Spatio-Temporal Crime Patterns and Event-Driven Fluctuations in NYC

Jan Tietz
Email: jtietz3060@floridapoly.edu

Niklas Grau
Email: ngrau3035@floridapoly.edu

Manuel Rodriguez
Email: mrodriguezberdud318@floridapoly.edu

***Abstract*—This project investigates how crime patterns in New York City evolve across space and time, with a particular focus on the short-term effects of major public events. Using 2024 NYPD arrest data and NYC's permitted event records, we analyze 5 crime categories most relevant to residents and visitors: Robbery, Assault, Dangerous Drugs, Criminal Trespass, and Petit Larceny. We combine spatial clustering (DBSCAN) with kernel density estimation to compare persistent year-round hotspots with event-specific fluctuations. Our findings show that Manhattan's commercial and transit corridors remain the city's dominant crime centers, while major events unexpectedly reduce overall arrest volumes by an average of 36.7%. However, crime composition shifts markedly during these periods, with assaults becoming more prevalent and drug-related arrests declining. Spatial density maps reveal that some events redistribute crime toward central Manhattan or along event routes, while others produce uniform reductions citywide. Together, these results demonstrate that crime in NYC is shaped by both stable environmental factors and dynamic event-driven influences, highlighting the importance of multi-scale spatio-temporal analysis for improving public safety communication and supporting safer mobility decisions for travelers.**

***Index Terms*—Data Mining, crime analysis, event impact, spatio-temporal clustering, New York City**

I. Introduction

Urban crime remains a significant concern in large metropolitan areas, where the interplay of high population density, tourism, and public events often creates complex patterns of risk. New York City (NYC), one of the most visited destinations in the world, provides an especially compelling setting for studying the relationship between crime and urban dynamics. Each year, the city hosts a wide range of major events, including holidays, large sporting competitions, parades, and cultural gatherings, which attract millions of participants and visitors. While these events contribute positively to the city's cultural and economic life, they may also lead to shifts in crime patterns, particularly in categories such as Petit Larceny, Robbery, Assault, Drugs, and Criminal Trespass that directly affect public safety.

Although crime data is frequently collected and published by law enforcement agencies, it is often underutilized in providing actionable insights to residents, city officials, and visitors. Traditional crime statistics tend to emphasize long-term trends rather than short-term fluctuations associated with specific events. This creates a gap in understanding how crime risk evolves in response to changing urban conditions and where spatial hotspots may emerge during periods of increased crowd activity. For tourists in particular, safety is a critical factor that is often overlooked in travel recommendations, which usually focus on attractions and popularity rather than spatio-temporal risks. Incorporating safety metrics into the context of tourism and event activity can therefore provide more reliable and actionable guidance for mobility decisions. To address this gap, our project will combine NYC crime data with information on major events such as holidays, parades, and sporting competitions.

Contributions. The contribution of this project lies in: (i) introducing an *event-aware perspective* to the analysis of urban crime patterns in NYC, moving beyond long-term averages to highlight short-term fluctuations (ii) developing a framework for communicating spatio-temporal patterns in an accessible way that supports safer mobility decisions for travelers.

II. Related Work

Prior research shows that crime varies across space and time, and that such variations can be systematically modeled. Uittenbogaard and Ceccato [1] examined crimes in Stockholm using space–time scan statistics, showing that property crimes peak in commercial areas during the day, while violent crimes concentrate at night and on weekends, with some seasonal variation. This highlights the importance of incorporating both spatial and temporal perspectives when analyzing crime.

Delgado and Sánchez-Delgado [2] studied property crimes in Barcelona and demonstrated that crime levels follow cyclical seasonal patterns, with summer months showing higher intensity. Their results emphasize the predictive value of temporal regularities and the distinction between long-term seasonal cycles and short-term event-driven fluctuations.

In New York City, Kumar *et al.* [3] applied machine learning to NYPD data, identifying crime hotspots and forecasting trends while addressing ethical issues such as bias and privacy. Similarly, Nalluri *et al.* [4] analyzed transit safety and showed how crowd concentrations around transportation hubs reshape urban risk surfaces.

Other U.S.-based studies reinforce the role of advanced analytics in crime forecasting. Mohler *et al.* [5] modeled crime as a self-exciting point process, demonstrating how past incidents can predict near-future hotspots with high accuracy. Wang *et al.* [6] applied deep learning to Chicago crime data, showing that recurrent neural networks capture temporal dependencies more effectively than traditional statistical models.

These works highlight the range of methodological approaches that can inform predictive policing and public safety tools.

Event-driven crime research has also gained attention. Andresen and Malleson [7] found that crime patterns shift significantly around holidays and large sporting events, with spikes in thefts and assaults near entertainment districts. Ceccato and Uittenbogaard [8] further emphasized that festivals and other mass gatherings temporarily reshape urban risk landscapes, making time-sensitive modeling essential.

Finally, connections between safety and tourism have been documented in broader urban studies. Brantingham and Brantingham [9] introduced the concept of "environmental backcloth," noting that mobility patterns (e.g., flows of visitors to attractions) directly influence crime opportunities. More recent work by Yang *et al.* [10] analyzed Airbnb listings and neighborhood crime in NYC, showing how perceptions of safety influence traveler decisions.

In addition to commonly used methods such as DBSCAN or KNN, the ST-DBSCAN algorithm proposed by Birant and Kut [11] extends density-based clustering specifically to spatio-temporal data. While traditional methods often only consider spatial dimensions, ST-DBSCAN integrates both spatial and temporal neighborhoods. This makes it particularly suitable for detecting dynamic patterns such as event- or season-dependent crime clusters.

Together, these studies establish that crime clustering is both spatial and temporal, that seasonal and event-driven dynamics provide complementary insights, and that NYC offers the data richness to integrate these perspectives into safety-aware analyses relevant to travelers.

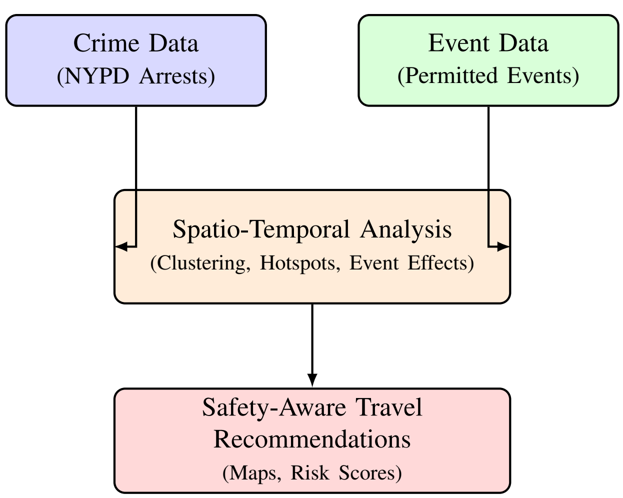

Fig. 1: Workflow of the study: combining NYC arrest and event data to generate safety-aware travel recommendations.

