Predictive Modeling of Intimate Partner Violence Using NYC Police Data

Ash Todd, Eaint May, Nora Phelan, Zack Taylor

May 14, 2025

Abstract

Intimate partner violence is a complex issue that is difficult to predict due to the many factors that may influence individuals to commit the violence. In this study we develop a predictive model using publicly available NYPD domestic violence reports to identify precincts at elevated risk of IPV. We take a place-based approach which focuses on policing by area and therefore mitigates broader biases involved with race, gender, and other personal characteristics. Our analysis shows that even a high-accuracy model still mirrors and magnifies existing disparities as racial and economic factors are disproportionate in areas of high police activity. Our findings present a fundamental tension in predictive policing: algorithmic performance metrics can mask deeper injustices perpetrated in historical policing data. Without confronting systemic issues such as racial bias and socioeconomic disparity, any deployment of IPV predictive policing risks perpetuating the very inequities it seeks to prevent.

1 Key Words

Predictive Policing According to Brennan Center for Justice, "Predictive policing involves using algorithms to analyze massive amounts of information in order to predict and help prevent potential future crimes." (Diaz, 2019)

DIR Short for Domestic Incident Report, the report required to be filled out by the NYPD when investigating a domestic issue.

IPV Short for Intimate Partner Violence, defined by the WHO as "behaviour within an intimate relationship that causes physical, sexual or psychological harm, including acts of physical aggression, sexual coercion, psychological abuse and controlling behaviours."

Intersectionality The Merriam-Webster Dictionary defines intersectionality as "the complex, cumulative way in which the effects of multiple forms of discrimination (such as racism, sexism, and classism) combine, overlap, or intersect especially in the experiences of marginalized individuals or groups".

Borough A subdivision of the New York City government. "New York is composed of five boroughs – Brooklyn, the Bronx, Manhattan, Queens and Staten Island" (NYC.gov).

Community District The New York City Charter defines a community district as an area of the city which is "used for the planning of community life within the city, the participation of citizens in city government within their communities, and the efficient and effective organization of agencies that deliver municipal services in local communities and boroughs." There are 59 in total, each residing within a single borough. (nyccharter.readthedocs.io)

Police Precinct An area of New York City with a dedicated police department. Each precinct is roughly equivalent in population to "the size of a midsize American city" (NYC.gov).

2 Introduction

Domestic violence is a deeply complex issue, rooted in histories of inequality, shaped by intersecting forces of gender, economic status, and systemic neglect. Domestic violence is officially defined as "physical, sexual, emotional, economic, psychological, or technological actions or threats of actions or other patterns of coercive behavior that influence another person within an intimate partner relationship." (OVW). The variety of social and psychological factors which can push someone to commit violence within the home make it difficult to predict future offenses. Additionally, responses to domestic violence are often inadequate and lacking in resources required to protect the victim from recurring violence. These failures incentivise other victims against reporting violence, and therefore makes fixing the broader issues even tougher. Unfortunately, as many of these cases go unreported, our research is constrained to reported cases of domestic violence.

This project aims to decipher the ways in which police data may be used to allocate resources to combat intimate partner violence through a predictive model lens. Through the analysis of the New York City Police Department's official reports of domestic violence, we were able to create a model of a place-based predictive algorithm pertaining to IPV reports that finds precincts with a "high report count" and flags precincts with "high felonies", which occurs when a precinct falls into the top-10 precincts by felony-level incident count over the target window. To achieve this we began by detailing the steps that a city such as New York might take to develop a predictive model for IPV: what data would be collected, how it would be processed, and what features might be prioritized. Using publicly available police report data, we explore the patterns and correlations that might inform such a model. Our primary research question is: how accurately can a place-based predictive algorithm–trained on NYPD domestic violence reports–identify the precincts that will exceed the "high report count" and/or "high felonies" threshold the following week? Additionally, we will analyze which neighborhood factors (racial composition and median income) drive these predictions. We expect to find that a model trained on this data would disproportionately associate domestic violence with predominantly Black and economically disadvantaged neighborhoods. We will also interrogate the assumptions behind this data and ask what gets lost, distorted, or reinforced when IPV is approached through a predictive lens. This is an especially important step when using police report data which, as studies have shown, is likely to be influenced by many biases (Güss et al., 2020). Furthermore, what does it mean to make violence "predictable" using data that itself is shaped by uneven policing and systemic bias? And who bears the risk when predictive tools are used in already-vulnerable communities?

