



**QMST 5336 – ANALYTICS**  
**GROUP PROJECT REPORT**  
**GLOBAL TERRORISM IN NORTH AMERICA**

BY SIAOSI FINAU, DEEPTHI KANDREGULA AND LIMLA PAUL

## **Problem Definition:**

**Problem:** LIMITING INCIDENTS OF TERRORISM IN NORTH AMERICA

### **Introduction**

Terrorist attacks has had a negative effect in the world for the past few decades. Some terrorist attacks were caused by wars, political views, socioeconomic, and many related reasons. Therefore, we decided to use a database to show the trend of terrorist attacks for the past few years to show where and the cause of these attacks. Our team used the Global Terrorism Database set that we retrieved from Kaggle (START Consortium). This database is considered reliable as it is maintained by researchers at the National Consortium for the Study of Terrorism and Responses to Terrorism (START), headquartered at the University of Maryland (National Consortium for the Study of Terrorism and Responses to Terrorism, 2018) This database populated over 180,000 events with more than 100 variables from years 1970 through 2017. Each event provided a wide range of information, including year and location of the incident, the weapons used, nature of the target, and the number of casualties. With this information, we believe that we can provide useful information to people about where most attacks are occurring and the reasoning behind the attacks for personal use, like traveling purposes for example. With our research and the information provided in the global terrorism database, we have a theory of how we can minimize terrorist attacks. With this, we focused particularity on New York City as it's a popular city with high crime rates. Using the global terrorism database to determine the trend in attacks and translating the information to specifically focusing on New York City can help our team provide useful information to help cities mitigate terrorist/crime events.

### **Visualization:**

## Descriptive Analysis

Our descriptive analysis consists of graphs to help interpret our data visually. These graphs will show the trend with the number of attacks from years 1970 – 2017, the type of weapons used, if the attacks were successful, and the people/groups that were targeted. As mentioned in our introduction, we will use the graphs to help focus on mitigating attacks in North America, particularly New York City. The Global Terrorism Database was filtered to get data based on the region 'North America' for which the doubt-terrorism column showed '0' which means there is no doubt that the attack is a terrorist activity.

North America includes United States, Canada and Mexico. There were a total of 2675 records out of the 180,000 records in the database. The main columns of interest were as follows: year, approximate date, doubt-terrorism, region, city, attack types, success, weapon types, target types, and property damage in USD (US dollars). Out of the 11 columns, we had to remove two columns as they had lot of missing data (approximate date – 2,646 missing, property damage – 1,810 missing). Keeping the above two columns and removing only the missing fields would reduce the quantity of data. You may refer to the appendix section to view the graphs that were populated throughout our research. Also, please note that the terrorist attacks are not only considered attacks that were taken action from another country, but attacks that occurred within its own country, like local shooting, bombing, and cyber-attacks.

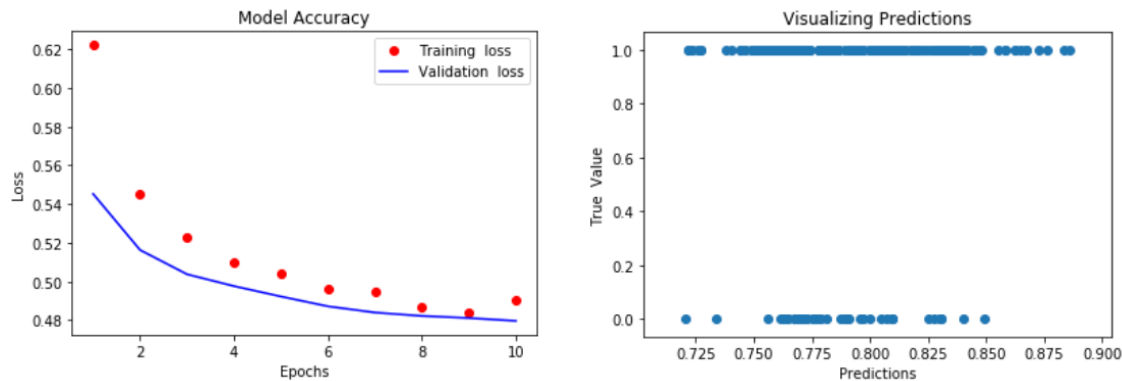
Graphs can be found in the Appendix and it focuses on the trend of attacks that occurred in North America (USA, Canada, and Mexico). There's been a slight increase since 2010 due to political, economic, religious, or social purposes. Public or private property damages are to be considered when dealing with our data.