III. Planned Experiment

To evaluate the effectiveness of our proposed approach, we will conduct experiments that investigate the relationship between crime records and seasonal events in New York City. The central hypothesis is that crime rates, particularly theft, robbery, and assault, rise in both frequency and concentration during major public events compared to baseline periods. We further anticipate that these increases will be most pronounced in areas with high tourist activity, including Times Square, Central Park, and major transportation hubs.

IV. Data and Preprocessing

A. Data Sources

Two primary datasets from the NYC Open Data platform were utilized:

1) NYPD Arrest Data (2024): This dataset provides incident-level records with attributes such as arrest date and time, offense description, geographic location (borough, precinct, latitude/longitude), and charge classification. The data is publicly available via the NYC Open Data portal [16]. Because this project focuses on visitor safety, we filtered the arrest dataset to include crime categories most relevant to tourists and public gatherings, as suggested in prior urban safety studies [7]. We selected the following five high-impact categories and briefly define each below:

- **Robbery:** the unlawful taking of property from a person through force, intimidation, or threat. This includes incidents such as mugging and street theft with confrontation.
- **Assault and Related Offenses:** physical attacks or threats of harm, including fights, domestic violence, or attacks with a weapon, that result in injury or pose a risk to public safety.
- **Dangerous Drugs:** the possession, sale, or manufacturing of controlled substances such as cocaine, heroin, or synthetic drugs, often concentrated near nightlife or transit areas.
- **Criminal Trespass:** unauthorized entry into buildings or restricted spaces, frequently involving attempts to access residential, commercial, or event-related premises without permission.
- **Petit Larceny:** the theft of property valued under $1,000 without the use of force, typically including pickpocketing, shoplifting, or unattended item theft.

These definitions align with NYPD arrest classifications and help clarify the nature of offenses considered in our analysis. Filtering to these categories improves interpretability and aligns with prior findings that reducing noise enhances both pattern detection [3].

2) NYC Permitted Events Information (2024): The event dataset is maintained by the Office of Citywide Event Coordination and Management (CECM) and provides structured information on authorized public events in New York City [17]. Each entry includes event type, date, location, and duration.

For this study, we focused on five major, recurring events that attract large crowds and are representative of seasonal variation:

- Halloween (Oct 31)
- Thanksgiving (Nov 28)

- Independence Day (Jul 4)
- New Year's Eve (Dec 31)
- NYC Marathon (Nov 3)

These events were selected due to their high visibility, strong public participation, and potential to influence local crime dynamics. Their clear temporal and spatial signatures make them ideal candidates for studying short-term shifts in crime patterns using spatio-temporal clustering techniques.

B. Challenges and Considerations

As highlighted in prior work [3], [12], preprocessing crime data in dynamic urban contexts like NYC presents several key challenges:

- **Reporting Biases:** Crime reporting varies across boroughs and demographic groups, and holiday periods may distort arrest documentation.
- **Temporal Resolution:** Arrest timestamps may not reflect the exact time of the incident due to processing and administrative delays.
- **Confounding Factors:** Weather, spontaneous or unpermitted events, and shifting police deployments can influence crime patterns independently of formal public events.

Recognizing these limitations helps situate our findings in a realistic context and supports responsible interpretation of the results.

Figure 2 presents two complementary views of seasonal dynamics in New York City. Panel (a) displays the average weekly arrests for five major crime categories in 2024. Panel (b) shows the average weekly number of permitted public events in the same year. Both plots cover the full calendar year from January to December and report weekly mean values.

Initial exploration of the relationship between all permitted events and arrest patterns revealed no clear systematic correlation across the full event calendar. While permitted events peak dramatically during summer months (June–August), arrest volumes remain relatively stable throughout the year with only modest seasonal variation. This absence of a broad correlation suggested that most small-scale permitted events do not substantially influence citywide crime patterns. Consequently, we refined our analytical focus to five major, high-visibility events, Halloween, Thanksgiving, Independence Day, New Year's Eve, and the NYC Marathon, which are more likely to affect urban dynamics through large-scale crowd concentration, altered routine activities, and heightened police deployment.

V. METHODOLOGY

Our methodology is designed to systematically detect and interpret spatio-temporal crime patterns in New York City.

A. Density-Based Clustering Algorithms

To identify crime hotspots that evolve over time, we employed two **density-based clustering algorithms**, which group observations according to regions of high point density and treat isolated incidents as noise. These methods are particularly suitable for crime data because they make no assumptions about cluster shape and are robust to outliers. Below, we describe the two main algorithms used in this study: DBSCAN and ST-DBSCAN.

1) DBSCAN (Density-Based Spatial Clustering of Applications with Noise): The DBSCAN algorithm [13] detects clusters based on two key parameters: the neighborhood radius ε and the minimum number of points $MinPts$ required to form a dense region. Given a dataset $D = \{x_1, x_2, \ldots, x_n\}$, DBSCAN defines:

$$N_\varepsilon(x_i) = \{x_j \in D \mid \|x_i - x_j\| \leq \varepsilon\}$$

as the ε-neighborhood of point x_i. A point is a *core point* if $|N_\varepsilon(x_i)| \geq MinPts$. Clusters are formed by iteratively connecting core points and their density-reachable neighbors. Points that do not belong to any cluster are labeled as noise.

In our study, DBSCAN provides a baseline for identifying spatial crime hotspots based solely on geographic proximity. Its non-parametric nature allows the discovery of irregularly shaped crime zones, such as dense clusters around entertainment districts or transit hubs.

2) ST-DBSCAN (Spatio-Temporal DBSCAN): While DBSCAN captures spatial density, it ignores temporal correlations. The ST-DBSCAN algorithm [11] extends DBSCAN by integrating time as an additional dimension. A data point p_i is represented as (x_i, y_i, t_i), and its neighborhood is defined by both spatial and temporal radii:

$$N(p_i) = \{p_j \in D \mid dist_s(p_i, p_j) \leq \varepsilon_1 \ \wedge \ dist_t(p_i, p_j) \leq \varepsilon_2\}$$

where $dist_s$ and $dist_t$ denote spatial and temporal distances, respectively. The algorithm distinguishes between spatial density, temporal density, and attribute similarity (e.g., crime type) to classify points.

In theory, ST-DBSCAN should allow detection of clusters that are both geographically compact and temporally correlated. However, as discussed in Section VII-A, this approach proved impractical for our dataset due to data sparsity issues with single-day subsets containing only 200 arrests distributed across NYC's area.

