

New York City Shooting Victims: Data Analysis Of The Most At Risk Population

Erick Orellana
George Mason University
eorellan@gmu.edu

Arush Kukreja
George Mason University
akukreja@gmu.edu

Kening Zhu
George Mason University
Kzhu5@gmu.edu

Sohan Reddy Siddenki
George Mason University
ssiddenk@gmu.edu

Abstract—According to a report published by The City University of New York, it is estimated that in the year 2018 the number of visitors to New York City hit a record of 65.1 million (Tourism, 2020). New York City also consistently ranks high in lists consisting of the most populated cities in the United States. Even with the city already being so densely populated every year there are many people opting to make New York City their home because of the city's historical, economic, and cultural impact. We use machine learning models to analyse victim data in this research.

Index Terms—Security, Machine learning, Data, Victims.

I. INTRODUCTION

Unfortunately, it has been reported recently in an article ABC 7 News that shootings in New York City have risen to levels that had not been seen in years (WABC, 2020). With so many people visiting and choosing to make New York City their home combined with the increased safety concerns around gun violence, it creates a problem of how one can understand where in New York City a person is most likely to become a victim of gun violence. New York City is split up into five boroughs. These boroughs are Queens, Bronx, Manhattan, Staten Island, and Brooklyn. There have been varying degrees of shooting incidents in the different neighborhoods over the years. When a shooting is reported, the New York City Police Department reports on the borough where the incident occurred as well as whether the shooting was fatal or not. This leads to the first question of this study:

1. How do the five boroughs of New York City compare in terms of total fatal and non-fatal incidents in 2019?

New York City prides itself in being a city with a lot of diversity in terms of ethnicities, cultures, and ages to name just a few factors contributing to the diversity. Due to the diversity of the population, victims of shooting incidents fall into different categories in terms of age group, race, and gender. This leads to the second question of this study:

2. Which group of people (as determined by gender, age group, and race) are at a higher risk of being a victim of shooting incidents for each borough?

The New York City Police department provides the time and date of all shooting incidents. A preliminary look at the data shows that incidents happen throughout different times of the day. This piece of information leads to the third question of this study:

3. For the most at-risk group, what time interval of the day is considered the most dangerous?

In recent years, due to the frequent occurrence of shootings, different government agencies in conjunction with the police department of each administrative region have worked together to prevent the frequency of shootings as to ensure the normal life of the people. Understanding which borough is having the most success may help in formulating a plan for those areas seeing less success. This leads to the following question for this study:

4. Has any borough seen a significant improvement in the prevention of shooting incidents over the last ten years for the most at-risk group?

In the current political climate, there is often a discussion surrounding gun control and gun safety in the United States. It is not uncommon to see on the news reports of high-profile shooting incidents in public places or in private settings. There are also many more shootings that go unreported to the public happening every day. New York City is no exception to these occurrences. Often times when these incidents happen there is a lot of effort that goes into creating a profile of the shooter and the public conversation turns to how to identify potential perpetrators of gun violence to prevent them from accessing guns. This project flips the script on that and is motivated by providing insights into what group of people are at a higher risk of becoming a victim of a shooting incident. The dataset may require pre-processing in order to be utilized by each of the tools evaluated. The pre-processing that occurs will be done in accordance with best practices documented by the tools being used.

This report can be used to aid individuals in making better informed choices regarding their safety when deciding to visit or reside in New York City. On a larger scale, this report can be used to uncover which boroughs, if any, have made improvements in gun safety and which have not. The pre-processing that occurs will not affect any intrinsic bias as to keep biases consistent across processed data sets and tools [1]–[9].

This has the potential of finding the factors that make the current solutions effective or ineffective in curbing the number of gun shooting incidents. Lastly, identifying the most at risk group of people can be used to ensure that the correct community health or social welfare programs are available so that that potential victims do not end up on the New York City's report of shooting incidents. [10] [11], [12] [13], [4] . [11], and [12] [14]. Social media investigation and Natural

language processing model is used to analyse crime data. [15]–[33].