## Decision Making:

### Predictive Analysis

From the descriptive analysis, we got a better understanding of our dataset. Hence, we decided to build our model with the following 5 features: success, year, attack types, target types, and weapon types. Classification model was used as we are dealing with categorical variables. Also, Neural Network Model was used to predict the attack in a region, and we wanted to classify the attack as success or no success. After cleaning our dataset, we had a total of 2675 records. Our model was divided randomly into 85% for training, 10% for validation and 5% for test (Training on 1925 samples, validating on 482 samples). Our simple neural network model had 1 input layer and 1 hidden layer with 128 units each and 1 output layer with 1 unit. We configured our model with Adam optimizer and evaluated our model on the basis of binary cross-entropy loss. We used Relu activation for input and hidden layers and Sigmoid activation for output layer. To combat problem of over-fitting, we used dropout regularization. After 10 epochs of training, our model achieved 82.23% accuracy on training set and 82.78% accuracy on validation set. The gap between training and validation set accuracy is small, but our model is slightly overfitting as seen from the graphs below. Our testing set accuracy is 83%. Therefore, our model can predict the success of an attack in North America with 83% accuracy.

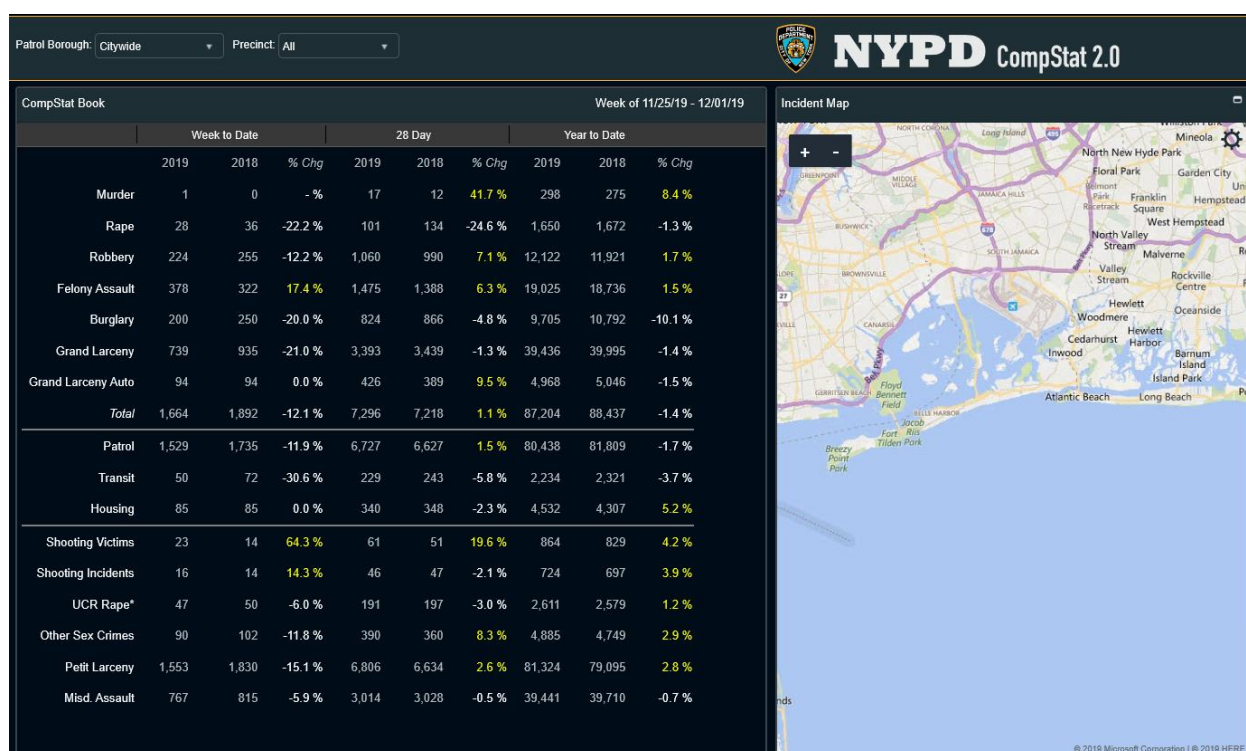
We tried different approaches to get better accuracy. In order to better improve our model, we would likely need more data and more features to increase training accuracy.