B. Kernel Density Estimation

To visualize spatial crime distributions, we apply Kernel Density Estimation (KDE) [18], which estimates the probability density function of georeferenced arrest locations. For n observed crime locations $\mathbf{x}_i$, the density at any point $\mathbf{x}$ is:

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \tag{1}$$

where $K(\cdot)$ is a Gaussian kernel and $h = 0.01$ degrees is the bandwidth parameter. We compute density surfaces for event and control days using identical parameters and geographic extents, enabling direct comparison through difference mapping.

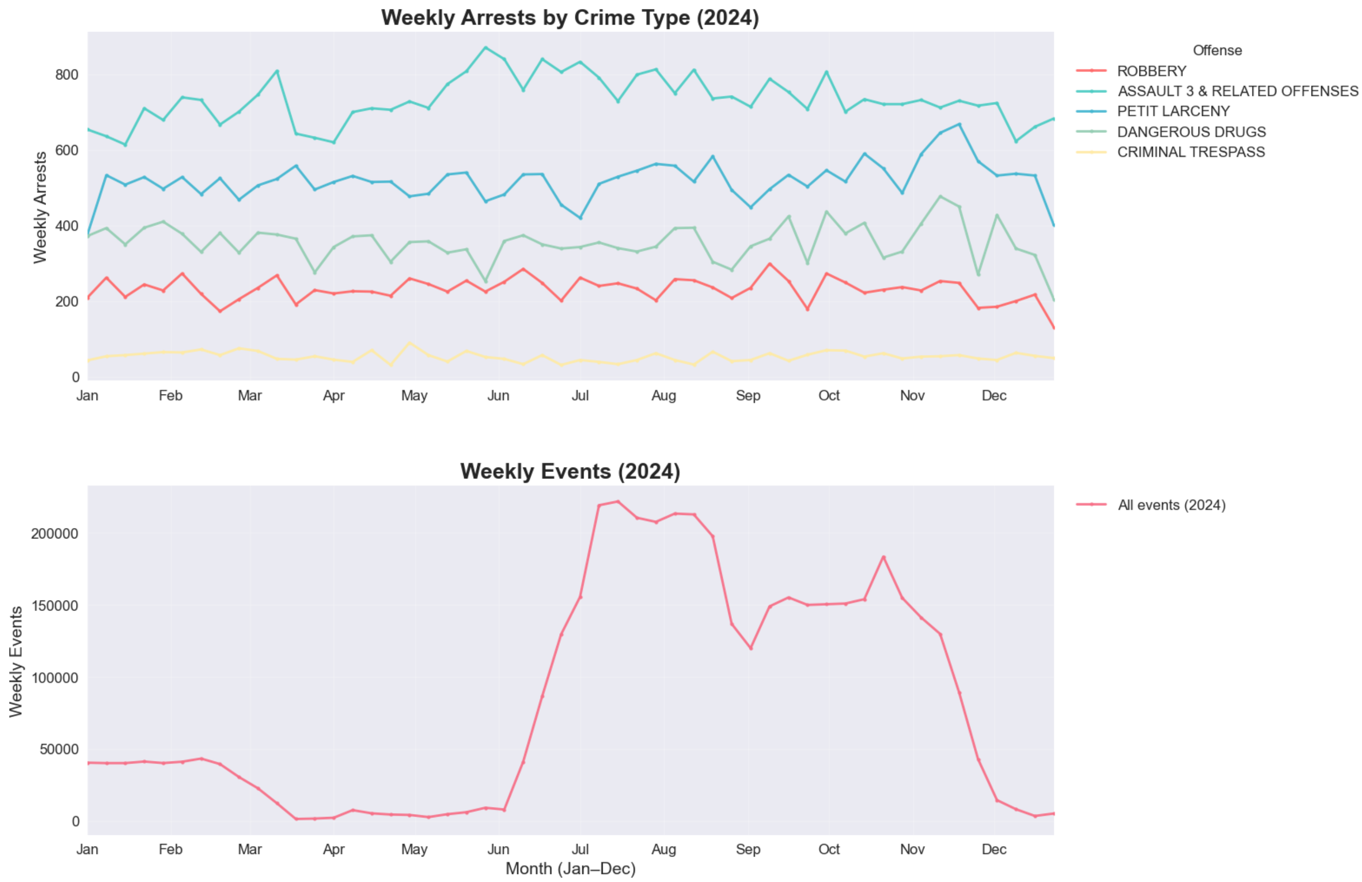

Fig. 2: Comparison of seasonal dynamics in crime (a) and public events (b) in New York City. (2024)

Methodological Rationale. We selected KDE over ST-DBSCAN for temporal event analysis because ST-DBSCAN proved impractical for our sparse single-day datasets. With approximately 200 arrests per event day. KDE offers superior robustness for sparse data, enables direct event-vs-control comparisons. This reveals spatial shifts in crime patterns attributable to event impacts.

We initially attempted to apply spatio-temporal clustering using ST-DBSCAN on single-day subsets (e.g., NYC Marathon on November 3, 2024). However, this approach proved unsuitable due to the sparse distribution of 200 daily arrests across NYC's large geographic area. Instead, we employed kernel density estimation to analyze temporal variations in crime patterns during major events compared to control periods.

VI. Results and Analysis

Our analysis follows a three-stage approach: (1) baseline spatial clustering to identify persistent crime hotspots across 2024, (2) multi-event comparative analysis to examine crime distribution shifts during major public events, and (3) spatial density analysis to detect geographic redistribution patterns during events. This section presents the findings from each stage and discusses their implications for understanding urban crime dynamics.

A. Data Preprocessing and Sampling Strategy

The complete 2024 NYPD arrest dataset contained over 350,000 records. After filtering to five crime categories (Robbery, Assault 3, Dangerous Drugs, Petit Larceny, Criminal Trespass) and removing invalid coordinates, approximately 100,000 arrests remained.

To manage computational constraints while preserving spatial and crime-type distributions, we applied stratified sampling using an 8×8 spatial grid (64 bins) combined with crime categories, creating up to $64 \times 5 = 320$ strata. where $f = 30{,}000/100{,}000$. This reduced the dataset by 70% (to 30,000 records) while preserving crime-type distributions within 0.5 percentage points and maintaining geographic coverage across all NYC regions.

