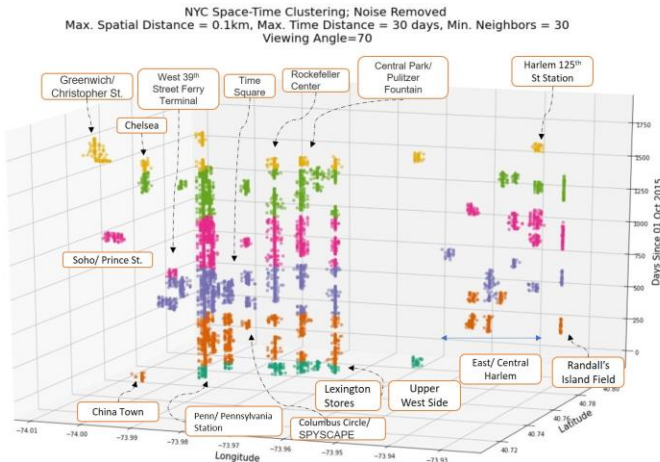
Spatial Displacement Over Time

Fig. 10 Displacement of Hotspots Over Years (by Number of Days since Oct 2015, Latitude & Longitude of Spots); Radius<=0.1km, Time Difference <=30 days, Neighbors >=30

Previous findings suggest temporal pattern by week, month and seasons, and discover hotspots cluster with 1km radius within dense spots of Manhattan. Therefore, initial parameters for Space-Time clustering are set to 1km with range (0.1-2km), time parameter as 7, 14, 30 and 60 days, and neighbours initially set as 50. Tuning stops when clear separation between clusters and coherence are generated with minimum noise. Clusters in categorical palette show annual shift of hotspots.

Long-terms hotspots are discovered as high values are reported annually, for instance, Columbus Circle/ SPYSCAPE, Randall's Island Field, Harlem 125th Street Station and Lexington stores. Some hotspots even report high value throughout years, including Pennsylvania Station, Rockefeller Centre, and place next to Pulitzer Fountain in Central Park. They are mostly transport hubs or popular sites with higher pedestrian flows.

Crimes may shift to area in vicinity for more opportunities [1]. This is observed in Harlem, where hotspots are shifting between East and Central Harlem. This echoes with previous study in Harlem, which crimes spread across neighbourhoods.

There are hotspots that solely appear in one year, such as Soho/ Prince Street (2018), China Town (2015) and Christopher Street in Greenwich (2020), which are all regions with high pedestrian flow. Yet it does take time in future to observe if they are noticeable hotspots.

4.3   Results

Appreciation on crime pattern contributes on crime reduction and prevention, as this study reveals clear temporal and spatial patterns on felony crimes. Temporal analysis indicates seasonal periodicity over different offense, whereas higher crime is found in Summer-Autumn, and lower in Winter-Spring (Fig.4&5). Crimes are more likely to occur in weekend midnight (11pm-4am) and weekday evening (4pm-11pm). This pattern varies by felony offenses (Fig.6).

Spatial analysis reveals hotspots by offense, as offense involving human casualties are mostly in Upper Manhattan and Brooklyn, while possession others' belongings occurs in

Lower Manhattan (Fig.8). Crimes are associated with popular sites with high pedestrian flow such as walking streets, malls and transport hub, or famous neighbourhoods like Garment and Greenwich (Fig.7&9).

As patterns by offenses are depicted, it informs the allocation of police forces on patrolling, for instance, police teams which handles assault/ murder cases could patrol more in hotspots in Upper Manhattan and Brooklyn in weekend midnight, and burglary/ larceny team patrols more in Lower Manhattan in weekday afternoon.

Hotspot displacement analysis suggests long-term hotspots whereas some exist throughout the whole year in past 5 years. This gives insights to strategic planning on resource allocation. For instance, set up police stand in these spots.

5   CRITICAL REFLECTION

This study offers a preliminary analysis on spatial-temporal pattern of felony crime in NYC by generalizing all felony crimes complaints in past 5 years with visual analytic approach, assuming no false complaints. Stable temporal patterns are revealed with 5 years data, while detailed spatial patterns disclosed with only 2 years data due to difficulties in performing complicated computations on 707,090 data.

Two drawbacks remain for spatial analysis. First, as Manhattan is far denser than other 4 boroughs in terms of spatial densities, crime numbers and pedestrian flows, it imposes difficulties in performing clustering across 5 boroughs and thus density-based clustering is only deployed in Manhattan in this study. It is advised to perform separate analysis on NYC and Manhattan alone in future.

Furthermore, as per the dense Manhattan, spots of current resolution (6 digits) are either plotted on same point on map, or having even distance between two points, which may result in inaccuracy in clustering. This study adapts three visualization methods for hotspot mapping. It is hard to identify and further investigate hotspots with Density Heatmap given the dense Manhattan. Clustering is best in identifying street-level hotspots, and further investigate with the assigned cluster ID, yet requires high computational power for this large dataset. Grid Thematic Mapping is useful to analyse hotspots with fixed grid size, yet reduction of cell size may impact visualization pattern. Considering map resolution, other approaches such as smoothing kernel density per grid cell, which measures intensity level per grid, may improve computational time and hotspots mapping in future.