## Prescriptive Analysis:

In order to design effective terrorism prevention programs, we must begin with the nature and parameters of the threat those efforts are seeking to address. An understanding of the threat provides a baseline for assessing whether terrorism prevention efforts are adequate and serves as a point of departure for determining what types of additional terrorism prevention programming are needed. Mass-casualty attacks have been prominent in national experience of terrorism, including the 1995 bombing of the Alfred P. Murrah Federal Building in Oklahoma City that killed 168 people and the September 11, 2001 (9/11), attacks that resulted in the death of almost 3,000 individuals and the injury of many more people, including those with long-term health effects from the attacks themselves and from subsequent response operations. 9/11 galvanized the national response to terrorism, leading to the formation of DHS, among other national actions and policy changes. DHS is charged with addressing not just the risk of similar large-scale attacks, but also smaller-scale, and more-frequent terrorist threats. Although the potential for individuals to radicalize to violence had previously raised concern, efforts in the wake of 9/11 built the foundation for CVE in the United States and the subsequent expansion in the succeeding years.

In this age of urban surveillance, it's hard to believe that a decade ago the New York Police Department didn't have a single camera in Times Square — or on any city street, for that matter. It wasn't until 2006 that the NYPD deployed its first 500 street-level security cameras to the tune of \$9 million. They quickly multiplied, and today, the department has about 2,000 cameras scanning the streets, sidewalks, rooftops, parks, bridges and tunnels of New York night and day. There's 7,000 more in public housing and another 4,000 in the subway. Last year the city had 335 murders, its second fewest in the modern era. The number of shootings fell below 1,000 for the first time since record keeping began, and the number of felonies was the smallest since [CompStat](#) crime tracking started 23 years ago. Our dataset contains information from 1970-2017. And as we see that the focus of predictive and descriptive is towards North America, below are some interesting findings for the preventive measures on terrorism in New York.



**Figure 7:** The above figures show the New York city patrolling data to find different incidents.

CCTV cameras and police patrolling are preventive measures for Terrorism, but which one will be more cost efficient? How can we find the maximum number of attacks and also manage the budgets. What are the budgets which need to put approximately for preventing an attack? All these questions are answered on analyzing the data from various sources for New York City crime index through police patrolling and CCTV camera footage. First, we get the data and then we categorized our data into Terrorism by taking all the suspects under acts of terrorism like weapons, bombing/explosion, hostage etc. On analyzing the data above we follow a linear programming approach to formula our data with decision variables as, number of attacks/suspects using police patrols i.e. New York arrest rate and number of attacks/suspects using CCTV cameras i.e. New York number of attacks/suspects. Our objective is to find the maximum number of suspects and be cost efficient on CCTV cameras and police patrols. So, we found the maximum number of suspects as 16750 for the year 2017 with a cost of 50USD per camera and 30USD per cop. We have 4.23 officers per 1K population and 1.28 cameras per 1K population with a maximum crime index risk of 43.8. We can prevent the attacks by increasing the cost of cameras by 335USD and keeping the police patrol amount the same.

### **Analytics & Recommendations:**

As per the analytics shown above, we can minimize the attacks yearly by using below methods:

- Doing Descriptive analysis of the terrorism dataset using various visualization tools
- Performing Predictive analysis of future attacks using Neural Network Model
- Executing Prescriptive Analysis with Linear Programming Model

We recommend retaining the cost of each police patrol officer per hour to be maintained the same while maximizing the number of CCTV cameras in New York as they are more effective in minimizing attacks and its cost efficient for the New York Police Department.

### **Critical Thinking:**

- 1) Dataset had missing values for 'approximate date' column which made it challenging to do time-series predictive forecasting for New York City. Having robust data in the "approximate date" column for future events may help in formulating better models for predicting the success of terrorist attacks in the future.
- 2) As mentioned in the predictive analysis - we would likely need more accurate data for North America and more features to increase the training accuracy of our model.
- 3) With prescriptive analysis, it's hard to find data related to preventive measures such as police patrolling and approximate number of CCTV cameras for entire North America. Hence, the prescriptive analysis was centered on New York City. Moreover, a lot of research was required to gather information related to number of suspects (accused) and the cost estimates for preventive measures.