B. Baseline Spatial Clustering with DBSCAN

To establish a baseline understanding of persistent crime patterns, we applied DBSCAN to the complete 2024 arrest dataset using only spatial coordinates. We employed the haversine distance metric with parameters $\varepsilon = 0.0001$ radians and $MinPts = 150$ arrests. These stricter parameters were selected to identify highly concentrated, dense crime clusters while filtering out more dispersed criminal activity. The low epsilon value ensures that only arrests occurring in very close

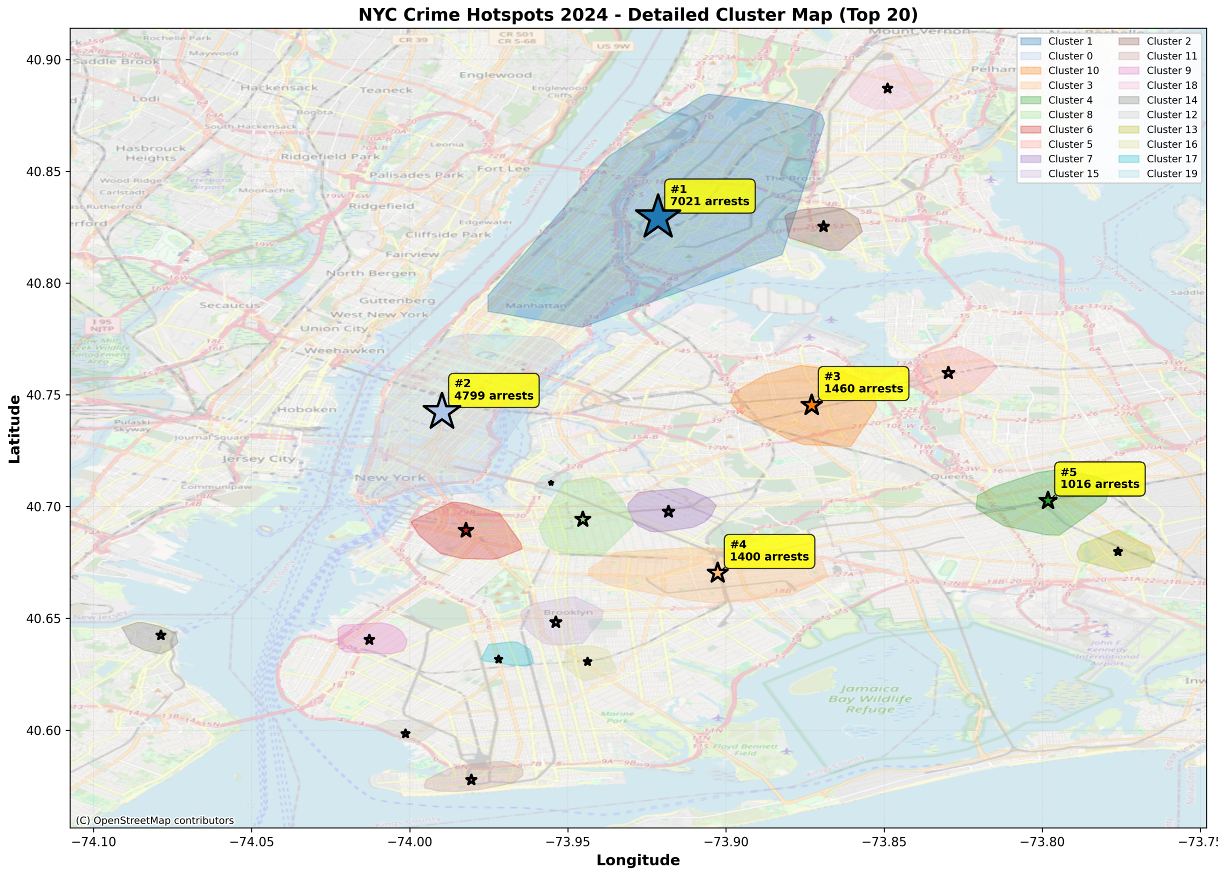

Fig. 3: Baseline spatial crime clusters in NYC for 2024 using DBSCAN ($\varepsilon = 0.0001$, MinPts = 150). The map shows the top 20 persistent crime hotspots across all five analyzed crime categories, overlaid on an OpenStreetMap basemap for geographic context. The five largest clusters are labeled with their total arrest counts and centroids marked with star symbols. Manhattan exhibits the highest concentration of large-scale clusters, with Cluster 1 in Midtown representing the dominant hotspot.

geographic proximity are grouped together, revealing the most intense hotspot cores.

Figure 3 presents the resulting spatial crime clusters across New York City, with cluster boundaries overlaid on an OpenStreetMap basemap to provide geographic context. The analysis identified 24 distinct clusters, with the visualization highlighting the top 20 by arrest count. The clustering results reveal several key patterns:

Manhattan Dominance. The two largest clusters are located within Manhattan, accounting for approximately 12,000 arrests combined. Cluster 1, centered in Midtown Manhattan near Times Square and major transportation hubs, represents the single largest hotspot with over 7,000 arrests, nearly 50% more than the second-largest cluster. This extreme concentration aligns with prior findings that high-density commercial and tourist areas experience elevated crime rates [3].

High-Density Core Hotspots. The stringent parameters successfully identify intense crime concentration zones, filtering out more dispersed activity. This reveals the true "hotspot cores" where criminal activity is most systematically concentrated, as opposed to broader areas with elevated but less focused crime patterns.

Borough Distribution. While Manhattan dominates, crime hotspots exist throughout the city. Cluster 7 in Queens and several smaller clusters in Brooklyn indicate that criminal activity concentrates in commercial and transit-oriented neighborhoods across all boroughs, though with lower intensity than Manhattan's core areas.

Noise Points. Approximately 40-50% of arrests were classified as noise (not belonging to any cluster) due to the strict clustering parameters, representing either isolated incidents or lower-density crime activity that does not meet the threshold for highly concentrated hotspots.

Visualization Enhancement. The use of OpenStreetMap as a basemap layer (implemented via the `contextily` Python library) provides essential geographic context, allowing immediate identification of landmarks, neighborhoods, and infrastructure that correspond to detected crime clusters. This integration enhances interpretability and supports practical decision-making for resource allocation.

The baseline DBSCAN analysis provides essential context for interpreting event-driven fluctuations. While Figure 3 reveals the geographic distribution of crime hotspots, the composition of criminal activity varies substantially across

TABLE I: Crime Category Distribution in Major Spatial Clusters (DBSCAN, 2024)

Cluster ID	**Size**	**Robbery (%)**	**Assault (%)**	**Drugs (%)**	**Larceny (%)**	**Trespass (%)**
1	7,021	13.4	40.9	23.8	19.6	2.3
0	4,799	10.5	22.2	15.8	**47.9**	3.6
10	1,460	14.2	41.3	11.9	31.1	1.4
3	1,400	**15.3**	**44.4**	29.1	7.4	**3.7**
4	1,016	9.3	28.0	**36.4**	24.6	1.8
Other Clusters (19 total)	3,304	12.8	32.5	17.3	31.9	**5.5**
Citywide Average	19,000	12.4	34.2	20.1	28.7	4.6

clusters. Table I presents a detailed breakdown of crime categories within the five largest clusters, which collectively account for 82% of all arrests in the baseline analysis.

The five major clusters exhibit distinct crime profiles that reflect their underlying urban contexts. Cluster 3 emerges as the most violent hotspot, with assault offenses comprising 44.4% of arrests, significantly above the citywide average of 34.2%. This cluster also shows elevated drug-related arrests (29.1%) and the highest robbery rate (15.3%), suggesting a concentrated area of serious interpersonal and drug-related crime.