3 Data

The dataset used in our analysis was released to the public by the New York City Police Department. It contains data recorded from domestic violence-related offenses in New York City from 2020 and 2021. Each report includes information on the type of offense, the date it was reported, the precinct code and borough in which it occurred, if the offense involved an intimate partner, the race, sex, and age of the victim and offender, and information on the financial state of their area. The type of offense is listed as either a Domestic Incident Report (DIR), felony rape, or felony assault. According to the New York State DCJS's model policy for domestic incidents, a DIR is a state-issued form "for officers to complete for all calls for service where police intervention is requested for a domestic incident." They are taken at the time of the response and contain no information on further conviction or other legal interventions involving either party (suspect or victim).

Before cleaning, the dataset contained missing values in multiple columns. The columns indicating high poverty, low median income, and high unemployment either contained a 1 or were left blank. These NaNs were changed to 0. Additionally, the columns containing the reported ages of the suspect and the victim had a significant number of blank entries. This could be due to insufficient information at the scene of the incident or mistakes in the reporting process. Regardless, we felt this missing information could cause inaccuracies in our analysis when it came to studying overall themes in age difference against other factors. For this reason, we took the average age in each column and replaced the missing values with those averages. This way we could take the ages across the entire dataset against other factors and still produce reasonably accurate results. Any remaining rows containing

NaNs in other columns, as well as one incorrect age (likely a human error), were removed from the dataset.

The datatypes of multiple columns also needed to be changed to be used in our analysis. The suspect and victim age columns were changed from strings to integers, with 0 indicating female and 1 indicating male. Suspect and victim ages were also changed to integers, and the report date column was changed to DateTime.

For our analysis, we needed a way to determine if the reported offense was a felony or not. The existing offense type column has three possible values: DIR, Rape, or Felony Assault. The latter two constitute felonies, but a DIR can result in a misdemeanor, a felony, or no arrest. To simplify this data, we created a new variable called Felony Offense. This variable attaches a boolean value to the report: 1 for a reported felony (either rape or felony assault) or 0 for no reported felony (DIR). This way we can easily make connections between a reported crime and other variables within the data.

4 Methods

According to a report on New York City Surveillance Technology by the Brenna Center, there are two types of predictive policing in place: place-based and person-based. Place-based predictive policing uses algorithmic systems to analyze datasets to try to predict where certain crimes are likely to occur. Police presence is then deployed based on the predictions. Person-based predictive policing uses algorithmic systems to analyze datasets to generate a list of individuals who are likely to commit a crime. (Diaz, 2019).

A model of place-based predictive policing was created by finding the ten precincts with the highest number of DIRs for a given month as shown by Table 1 through 6. A precinct appearing on this list may indicate a place with a high number of domestic incidents. The ten precincts with the highest number of felonies for the given month were also tracked. These two lists of precinct codes were compared and an additional flag was placed next to high-reporting precincts that also contained a high number of felony reports. This information could be used by the New York City Police Department to predict "hotspot" locations. The department's response would then be to allocate additional resources to these locations and increase patrol units in the area. (Developing the NYPD's Information Technology).

To move from descriptive flags to a forecasting tool, we built a binary classifier using logistic regression. For each precinct and the first week of each month in 2020, we computed two features: (1) the count of DIRs and (2) the count of felony-level incidents. We then labeled each precinct "1" if it appeared among the top 10 by DIRs in the following week and "0" otherwise. We repeated this process separately for felony-only counts to create a second label. After aggregating to one record per precinct-week, we split the data into 80% training and 20% test sets, fit two logistic models (one for overall DIR flags, one for felony flags), and evaluated their accuracy. This approach allowed us to quantify which precincts are likely to become DIR hotspots.

Another area of our analysis was the racial makeup of the suspects and victims from each precinct. We created a chart to show the racial proportions of the suspects from each precinct (Fig. 4), and then a similar chart showing the racial proportions of the victims (Fig. 5). Displaying this information alongside the volume of reports in each precinct can help determine a connection between predictive policing and racial bias. It could be used in these "hotspot" precincts to determine if there is a racial bias in the policing of these areas or, when compared to the area's actual population, if the racial makeup of those precincts is driving that assumption.