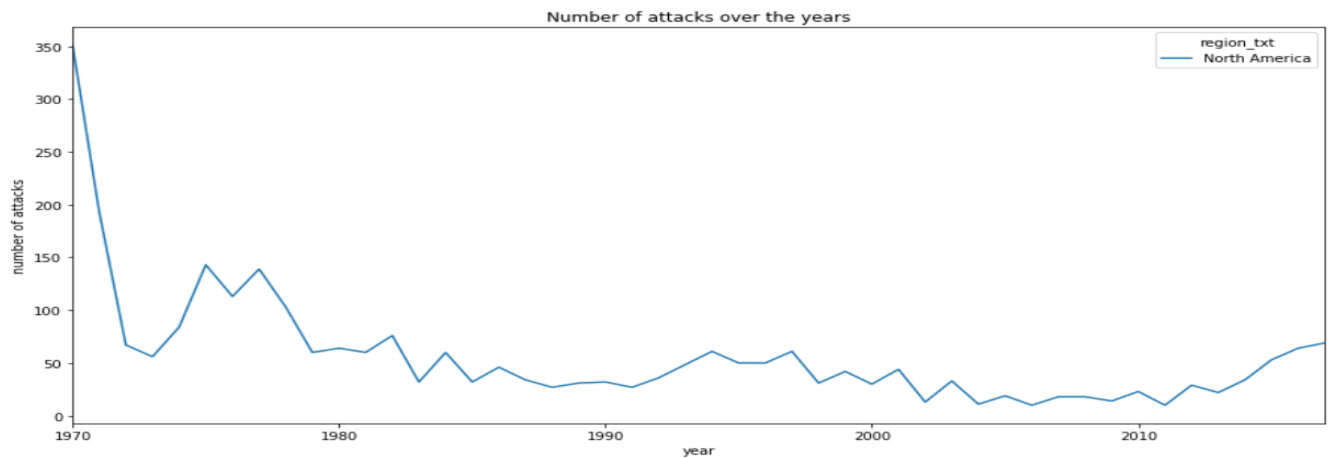
### **Conclusion:**

With the information gathered from the Global Terrorism Database and further research, we have provided valuable information for potential users, such as the Department of Defense and City Officials of New York. Users can see the trend in the number of attacks that occurred in North America over the years and the most preferred weapon types used. Our predictive analysis suggests that the successful rate of an attack occurring in North America is highly likely and therefore precautions should be set in place. Similarly, we believe adding CCTV to cities will

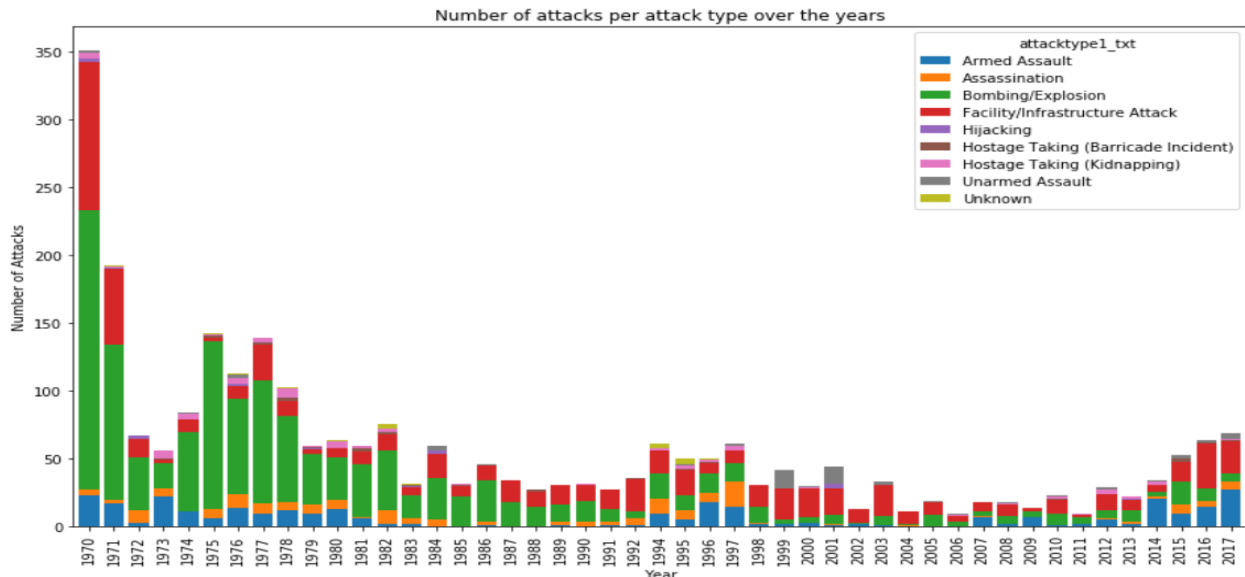


help decrease the amount of attacks/crime. Terrorist attacks are inevitable, as people around the world have different beliefs, interest, or desires and so our research is not to prevent the inevitable, but to help mitigate these events.