In contrast, Cluster 0 demonstrates a markedly different pattern, with petit larceny dominating at 47.9%, nearly double the citywide average of 28.7%. This larceny-focused profile is consistent with commercial or tourist-heavy areas where pickpocketing, shoplifting, and theft of unattended items are prevalent.

Cluster 4 stands out for its drug-related arrests, which constitute 36.4% of all incidents, which is the highest proportion among major clusters and substantially above the citywide rate of 20.1%. This suggests a persistent drug market or trafficking corridor within this geographic area.

These findings demonstrate that crime hotspots are not homogeneous in their criminal profiles. Effective policing strategies must therefore be tailored to the specific crime mix of each cluster rather than applying uniform approaches across all high-crime areas.

C. Multi-Event Comparative Analysis

To investigate whether major public events alter crime patterns, we conducted a comparison of arrest data from five events against baseline periods. For each event, we extracted arrest records from the event day and compared them to 3-4 control days of the same weekday type from surrounding weeks. This approach controls for day-of-week effects while isolating event-specific influences.

1) Total Arrest Volumes: Contrary to our initial hypothesis, the analysis revealed a consistent and substantial decrease in total arrests during major public events compared to baseline periods. Table II summarizes the findings across all five analyzed events, showing an average reduction of 36.7%.

This "holiday effect" manifests most dramatically during Thanksgiving (−61.6%) and New Year's Eve (−43.0%), both of which are traditional family-oriented holidays. The NYC Marathon, while attracting massive crowds, shows the smallest reduction (−14.2%), suggesting that large-scale sporting events may have less impact on overall crime rates than holidays that alter routine activities citywide.

TABLE II: Total Arrests During Major Events vs. Baseline Periods

Event	**Arrests on Event Day**	**Baseline (Ø/day)**	**Difference**	**Change (%)**
Halloween (Oct 31)	215	310	−95	−30.6%
Thanksgiving (Nov 28)	127	330	−203	−61.6%
Independence Day (Jul 4)	211	287	−76	−26.6%
New Year's Eve (Dec 31)	159	279	−120	−43.0%
NYC Marathon (Nov 3)	196	228	−32	−14.2%
Average				−36.7%

Several factors may explain this counterintuitive finding. First, many residents leave the city during major holidays, reducing the pool of potential offenders and victims. Second, increased police deployment during high-profile events may deter criminal activity or shift enforcement priorities. Third, the social dynamics of holidays, particularly family gatherings and altered routines—may temporarily disrupt the environmental conditions that facilitate certain types of crime.

2) Crime Category Composition: While total arrests decreased, the composition of crime categories showed significant event-specific shifts. Table III presents the detailed breakdown of crime category distributions for baseline and event days across all five events.

Table IV summarizes the average changes across all events. The most striking pattern is the consistent and substantial increase in the proportion of Assault 3 arrests during events (average +13.27 percentage points), reaching an extreme of +34.2pp during Thanksgiving. This suggests that while overall arrests decline, interpersonal conflicts and altercations become relatively more prominent during holiday gatherings, potentially reflecting alcohol consumption, family tensions, or crowded conditions.

Conversely, Dangerous Drugs arrests show a consistent proportional decline (average −8.13pp), with particularly strong decreases during Thanksgiving (−14.3pp) and New Year's Eve (−14.2pp). This may reflect shifted police enforcement prior-

TABLE III: Crime Category Distribution: Baseline vs. Event Days (Percentage of Total Arrests)

Crime Type	**Halloween**		**Thanksgiving**		**Independence**		**New Year's**		**NYC Marathon**	
	Base	Event	Base	Event	Base	Event	Base	Event	Base	Event
Robbery	12.6%	16.3%	12.8%	15.7%	14.0%	10.4%	13.0%	9.4%	8.2%	9.2%
	(+3.7pp)		(+2.9pp)		(−3.6pp)		(−3.6pp)		(+1.0pp)	
Assault 3	34.5%	41.9%	32.7%	66.9%	35.0%	40.3%	33.6%	46.5%	42.9%	49.5%
	(+7.3pp)		**(+34.2pp)**		(+5.2pp)		(+12.9pp)		(+6.6pp)	
Drugs	21.0%	21.9%	19.8%	5.5%	22.6%	13.7%	23.0%	8.8%	13.3%	9.2%
	(+0.8pp)		(−14.3**pp)**		(−8.9pp)		(−14.2pp)		(−4.1pp)	
Petit Larceny	28.2%	17.2%	31.8%	11.0%	26.5%	31.8%	27.4%	30.2%	34.2%	30.1%
	(−11.0pp)		(−20.8**pp)**		(+5.3pp)		(+2.8pp)		(−4.1pp)	
Trespass	3.6%	2.8%	2.9%	0.8%	1.9%	3.8%	3.0%	5.0%	1.5%	2.0%
	(−0.8pp)		(−2.1**pp)**		(+1.9pp)		(+2.1pp)		(+0.6pp)	

Note: Values in parentheses show percentage point changes (pp). Bold values indicate the most extreme changes for each crime type.

TABLE IV: Average Crime Category Changes Across All Events

Crime Category	**Avg. Change (pp)**	**Direction**
Assault 3 & Related	+13.27	↑ Increase
Dangerous Drugs	−8.13	↓ Decrease
Petit Larceny	−5.55	↓ Decrease
Robbery	+0.09	≈ Stable
Criminal Trespass	+0.32	≈ Stable

ities during major events, reduced street-level drug activity as dealers and users alter routines, or simply that drug arrests represent a smaller share of the reduced total arrest volume.

Petit Larceny exhibits event-specific variation. Halloween and Thanksgiving show substantial decreases (−11.0pp and −20.8pp respectively), while Independence Day and New Year's Eve show modest increases. This variability likely reflects differences in crowd behavior and opportunity structures: holidays focused on private gatherings may reduce theft opportunities, while public celebrations with large crowds may create conditions favorable to pickpocketing and opportunistic theft.

D. Spatial Density Evolution

We visualized arrest distributions using kernel density estimation (KDE) to compare crime patterns across temporal periods around each major event. Figure 5 presents spatial heatmaps for three time periods: one week before (baseline), event day, and one week after (recovery).

Stable Geographic Patterns. Across all events, spatial distributions remain remarkably consistent. Central Manhattan (Midtown) and northern Brooklyn maintain elevated density levels regardless of time period. This stability indicates that major events do not fundamentally alter crime geography—the same areas remain hotspots throughout.