II. LITERATURE SEARCH

When a shooting incident occurs, there are various possible outcomes for the victim of the act of violence. One possible outcome is the life of the victim ending because of the incident. Another outcome is the victim surviving the shooting incident with the potential of having short term or long term physical and mental trauma. It is reported by the Substance Abuse and Mental Health Services Administration that there are three stages that a victim goes through after an incident: “the acute phase immediately after the event, the intermediate phase several days to weeks afterward and the long-term phase” (Novotney, 2018). There are shooting victims that come out of these three stages and can come to an acceptance of it and find a way to integrate back into their everyday life without major impacts to their mental health.

However, there are other victims that are not able to recover mentally from the incident. According to a study published by the American Psychological Association surviving victims of shooting incidents have an increased risk of experiencing lasting mental health problems such as depression, substance abuse and post-traumatic stress (Novotney, 2018). Another study by The National Child Traumatic Stress Network discovered that in some cases surviving victims of shooting incidents “become overly confrontational or aggressive, or engage in high-risk behaviors (e.g., driving recklessly, using drugs and alcohol)” (NCTSN, 2014). It is evident that surviving a shooting incident may cause a major decline to the quality of life of the surviving victim which is why it is important to ensure that potential shooting incident victims are protected before becoming actual victims.

While there are proven negative psychological impacts to a surviving victim of a shooting incident, there are also negative impacts to the public health and safety of the community. It is reported that after a shooting incident, some surviving victims fall into a depression that can lead them to feel worthless or hopeless (NCTSN, 2014). This feeling of hopelessness that a shooting victim may be experiencing along with being driven to becoming aggressive or to developing substance abuse problems can put the safety of themselves and of the public at risk. A feeling of hopelessness combined with mental issues such as substance abuse has been proven to lead individuals to becoming the perpetrators of shooting incidents. Being able to prevent a person from becoming a victim of a shooting incident can also prevent them from becoming a risk factor for future shooting incidents which can ultimately help make communities safer.

The Prevention Institute has published a list of suggestions that address ways in which a potential victim of a shooting incident can be provided support that would keep them from becoming victims. The Prevention Institute explains that “people with mental illness are more likely to become victims of violence than perpetrators” (Prevention). For this reason, it is important that people with mental health issues are being

provided the proper support and resources to prevent them from becoming victims of gun violence. This support could be giving potential victims the knowledge needed to detect when a situation is escalating to a point where they are at risk of being in a gun shooting incident. To provide this education in the most effective way possible it is important to understand who the audience.

For this reason, it is useful to have a profile of the most at risk individuals on things such as age or location to ensure that information distribution is being given in a way that the potential victim will be able to process it. For example, depending on the age group most at risk, it may be more beneficial to distribute this information through social media rather than only distributing this information in grade school settings where the people who need it most will not get to it. By giving potential victims the tools to assess certain situations it could help in them removing themselves from the situation or seeking the best type of help to get them out of the situation.

The Prevention Institute also suggests that adequate community planning along with implementing community safety plans that focus on prevention and intervention are important to ensuring the safety of at-risk victims of shooting incidents. One intervention strategy is hospital-based support. According to the American Academy of Family Physicians bringing up the conversation of guns at home for households with children or bringing up the conversation with childbearing aged women of intimate partner violence can help to raise concerns of being at risk of being a victim of a shooting incident (AAFP).

Currently some states do not allow this conversation to be brought up but having the data show that these groups have statistically been victims of shooting incidents could help sway policy towards including this conversation in health screenings. Another strategy that could help prevent at risk people from becoming victims of shooting incidents is implementing upstream social programs such as youth employment, safe parks, reducing alcohol outlet density (Prevention). A community health program in Minneapolis focused on using upstream strategies to keep young potential at-risk victims safe and saw a 62% reduction of gun shooting incidents (Prevention). This again goes to show that it is important to have a profile on who the most at risk population is so that various prevention strategies can be tailored towards those at-risk of becoming victims of shooting incidents.