Another drawback on pattern recognition as initial temporal analysis suggests year 2020 as an abnormal year under influence of the ongoing pandemic situation, and political conflicts. They are identified as outliners in this study and still disclose similar periodicity. Yet given latest political riots and pandemic, it remains uncertain whether future pattern would change. More analysis should be done as if more data in 2021 is accessible.

Parameters of Space-Time-Clustering (Fig. 10) were set strictly to separate clusters, resulting in fewer hotspots than spatial clusters in 2019-20 (Fig. 9) and only most severe hotspots are depicted. They can be eased in future for more hotspots mapping.

It is also found that spatial-temporal patterns differ by felony offences. If further examination is to be made, separate analysis by offenses is advised to identify different hotspots, which is believed to differ. Separation can identify more accurate spatial-temporal patterns, and reduces the current computational burden caused by large dataset.

Future analysis should also focus on the association of spatial-temporal patterns are associated with locational features, and question on 'What has resulted in this pattern?'. Multiple factors are suggested such as urban land uses, demographic and socio-economic data of neighbourhoods, phenomena, or economic activities in vicinity (e.g., bar closing time, building density, etc. Adding these semantic attributes into analysis will help predict potential hotspots, and preventive measures can be enforced.

Table of word counts

Problem statement	249
State of the art	493
Properties of the data	499
Analysis: Approach	499
Analysis: Process	1441
Analysis: Results	200
Critical reflection	480
Total	3861

REFERENCES

[1] Alberto R. Gonzales, Regina B. Schofield and Sarah V. Hart, Mapping Crime: Understanding Hot Spots, U.S. Department of Justice, 2005

[2] K. Leong, Junco Li, S. Chan and V. Ng, "Dynamic Pattern Analysis Framework for cooperative crime prevention," *2008 12th International Conference on Computer Supported Cooperative Work in Design*, Xi'an, 2008, pp. 1053-1058, doi: 10.1109/CSCWD.2008.4537125..

[3] S. V. Nath, "Crime Pattern Detection Using Data Mining," *2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops*, Hong Kong, 2006, pp. 41-44, doi: 10.1109/WI-IATW.2006.55.

[4] A. A. Alkhaibari and Ping-Tsai Chung, "Cluster analysis for reducing city crime rates," *2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, Farmingdale, NY, 2017, pp. 1-6, doi: 10.1109/LISAT.2017.8001983.

[5] A. V. Hazarika, G. J. S. R. Ram, E. Jain, D. Sushma and Anju, "Cluster analysis of Delhi crimes using different distance metrics," *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, 2017, pp. 565-568, doi: 10.1109/ICECDS.2017.8389500.

[6] A. M. Aryal and Sujing Wang, "Discovery of patterns in spatio-temporal data using clustering techniques," *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, Chengdu, 2017, pp. 990-995, doi: 10.1109/ICIVC.2017.7984703.

[7] C. ROBBINS, "Live Protest Updates: Protester Says Curfew Makes It Legal For Cops 'To Brutalize Us Now'." *Gothamist*, Jun. 01, 2020. Available at https://gothamist.com/news/live-protest-updates-nyc-faces-prospect-curfew-wake-soho-looting

[8] NYPD, NYPD Complaint Data Historic, 2006-2019, Published by NYC OpenData. Accessed on: 29/12/2020. Available at: https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i