Finally, the reports in the dataset were examined on the basis of socioeconomic standing. We calculated the percentage of community districts in each borough that are economically disadvantaged, including the percentage of community districts experiencing high poverty, high unemployment, and low median income (Table 8). We also calculated the percentage of the total number of reports in each borough that occurred in an economically disadvantaged community district, again based on poverty, unemployment, and median income (Table 7). This information allows us to draw connections between the rate of policing and economic inequality.

5 Results

In the created model as shown in Table 1 through 6, a total of 6 months were analyzed–February through April in both 2020 and 2021–and 8 precincts were consistently present in the resulting reports: precincts 75, 43, 47, 40, 46, 73, 52, 67. Precincts 75, 73, and 67 are located in Brooklyn and 43, 47, 40, 46, and 52 are located in the Bronx. All of these precincts were also found to have a high felony rate for at least one out of the three months analyzed for each year.

From the model we built using logistic regression, we held out 20% of the precinct-week observations and then computed a confusion matrix (Fig. 3) to assess performance. The matrix showed:
True Positives: 72 precincts correctly flagged as next week's hotspots
True Negatives: 373 precincts correctly flagged as next week's non-hotspots
False Positives: 39 precincts incorrectly flagged as next-week
False Negatives: 29 precincts the model failed to flag despite being actual hotspots. From these results, we can report an overall accuracy of 86.7%, an overall precision of 64.9%, and an overall recall of 71.3%.

We also tracked the accuracy, recall, and precision of the model by month (Fig. 4-6). Accuracy fluctuates between 0.814 and 0.911, peaking in March and dipping lowest in September. Precision ranges from a low of 0.50 in April to a high of 0.75 in March and December indicating variable false-alarm rates across months. Recall spans 0.60 to 0.8333, with the model best at catching true hotspots in November and weakest in September. Overall, the model exhibits reasonably high and fairly stable accuracy throughout the year.

Based on our bar graphs detailing the racial makeup of the reports from each precinct (Fig. 7 and 8), we did not observe any significant difference between victims and suspects. Notably, the graphed reports showed most of the precincts as having a black majority. It should also be noted that a few precincts had a strong white majority. These majorities remained true across both victims and suspects for each precinct. Additionally, we observed one precinct with a majority–both victims and suspects–reported as Asian/Pacific Islander: precinct 106.

Our socioeconomic data shows that the Bronx, Brooklyn, and Manhattan had the highest percentages of community districts with a high poverty rate and a low median income; while the Bronx, Manhattan, and Queens had the highest percentages with high unemployment (Table 8). These rankings were similar to the boroughs with the highest percent of reports occurring in economically disadvantaged community districts (Table 7); though their order did vary, and the percentage of reports was generally higher than the percentage of districts.

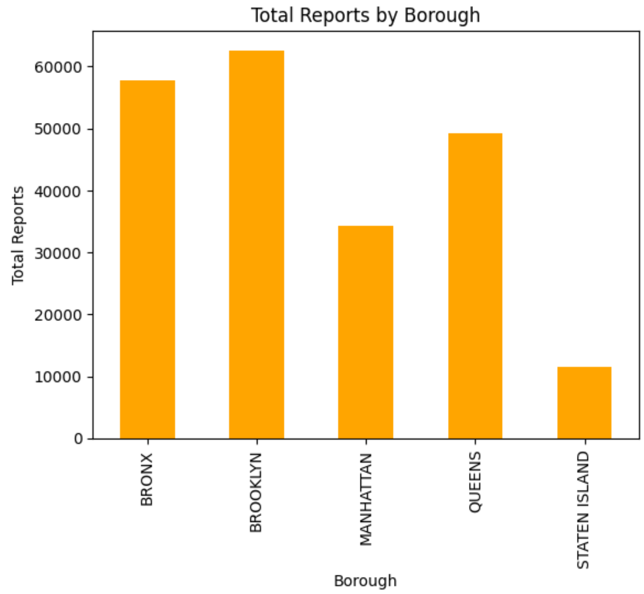

Figure 1: Total Reports Distribution by Borough.