Another way to protect victims of shooting incidents from becoming involved in those situations is to deploy law enforcement resources to the correct places at the correct times. One example of this is Los Angeles’ “Operation Cul-de-Sac.” A study was conducted, and it was determined that there were certain geographic locations in Los Angeles where there was an increased number of drive-by shootings.

The police department took this information and installed traffic barriers in specific locations which helped to block drive-through traffic and effectively reduce the number of shooting incidents in the area (Peterson, 2020). Another example of using the geographic profile of the most at-risk population occurred in Chula Vista, California, where analysis

was done to find motels and bars where there was a high level of shooting incidents. This information was then used to police those motels and bars more closely. The information was also used to get the property management companies of those establishments to update their policies and responses to acts that put victims at risk of shooting incidents. Having a profile of the most at-risk individuals along with the geographic location of those at-risk of becoming victims can be effective in reducing the overall number of shooting incidents.

Repeatedly it has been proven that to have a successful strategy towards preventing the potential shooting victims from becoming involved in the situation requires having a profile on the at-risk population. This means understanding things such as the age, race, and gender of a potential victim. This also means understanding things such as geographical location of the most at-risk population or time of day where they are most at risk. In this project a profile is being built for shooting victims in New York City to get a better understanding of what prevention strategies could be the most effective in curving the growing number of instances of gun violence.

III. PROPOSED APPROACH

This project has the dataset titled “NYPD Shooting Incident Data (Historic)” at its core. This data set is retrieved from the New York Police Department’s Shooting incident data and this data can be used by the public to explore and gain knowledge about the shooting incidents in New York City. This data set

From there we will following the points of interest to get into a more granular level by combining points of interest that are revealed to have some relevance to each other in terms of building a profile for the most at risk demographic in terms of becoming a victim of a shooting incident. Lastly, after getting to a granular level with calculations and visualizations, we will be able to propose a profile for the most at-risk demographic for shooting incidents in New York City.

IV. PRELIMINARY RESULTS

New York City suffers from a high rate of shooting incident, which sometimes have a fatal outcome. Most effort by authorities is focused on perpetrators of crime, analyzing their social and demographic characteristics to identify factors that can provide indications for developing a support system. However, it is also important to understand if any demography is more susceptible to such attacks. In addition to this, information on how citizens can be affected in different areas of New York City can help deploy resources selectively while planning preventive measures, thus optimizing expenditure.

An overview of incidents indicates that most incidents took place in Brooklyn and Bronx. Staten Island is comparatively safe when looking at all five boroughs of New York City. The preliminary research has uncovered that overall, 19% of incidents are fatal for the victim. Most victims of shooting incidents are Black, constituting over 70% of the affected population, followed by Hispanics who contribute to 23% of the cases. At face value it is apparent that the Black population is disproportionately a victim of shooting incidents. This can

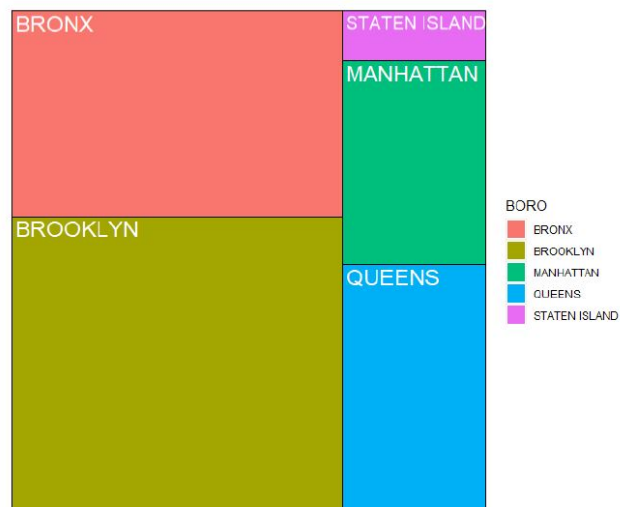


Fig. 1. Number of cases/deaths across time

be further investigated by seeing what portion of the over all population is Black as compared to other races to see if this disproportion is a reflection of the overall population of the five boroughs or if it is truly a large disproportion.