[9] NYPD, NYPD Complaint Data Current (Year To Date), 1/2020 – 30/09/2020, Published by NYC OpenData. Accessed on: 29/12/2020. Available at: https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Current-Year-To-Date-/5uac-w243
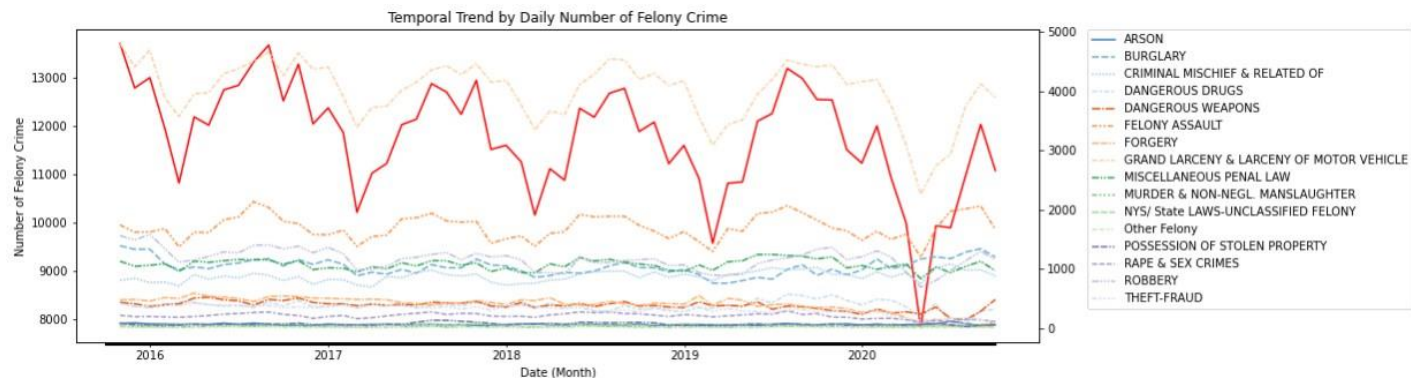
Fig. 4. Monthly Trend of Daily Number of Felony Crime by (Red Line), partitioned by offense (refer to Legend)
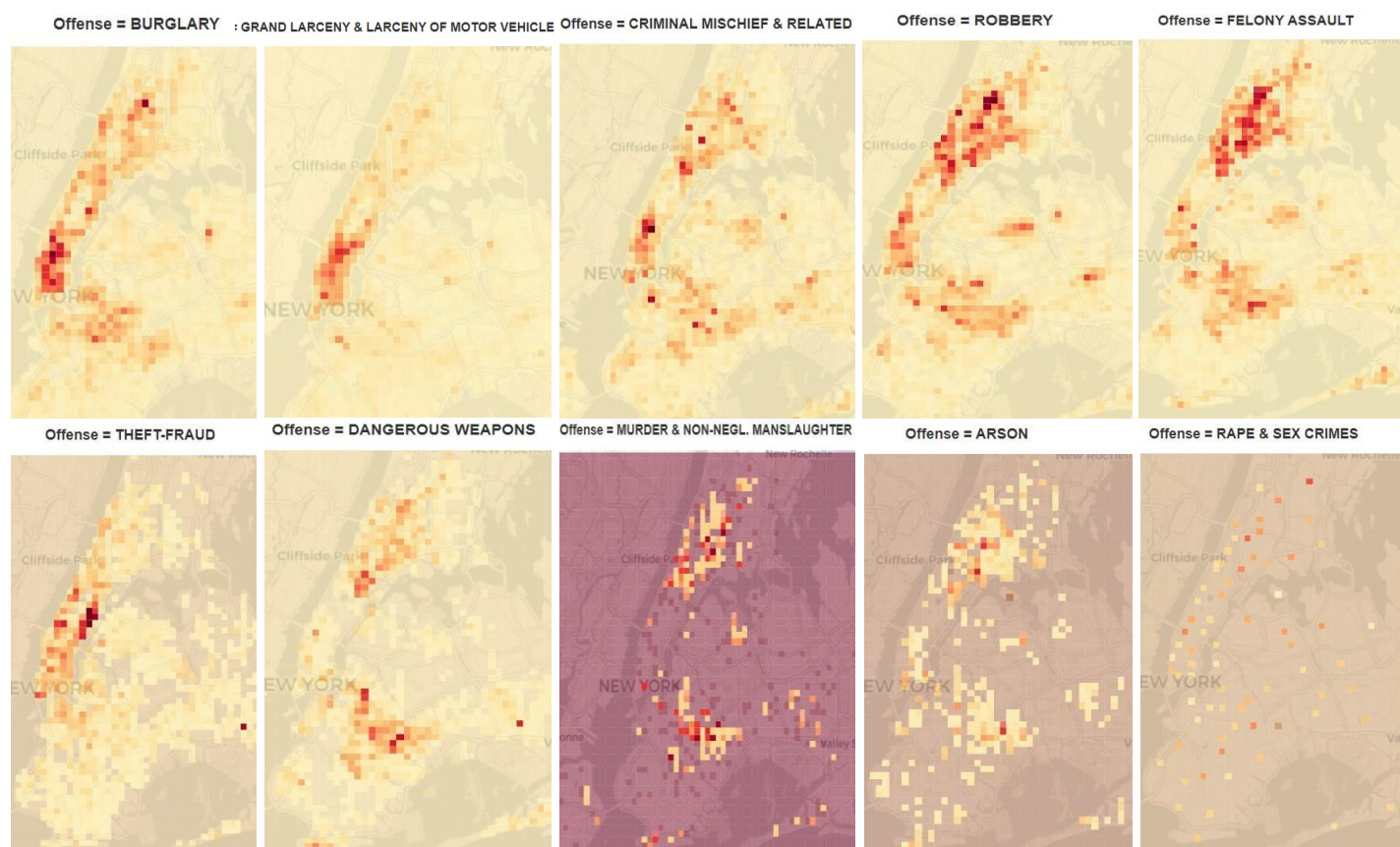

Fig. 8 NYC Hotspot Mapping by Crime Number with Grid Thematic Mapping (Cell Size = 0.62 km * 0.61km) by Offense Category