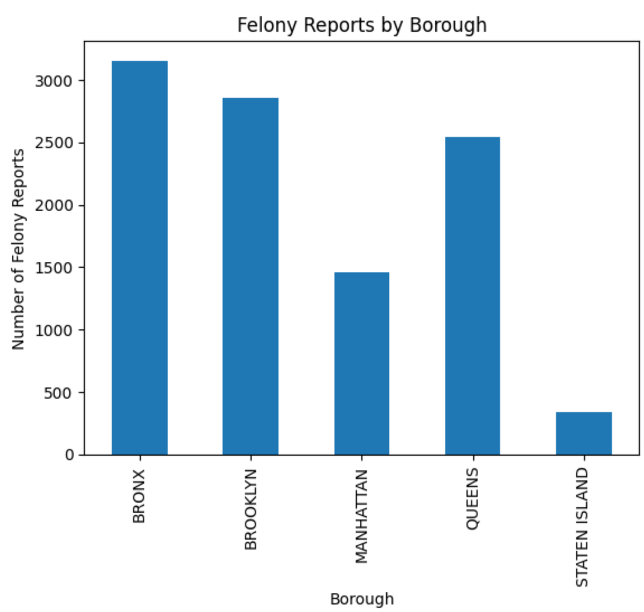

Figure 2: Total Felony Reports Distribution by Borough.

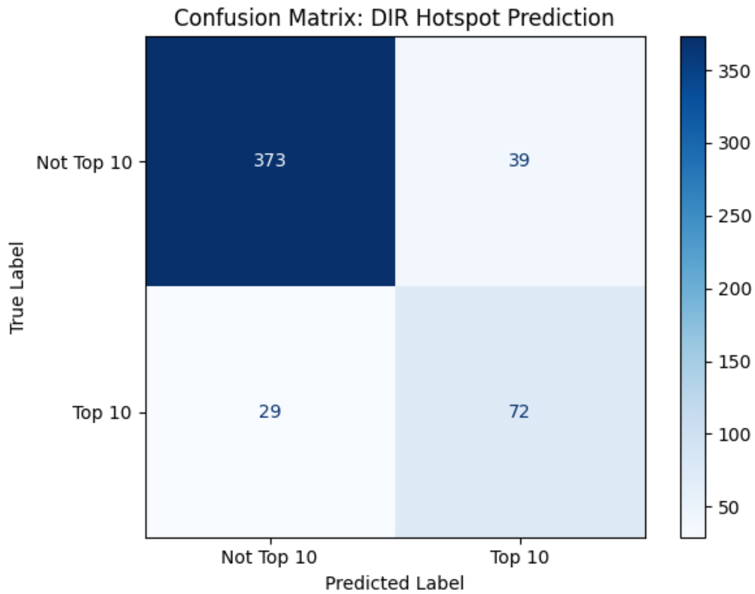

Figure 3: ConfusionMatrix of DIR Hotspot Prediction, Prediction by True Outcome.