Event-Driven Volume Fluctuations. While spatial patterns stay stable, arrest counts vary significantly. Independence Day shows 304 arrests (baseline) → 211 (event day, -30.6%) → 272 (post-event, +28.9% from event). Halloween exhibits 293 → 215 (-26.6%) → 329 (+53.0%). NYC Marathon shows minimal change: 197 → 196 (-0.5%) → 227 (+15.8%). Thanksgiving demonstrates the largest reduction: 306 → 127 (-58.5%) → 258 (+103.1% from event day, but -15.7% below baseline).

Post-Event Displacement Effects. Independence Day and Halloween both show substantial post-event increases (+28.9% and +53.0%), suggesting temporal displacement rather than genuine suppression of criminal activity. This pattern indicates that enforcement strategies should maintain elevated staffing not only during events but also in the immediate aftermath.

The consistency of spatial patterns, combined with varying volume effects, suggests major events in NYC primarily affect the *timing and intensity* of arrests within established hotspot zones rather than their *geographic distribution*. This has practical implications: law enforcement should adjust staffing levels in existing hotspots based on predicted volume changes rather than repositioning patrols to new areas during events.

E. Synthesis: Persistent vs. Event-Driven Crime Patterns

Combining the baseline DBSCAN, multi-event comparative analysis, and kernel density estimation provides a comprehensive view of how crime patterns operate across different temporal scales in New York City:

Persistent Spatial Concentration. The baseline analysis confirms that crime exhibits strong, stable spatial clustering throughout 2024, with Manhattan's commercial and transit districts serving as consistent hotspots. These persistent patterns likely reflect underlying environmental criminology factors such as routine activities, land use, and mobility infrastructure [9].

Event-Driven Volume Reduction. Contrary to expectations, major public events correlate with substantial decreases in total arrests (average −36.7%), suggesting that holiday routines and event-related behavioral changes temporarily suppress overall criminal activity. This "holiday effect" appears strongest for family-oriented holidays (Thanksgiving, New Year's Eve) and weakest for crowd-intensive sporting events (NYC Marathon).

Crime Category Redistribution. While total volumes decrease, the composition of crime categories shifts significantly during events. Assault arrests become proportionally more common (average +13.27pp), while drug arrests decline (−8.13pp). This redistribution suggests that events alter the opportunity structures and social dynamics that facilitate different crime types.

Spatial Redistribution During Events. Kernel density analysis reveals three distinct spatial response patterns: uniform density reduction across all areas (Thanksgiving), localized shifts maintaining density in central Manhattan (Independence Day, Halloween), and mobile corridor extensions along event routes (NYC Marathon). These patterns demonstrate how different event types reshape crime geography temporarily, with rapid return to baseline within one week post-event.

Methodological Implications. The complementary use of DBSCAN (persistent baseline), comparative statistics (overall event effects), and kernel density estimation (spatial event dynamics) demonstrates the necessity of multi-scale analysis. No single method captures the full complexity of urban crime dynamics; each provides essential pieces of the broader picture.

VII. Conclusions and Future Work

A. Failed Approaches

While ST-DBSCAN appeared theoretically well-suited for analyzing spatio-temporal crime patterns, it proved impractical for our dataset due to two critical limitations:

Data Sparsity. Single-day event subsets (e.g., NYC Marathon, Halloween) contained only 200 arrests distributed across New York City's approximately 780 km² area. This resulted in extremely low spatial point density when analyzed within short temporal windows (e.g., hourly or 4-hour intervals). The sparse distribution made it impossible to establish meaningful density thresholds that could identify genuine spatio-temporal clusters rather than arbitrary groupings.

Temporal Parameter Sensitivity. The temporal radius parameter created counterintuitive clustering behavior. When set to capture meaningful temporal patterns (e.g., incidents within the same 2-4 hour window), the algorithm grouped geographically distant incidents solely based on temporal proximity. For example, a robbery in Manhattan at 2 PM and an assault in Queens at 3 PM would be clustered together despite being 15 km apart, obscuring true spatial hotspots. Conversely, stricter temporal data excessively, yielding mostly noise points rather than meaningful clusters.

As a result, we relied on traditional spatial DBSCAN for year-round baseline hotspot analysis and kernel density estimation (KDE) for temporal event comparisons (Section VI.C-D). These methods proved more interpretable and robust for our sparse, event-driven dataset while still revealing meaningful spatial and temporal patterns.

B. Limitations

While our analysis provides insights into event-driven crime dynamics in New York City, several methodological and data-related limitations should be acknowledged.

Temporal Scope. Our study focuses exclusively on 2024 arrest data, limiting our ability to identify multi-year trends or assess whether event-specific patterns are consistent across different years. A longitudinal analysis spanning multiple years would strengthen conclusions about the reproducibility of observed effects.

Arrest vs. Crime Incidence. The NYPD arrest dataset captures enforcement actions rather than all criminal incidents. Arrests depend on police presence, reporting behavior, and enforcement priorities, which may shift during major events. Consequently, reduced arrest counts during holidays may partially reflect changes in policing strategy rather than actual crime reduction.

Event Selection Bias. We analyzed only five major recurring events with clear temporal boundaries. Smaller events, spontaneous gatherings, or overlapping activities were not considered, potentially underestimating the full spectrum of event-driven crime fluctuations. Additionally, unpermitted events (protests, flash gatherings) are not captured in our event dataset.

Spatial Resolution. Geographic coordinates in arrest records represent arrest locations, which may differ from actual crime scene locations due to suspect apprehension elsewhere. This introduces spatial uncertainty, particularly for mobile offenses like robbery.

Confounding Variables. Our analysis does not control for weather conditions, police deployment patterns, or concurrent events that may independently influence crime rates. For example, reduced arrests during Thanksgiving likely reflect multiple factors beyond the holiday itself, including population outflow and altered routine activities.

Sampling Trade-offs. To manage computational constraints, we employed stratified spatial sampling, reducing the dataset from $\sim$100,000 to $\sim$30,000 arrests. While this preserved spatial and categorical distributions, it may have reduced statistical power for detecting localized micro-patterns or rare crime types.

ST-DBSCAN Applicability. As discussed in Section VII-A, spatio-temporal clustering proved unsuitable for single-day event analysis due to data sparsity ($\sim$200 arrests/day across 780 km^2). Future work with higher-frequency data (e.g., complaint records including non-arrest incidents) may enable more granular spatio-temporal modeling.

C. Conclusion

Our analysis reveals patterns contrary to our initial hypothesis: arrests consistently decreased during major public events (average -36.7%), with family-oriented holidays showing the largest reductions (Thanksgiving: -61.6%). However, crime composition shifted significantly—Assault 3 arrests increased proportionally (+13.27pp average) while Dangerous Drugs arrests declined (-8.13pp).

These patterns suggest that major events alter both opportunity structures and enforcement priorities. The assault increase may reflect crowding effects, alcohol consumption, or interpersonal tensions during gatherings, while drug arrest reductions likely indicate shifted police focus toward crowd management and public safety rather than proactive drug enforcement.