REFERENCES

- [1] S. Zad, M. Heidari, J. H. J. Jones, and O. Uzuner, “Emotion detection of textual data: An interdisciplinary survey,” in *IEEE 2021 World AI IoT Congress, AIIoT2021*, 2021.
- [2] A. Adekunle, M. Meehan, D. Rojas-Alvarez, J. Trauer, and E. McBryde, “Delaying the COVID-19 epidemic in australia: evaluating the effectiveness of international travel bans,” *Australian and New Zealand Journal of Public Health*, vol. 44, pp. 257–259, July 2020.
- [3] M. Heidari, S. Zad, B. Berlin, and S. Rafatirad, “Ontology creation model based on attention mechanism for a specific business domain,” in *IEEE 2021 International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021*, 2021.
- [4] M. Bielecki, D. Patel, J. Hinkelbein, M. Komorowski, J. Kester, S. Ebrahim, A. J. Rodriguez-Morales, Z. A. Memish, and P. Schlagenhauf, “Air travel and COVID-19 prevention in the pandemic and peri-pandemic period: A narrative review,” *Travel Medicine and Infectious Disease*, vol. 39, p. 101915, Jan. 2021.
- [5] M. Heidari and J. H. Jones, “Using bert to extract topic-independent sentiment features for social media bot detection,” in *2020 11th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pp. 0542–0547, 2020.
- [6] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. y Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini, and A. Vespignani, “The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak,” *Science*, vol. 368, pp. 395–400, Mar. 2020.
- [7] S. Zad, M. Heidari, J. H. J. Jones, and O. Uzuner, “A survey on concept-level sentiment analysis techniques of textual data,” in *IEEE 2021 World AI IoT Congress, AIIoT2021*, 2021.
- [8] M. Heidari and S. Rafatirad, “Using transfer learning approach to implement convolutional neural network model to recommend airline tickets by using online reviews,” in *2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA)*, pp. 1–6, 2020.
- [9] M. Heidari, J. H. J. Jones, and O. Uzuner, “Misinformation detection model to prevent spread of the covid-19 virus during the pandemic,” in *IEEE 2020 International Symposium on Technology and Society (ISTAS20)*, ISTAS20 2020, 2020.