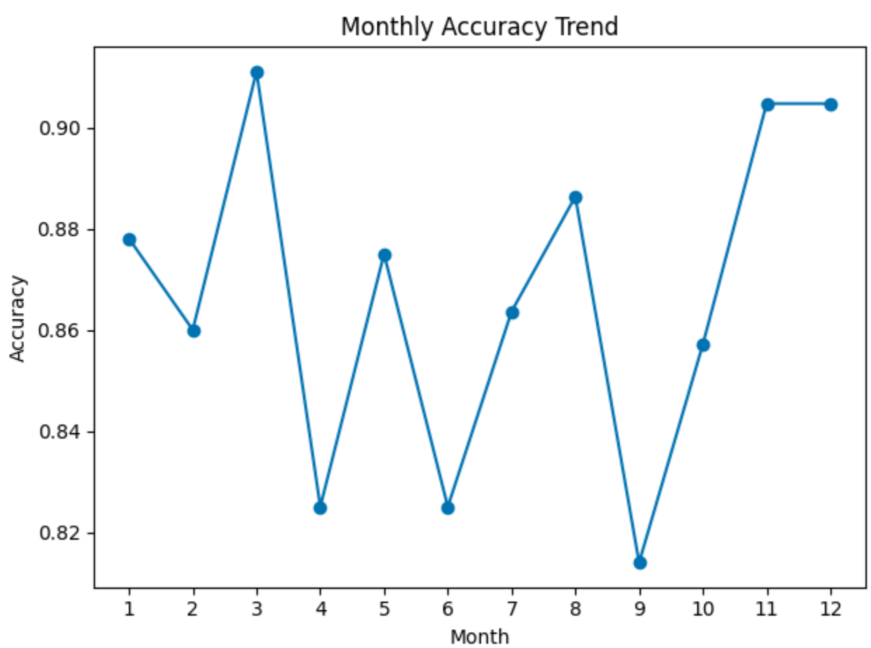

Figure 4: Accuracy Trend of Logistic Regression Model across each Month.

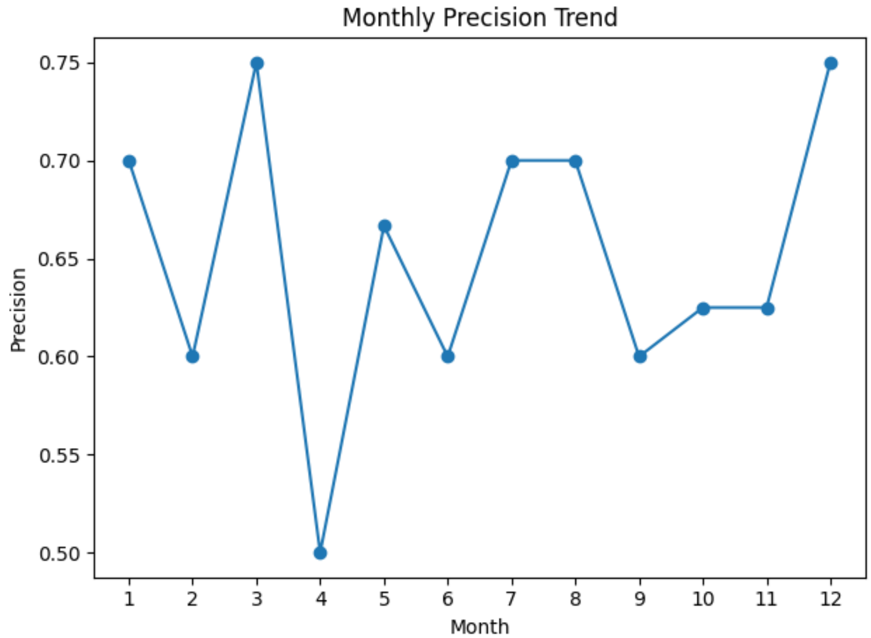

Figure 5: Precision Trend of Logistic Regression Model across each Month.

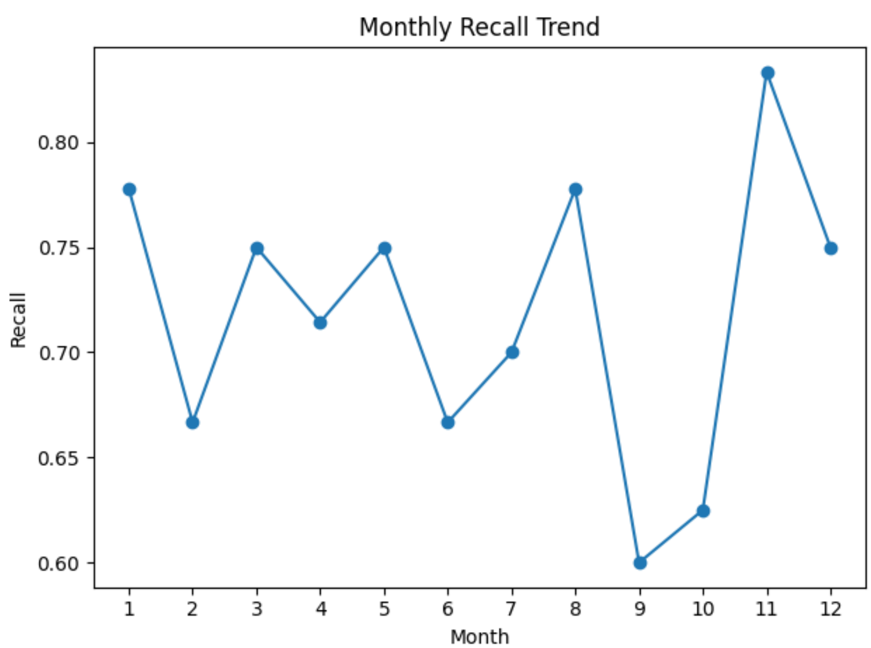

Figure 6: Recall Trend of Logistic Regression Model across each Month.