D. Practical Recommendations

Based on our findings, we offer the following recommendations for **For Tourists and Visitors:** Major public events in NYC (Thanksgiving, New Year's Eve, Halloween) correlate with significantly reduced arrest volumes (-36.7% average), suggesting these periods may be relatively safer for general mobility. However, the proportional increase in assault arrests (+13.27pp) warrants heightened awareness in crowded venues and event areas. Visitors should exercise particular caution during evening hours at large gatherings.

E. Future Work

Temporal Extension: Analyze 3-5 years of data to validate pattern consistency and detect year-over-year trends in event-driven crime dynamics.

Confounding Variable Control: Apply regression models controlling for weather, police deployment density, population mobility (e.g., cell phone data), and overlapping events to isolate pure event effects.

Comparative City Analysis: Extend to cities with different urban structures (e.g., Los Angeles, Chicago, Miami) to test generalizability of the "holiday effect" and assault concentration patterns.

Sociodemographic Integration: Link crime patterns to Census tract data to examine whether event impacts vary by neighborhood income, residential density, or tourism intensity.

VIII. Appendix

References

[1] A. C. Uittenbogaard and V. Ceccato, "Space–Time Clusters of Crime in Stockholm, Sweden," *Review of European Studies*, vol. 4, no. 1, pp. 148–161, 2012.

[2] M. Delgado and A. Sánchez-Delgado, "Seasonality in Property Crime: Entropy Measures and Predictability in Barcelona," *Applied Sciences*, vol. 9, no. 24, 2019, Art. no. 5486.

[3] J. S. Kumar, M. Amiruzzaman, A. A. Bhuiyan, and D. Bhati, "Predictive Analytics in Law Enforcement: Unveiling Patterns in NYPD Crime through Machine Learning and Data Mining," *Research Briefs on Information & Communication Technology Evolution (ReBICTE)*, vol. 10, Article No. 03, May 2024. DOI: 10.56801/rebicte.v10i.188. Available: https://rebicte.org/index.php/rebicte/article/view/188/218

[4] N. Nalluri, M. Bsales, and C. Nelson, "Data-Driven Security Measurements to Improve Safety in NYC and NJ Mass Transit," *Journal of Systemics, Cybernetics and Informatics*, vol. 21, no. 3, pp. 47–55, 2023. Available: https://www.iiisci.org/journal/PDV/sci/pdfs/ZA129PB23.pdf

[5] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, "Randomized Controlled Field Trials of Predictive Policing," *Journal of the American Statistical Association*, vol. 110, no. 512, pp. 1399–1411, 2015.

[6] X. Wang, R. Gerber, and J. Brown, "Spatio-Temporal Modeling of Crime Using Deep Learning," *Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 2115–2123, 2019.

[7] M. A. Andresen and N. Malleson, "Crime Seasonality and its Variations Across Space," *Applied Geography*, vol. 59, pp. 25–35, 2015.

[8] V. Ceccato and A. C. Uittenbogaard, "Space-Time Dynamics of Crime in Transport Nodes," *Annals of the Association of American Geographers*, vol. 104, no. 1, pp. 131–150, 2014.

[9] P. J. Brantingham and P. L. Brantingham, "Criminality of Place: Crime Generators and Crime Attractors," *European Journal on Criminal Policy and Research*, vol. 3, pp. 5–26, 1995.

[10] Y. Yang, L. Pan, and W. Song, "Tourism, Crime, and Safety Perceptions: Evidence from Airbnb Listings in New York City," *Tourism Management*, vol. 77, Art. no. 104014, 2020.

[11] Derya Birant and Alp Kut. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering*, 60(1):208–221, 2007. ISSN 0169-023X. doi:10.1016/j.datak.2006.01.013. URL: https://www.sciencedirect.com/science/article/pii/S0169023X06000218. (Intelligent Data Mining).

[12] B. Gebreamlak and D. Ochoa, "Heat Mapping Crime: A Data-Driven Approach to Policing in New York," Portland State University, altREU Projects, 2023. Available: https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1016&context=altreu_projects.

[13] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 226–231, AAAI Press, 1996.

[14] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-Based Clustering Based on Hierarchical Density Estimates," in *Advances in Knowledge Discovery and Data Mining (PAKDD 2013)*, Lecture Notes in Computer Science, vol. 7819, pp. 160–172, Springer, 2013. DOI: 10.1007/978-3-642-37456-2_14.

[15] Data Mining and Text Mining GitHub Repository. Available: https://github.com/Jatixx/Data_Mining_and_Text_Mining.git

[16] NYC Open Data, "NYPD Arrests Data (Historic)," 2024. Available: https://data.cityofnewyork.us/Public-Safety/NYPD-Arrests-Data-Historic-/8h9b-rp9u/about_data. [Accessed: Oct. 25, 2025].

[17] NYC Open Data, "NYC Permitted Event Information – Historical," 2025. Available: https://data.cityofnewyork.us/City-Government/NYC-Permitted-Event-Information-Historical/bkfu-528j/about_data. [Accessed: Oct. 25, 2025].

[18] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. New York, NY: Routledge, 2018.