- [10] M. Heidari, J. H. J. Jones, and O. Uzuner, "Offensive behaviour detection on social media platforms by using natural language processing models," in *IEEE 2020 International Symposium on Technology and Society (ISTAS20)*, ISTAS20 2020, 2020.
- [11] S. H. Bae, H. Shin, H.-Y. Koo, S. W. Lee, J. M. Yang, and D. K. Yon, "Asymptomatic transmission of SARS-CoV-2 on evacuation flight," *Emerging Infectious Diseases*, vol. 26, pp. 2705–2708, Nov. 2020.
- [12] E. M. Choi, D. K. Chu, P. K. Cheng, D. N. Tsang, M. Peiris, D. G. Bausch, L. L. Poon, and D. Watson-Jones, "In-flight transmission of SARS-CoV-2," *Emerging Infectious Diseases*, vol. 26, pp. 2713–2716, Nov. 2020.
- [13] N. C. Khanh, P. Q. Thai, H.-L. Quach, N.-A. H. Thi, P. C. Dinh, T. N. Duong, L. T. Q. Mai, N. D. Nghia, T. A. Tu, L. N. Quang, T. D. Quang, T.-T. Nguyen, F. Vogt, and D. D. Anh, "Transmission of SARS-CoV 2 during long-haul flight," *Emerging Infectious Diseases*, vol. 26, pp. 2617–2624, Nov. 2020.
- [14] T. W. Russell, J. T. Wu, S. Clifford, W. J. Edmunds, A. J. Kucharski, and M. Jit, "Effect of internationally imported cases on internal spread of COVID-19: a mathematical modelling study," *The Lancet Public Health*, vol. 6, pp. e12–e20, Jan. 2021.
- [15] S. Chen, S. Owusu, and L. Zhou, "Social network based recommendation systems: A short survey," in *2013 International Conference on Social Computing*, pp. 882–885, 2013.
- [16] S. Lin, C. Liu, and Z.-K. Zhang, "Multi-tasking link prediction on coupled networks via the factor graph model," in *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, pp. 5570–5574, 2017.
- [17] M. Heidari, J. H. J. Jones, and O. Uzuner, "Deep contextualized word embedding for text-based online user profiling to detect social bots on twitter," in *IEEE 2020 International Conference on Data Mining Workshops (ICDMW)*, ICDMW 2020, 2020.
- [18] Y. Chu, F. Huang, H. Wang, G. Li, and X. Song, "Short-term recommendation with recurrent neural networks," in *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 927–932, 2017.
- [19] C. Yang, X. Chen, T. Song, B. Jiang, and Q. Liu, "A hybrid recommendation algorithm based on heuristic similarity and trust measure," in *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, pp. 1413–1418, 2018.
- [20] S. Ji and J. Liu, "Interpersonal ties and the social link recommendation problem," in *2019 6th International Conference on Systems and Informatics (ICSAI)*, pp. 456–462, 2019.
- [21] M. Heidari and S. Rafatirad, "Bidirectional transformer based on online text-based information to implement convolutional neural network model for secure business investment," in *IEEE 2020 International Symposium on Technology and Society (ISTAS20)*, ISTAS20 2020, 2020.
- [22] J. Wang, H. Song, and X. Zhou, "A collaborative filtering recommendation algorithm based on biclustering," in *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 803–807, 2015.
- [23] S. Chen, S. Owusu, and L. Zhou, "Social network based recommendation systems: A short survey," in *2013 International Conference on Social Computing*, pp. 882–885, 2013.
- [24] M. Heidari and S. Rafatirad, "Semantic convolutional neural network model for safe business investment by using bert," in *IEEE 2020 Seventh International Conference on Social Networks Analysis, Management and Security, SNAMS 2020*, 2020.
- [25] A. Gatzoura, J. Vinagre, A. M. Jorge, and M. Sánchez-Marrè, "A hybrid recommender system for improving automatic playlist continuation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 5, pp. 1819–1830, 2021.
- [26] Z. Liao, Y. Song, Y. Huang, L.-w. He, and Q. He, "Task trail: An effective segmentation of user search behavior," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 12, pp. 3090–3102, 2014.
- [27] M. Heidari, J. H. J. Jones, and O. Uzuner, "An empirical study of machine learning algorithms for social media bot detection," in *IEEE 2021 International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021*, 2021.
- [28] C.-Y. Chi, Y.-S. Wu, W.-r. Chu, D. C. Wu, J. Y.-j. Hsu, and R. T.-H. Tsai, "The power of words: Enhancing music mood estimation with textual input of lyrics," in *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1–6, 2009.
- [29] A. Gatzoura, J. Vinagre, A. M. Jorge, and M. Sánchez-Marrè, "A hybrid recommender system for improving automatic playlist continuation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 5, pp. 1819–1830, 2021.
- [30] H. Yang, C. He, H. Zhu, and W. Song, "Prediction of slant path rain attenuation based on artificial neural network," in *2000 IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 1, pp. 152–155 vol.1, 2000.
- [31] M. Heidari, S. Zad, and S. Rafatirad, "Ensemble of supervised and unsupervised learning models to predict a profitable business decision," in *IEEE 2021 International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021*, 2021.
- [32] M. Thompson, G. Alshabana, T. Tran, and A. Chitimalla, "Predict covid-19 cases using opensky data." <http://mason.gmu.edu/~ttran81/>, 2021.
- [33] "Crime report in new york." <https://abc7ny.com/new-york-city-crime-nyc-shootings-gun/20violence/8495935/>, (2020, September 12).