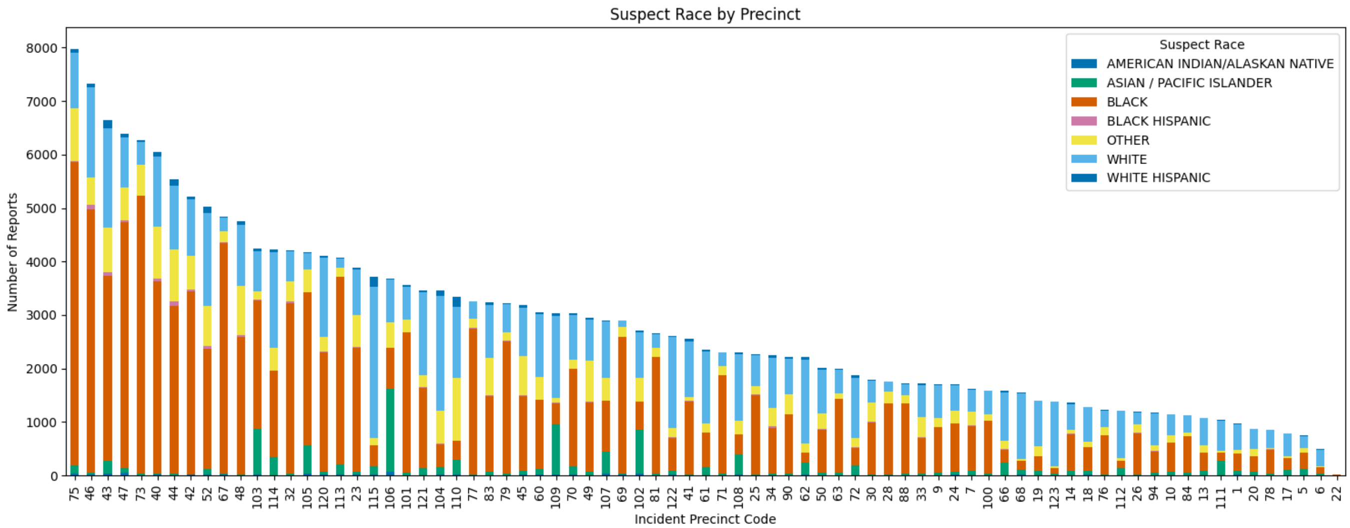

Figure 7: Stacked Bar Chart of amount of Reported Suspect's Race Organized by Precinct.