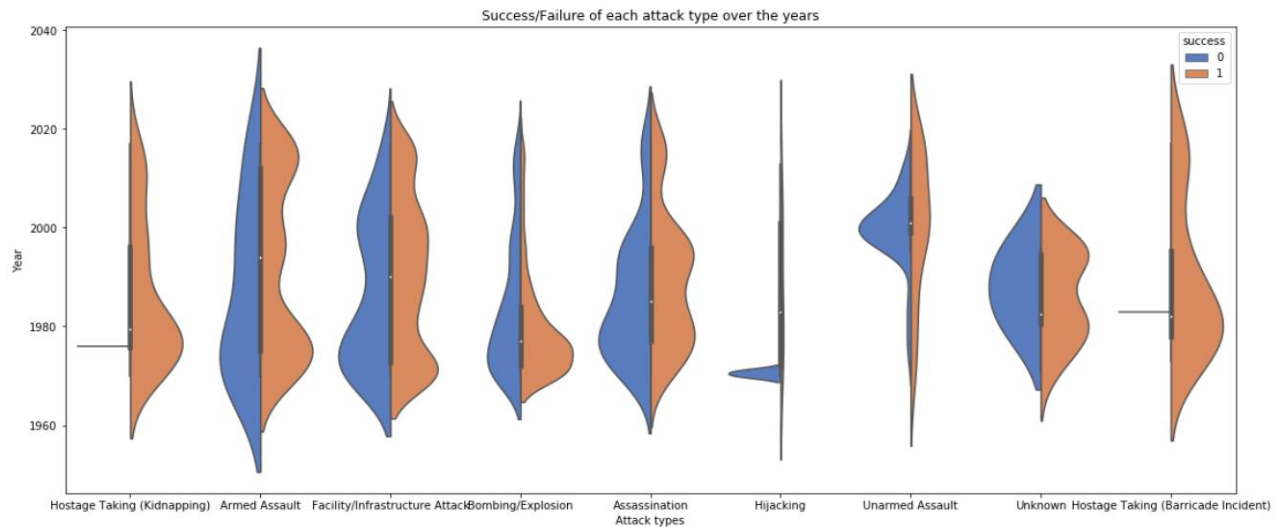
## Appendix



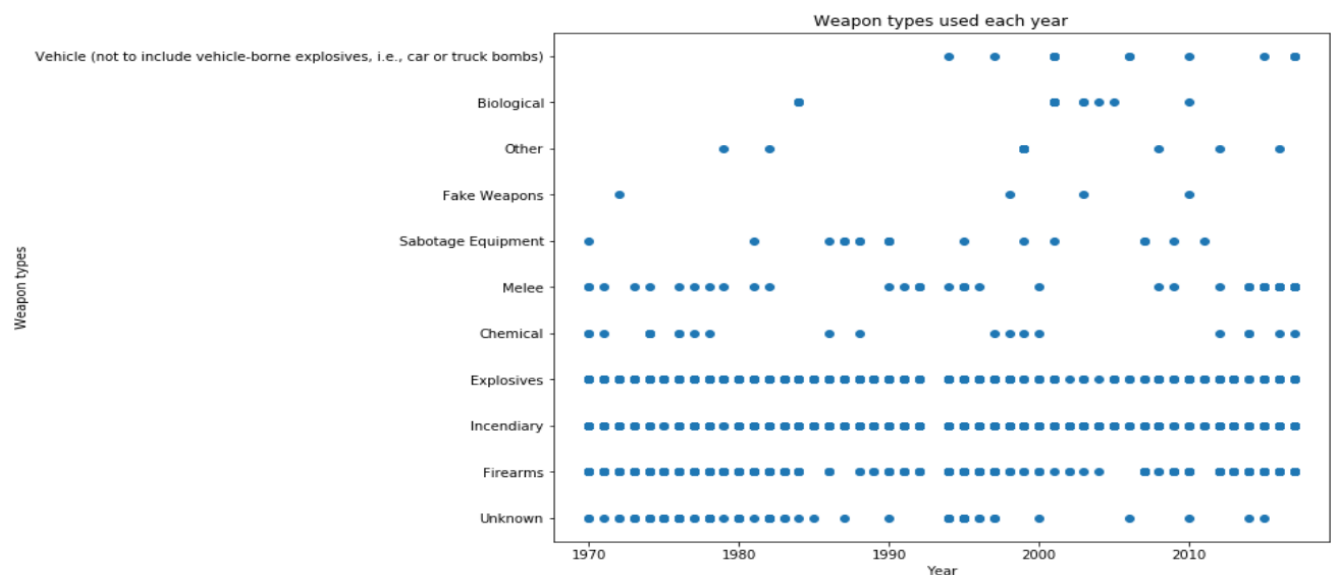
**Figure 1:** Line plot for the number of attacks has reduced considerably since the 1970's. But the trend shows that number of attacks have been slowly but steadily increasing after 2010.