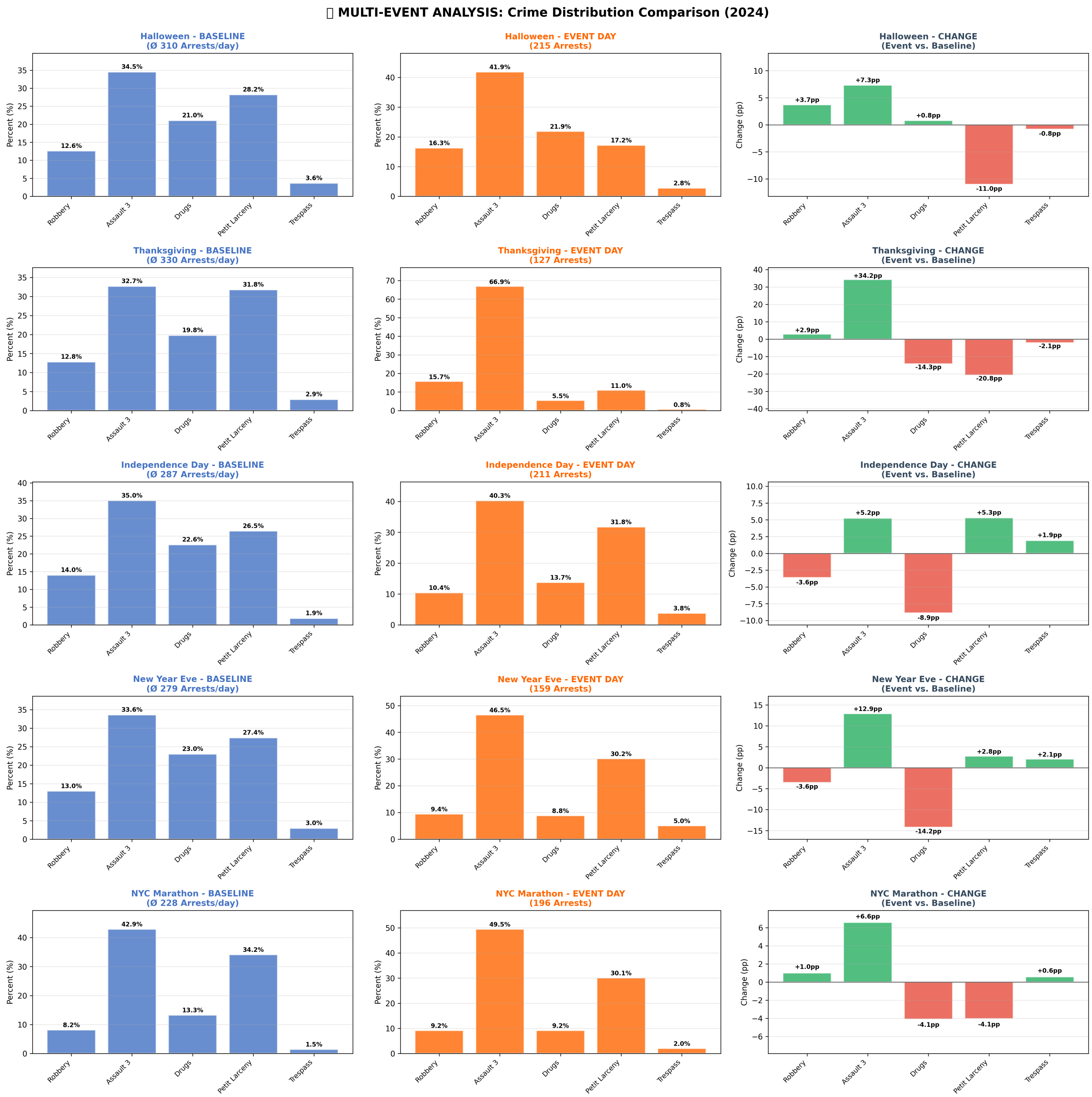

Fig. 4: Multi-Event Crime Distribution Comparison across five major NYC events in 2024. For each event (rows), three views are shown: baseline distribution from control days (left), event day distribution with total arrest counts (middle), and percentage point changes between baseline and event day (right, green = increase, red = decrease). All events show reduced total arrests (ranging from -14.2% to -61.6%) but significant compositional shifts, with Assault 3 consistently increasing and Dangerous Drugs declining proportionally.

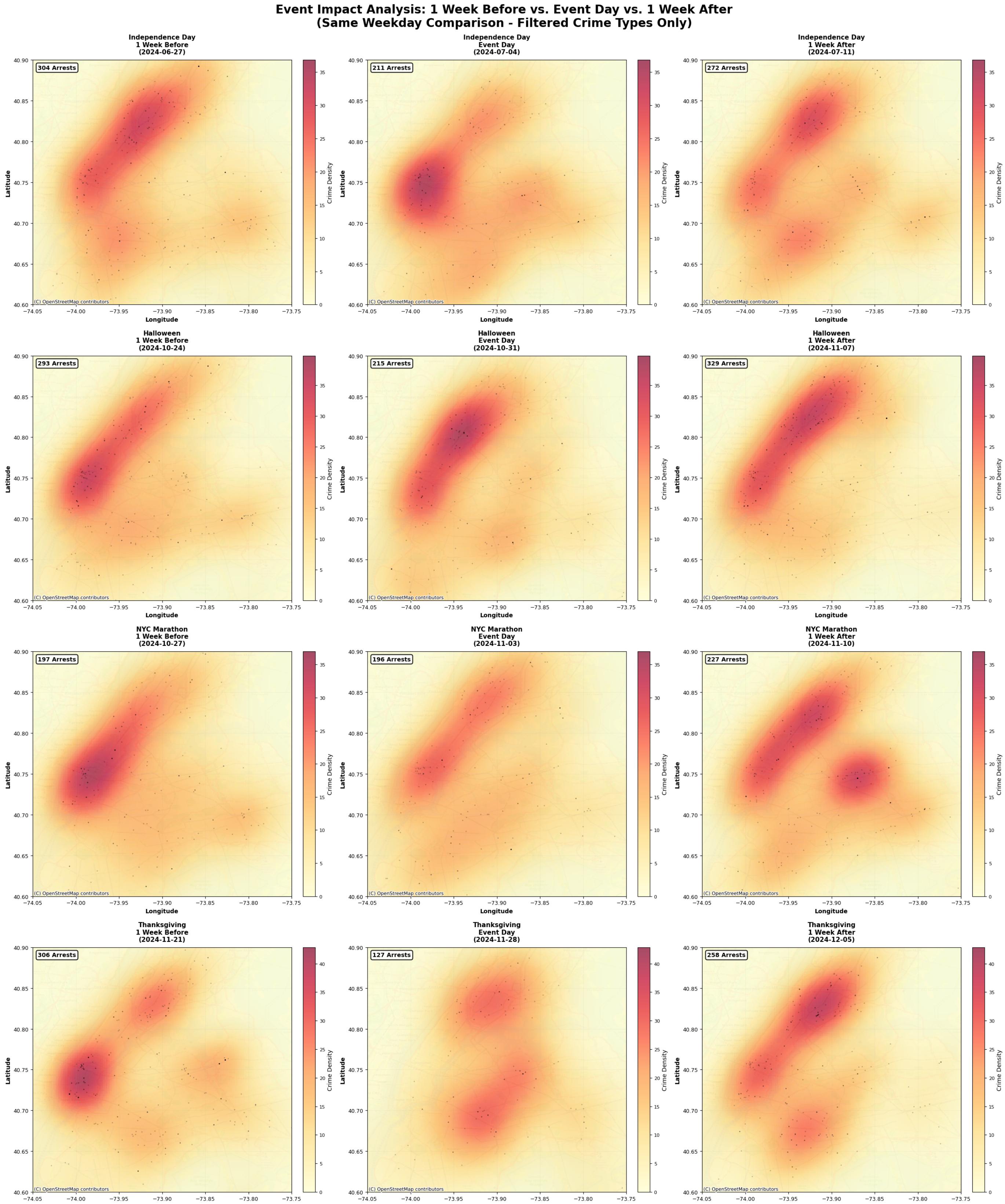

Fig. 5: Spatial density heatmaps comparing arrest patterns one week before (left), during (center), and one week after (right) five major NYC events in 2024. Each row represents one event with unified color scaling for direct comparison. Higher crime density is shown in red, lower density in yellow, with individual arrests marked as black dots.