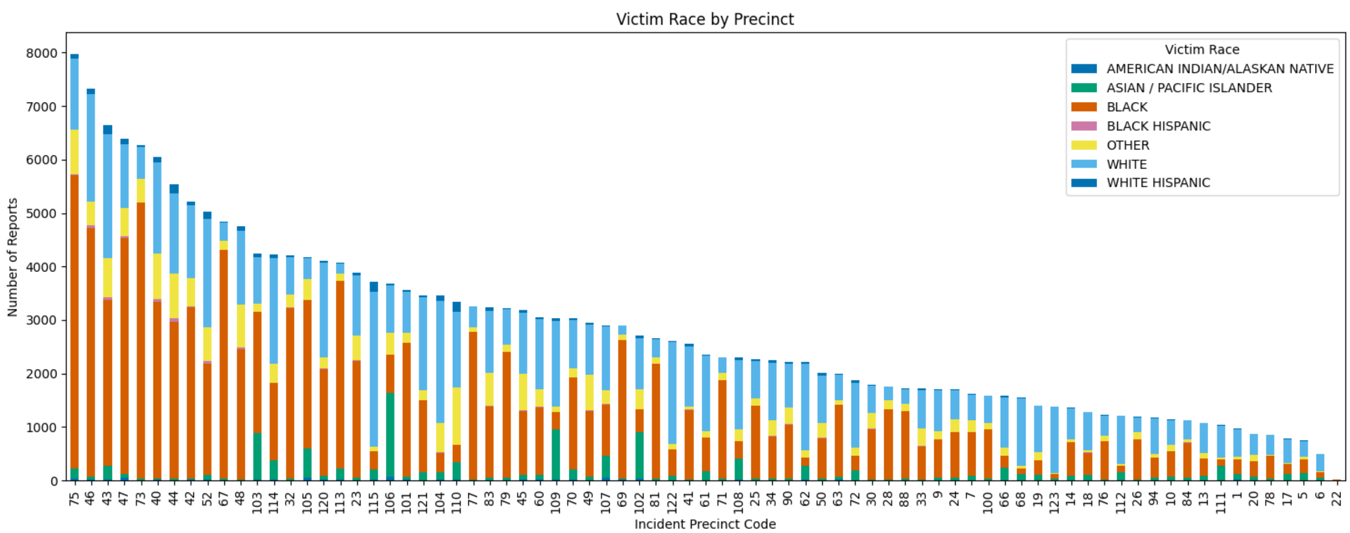

Figure 8: Stacked Bar Chart of amount of Reported Victims's Race Organized by Precinct.

Precinct	**Reports**	**High Felonies**
46	293	True
75	252	True
40	246	True
43	237	True
47	235	True
73	225	False
44	203	True
48	186	False
52	185	True
67	183	True

Table 1: Top precincts by report count and high felony classification (February 2020)

Precinct	**Reports**	**High Felonies**
46	324	True
75	285	False
73	283	False
43	259	True
47	238	True
40	227	True
42	225	True
44	203	True
52	201	True
67	190	False

Table 2: Top precincts by report count and high felony classification (March 2020)

Precinct	**Reports**	**High Felonies**
75	261	False
46	231	True
73	215	True
43	213	True
40	206	True
42	205	False
44	201	True
48	185	False
47	182	False
52	158	False

Table 3: Top precincts by report count and high felony classification (April 2020)

Precinct	**Reports**	**High Felonies**
75	272	True
46	271	True
43	218	False
73	210	False
67	210	True
47	207	True
44	199	True
42	189	True
40	181	False
52	178	False

Table 4: Top precincts by report count and high felony classification (February 2021)

Precinct	**Reports**	**High Felonies**
75	299	False
46	294	True
73	283	True
43	240	True
47	234	True
44	229	True
42	228	True
40	218	True
67	189	False
52	187	False

Table 5: Top precincts by report count and high felony classification (March 2021)

Precinct	**Reports**	**High Felonies**
46	333	False
75	297	True
73	264	True
43	249	True
40	237	True
67	196	False
47	196	True
42	192	False
52	190	True
105	188	False

Table 6: Top precincts by report count and high felony classification (April 2021)

	Poverty	Unemployment	Median Income
Bronx	0.747846	0.909770	0.747846
Manhattan	0.466064	0.470966	0.466064
Brooklyn	0.397933	0.099839	0.446663
Queens	0.000629	0.167266	0.000629
Staten Island	0.000000	0.000000	0.000000

Table 7: Percent of Reports Occurring in Economically Disadvantaged Community Districts

	Poverty	Unemployment	Median Income
Bronx	0.533333	0.666667	0.533333
Brooklyn	0.250000	0.050000	0.300000
Manhattan	0.230769	0.307692	0.230769
Queens	0.000000	0.055556	0.000000
Staten Island	0.000000	0.000000	0.000000

Table 8: Percentage of Community Districts at an Economic Disadvantage

6 Discussion

The consistency of high report numbers across specific precincts points to two boroughs of New York City: Brooklyn and The Bronx. This suggests that a predictive policing system based on similar data would disproportionately target these areas, reinforcing or creating a reputation for criminality. This, in turn, could lead to heightened suspicion of the residents. Notably, these boroughs also tend to experience higher levels of poverty and unemployment.

Additionally, the higher percentage of reports from economically disadvantaged community districts (Table 7), as compared to the percentage of districts at a disadvantage (Table 8), could be an indicator of systemic bias. If each community district had a relatively similar level of reports, the percentage of reports from disadvantaged districts would be the same as the percentage of districts at a disadvantage. The fact that the percentage of reports is higher indicates a higher police presence in poorer districts.

Our analysis of model performance over time shows that overall accuracy remains fairly high. However, even an 86% accurate model can be fundamentally unfair if its training data reflect long standing biases. Because law enforcement practices historically over-patrol and over-report in certain neighborhoods, particularly those with larger Black and low-income populations, the model effectively learns these disparities as if they were grounded truth. In practice, this means high numeric accuracy masks a deeper problem: the model will reproduce and even amplify biased policing patterns, sending more resources to the same over-policed communities and generating more data that reinforce its flawed assumptions.

By analyzing the data we discovered systemic issues related to the racial and socioeconomic makeup of New York City which could cause bias in predictive policing. This presents ethical issues with any attempt to provide a system of predictive policing as it may inflate these systemic problems. Attempts to increase police presence in an area based on a model trained on systemically unfair data will lead to more reports in that area, and therefore create more unfair data to inform said model.

Understanding how intersectionality affects domestic violence is critical in order to adequately address the issue in an effective manner. Factors such as race, gender, class, and welfare status shape lived experiences, struggles, and access to support systems, and therefore useful methods of intervention (Josephson, 2002). By understanding the impact of intersectionality, we can further identify the disparities in reporting, response, and prevention strategies and develop domestic violence interventions that address the needs of all affected communities.

Using entirely government-collected data presents limitations to our analysis. The high representation of Black individuals in both victim and suspect roles mirrors historic patterns of over-policing in the United States justice system. For example, according to a Harvard study on racialized mass incarceration, in 2007 Black males made up only 12 percent of the US male population, and yet they made up 39 percent of the total incarcerated male population (Bobo et al., 2010). A model trained on data that reflects these disparities may overfit to racially coded patterns of surveillance, reinforcing the notion that domestic violence is more likely to occur in Black communities- not because of actual prevalence, but because of systemic inequities in reporting and enforcement. This must be taken into consideration when presenting any findings from analysis of data such as this.

There are also limiting biases related to data collected specifically by law enforcement. The environment in which police officers work can generate bias due to extreme stress and administrative scrutiny, among other things. A study done by the Department of Psychology at the University of North Florida on the problems with police reports as data sources explains that findings in reports related to intimate partner violence "showed a discordance, with police reports widely overestimating the injury severity. ... The point we emphasize is that certain information in police reports—in this case the extent of an injury—should be interpreted with caution by researchers." (Güss et al., 2020). This makes it difficult to take the results of our predictive model as entirely accurate since it has been trained on data that may have been adversely affected by police bias.

The absence of population data, specifically the populations of different community districts, populations within the jurisdiction of precincts, and the racial makeups of these populations, limits the conclusions we can draw from our data. Differences in the number of arrests in poorer districts could be due to over-policing, but they could also be the result of differences in population. Incorporating this population data could help to rule out concerns of inaccurate findings based on differences in the distribution of races in the areas being studied.

References

[1] New York State, Capital View Office Park. Domestic incident policy: Model law enforcement policy language. Technical report, New York State Office of Justice Programs, 2023. Accessed April 12, 2025.

[2] Angel Díaz. New york city police department surveillance technology. https://www.brennancenter.org/our-work/research-reports/new-york-city-police-department-surveillance-technology, 2019. Brennan Center for Justice, Accessed April 12, 2025.

[3] NYC Government. Developing the nypd's information technology. Technical report, 2025. Accessed April 12, 2025.

[4] New York City Police Department. Nypd patrol services bureau. https://www.nyc.gov/site/nypd/bureaus/patrol/patrol-landing.page, 2025. Accessed April 12, 2025.

[5] C. Dominik Güss, Ma. Teresa Tuason, and João Teixeira. Problems with police reports as data sources: A researchers' perspective. *Frontiers in Psychology*, 11:582428, 2020.

[6] Lawrence Bobo and Victor Thompson. Racialized mass incarceration: Poverty, prejudice, and punishment, 2016. Accessed April 12, 2025.

[1] [2] [3] [4] [5] [6]