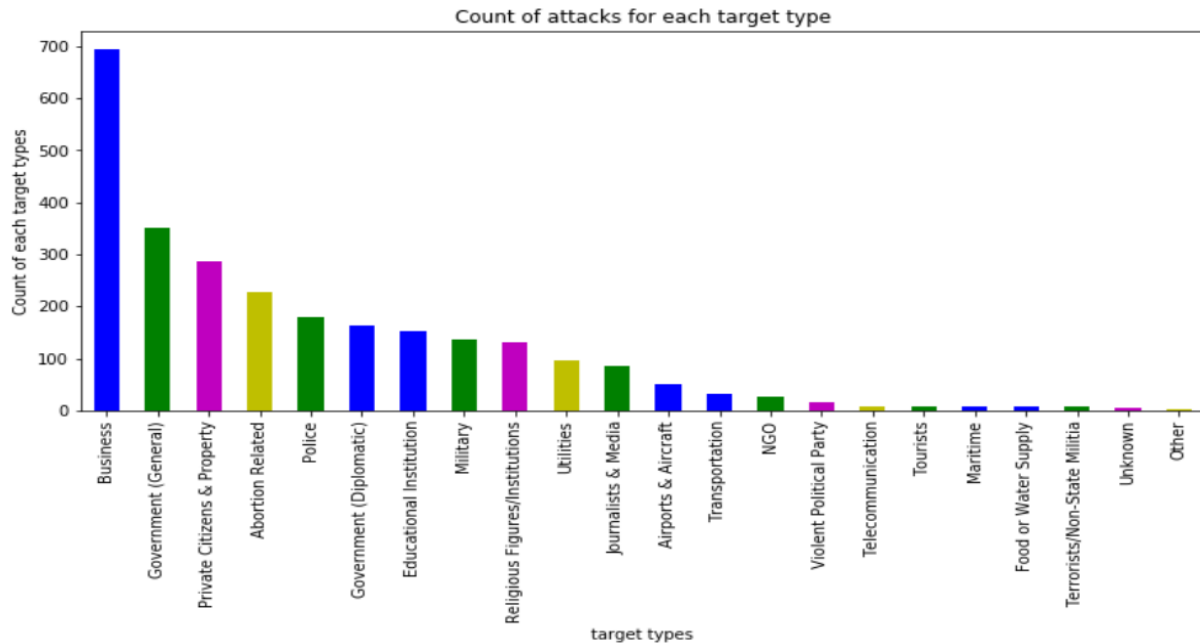
**Figure 2:** Bar graph showing that out of all the attack types, Bombing/Explosion is the preferred means of attack followed by Facility/Infrastructure attack by terrorists over the years.



**Figure 3:** Violin plots based on the success of attacks. Success = '1' (orange color) means the incident was successful. Success = '0' (blue color) means the incident was not successful. The plot shows that there is success for all types of attacks except for Hijacking over the years.



**Figure 4:** Scatter plot of weapon types used over the years. Explosives/Incendiary and Firearms are the most preferred weapon types by the terrorists.



**Figure 5:** Bar graph showing the number of attacks based on each target type over the years.



Mostly Businesses, Government, Private Citizens & Property are most affected over the years.

This damaged the economic growth of any region or country.

**Figure 6:** Correlation heat map showing strength of relationships between variables in the dataset. Success has weak positive correlation with factors such as year, target types and weapon types.

## Data Resource

1. "National Consortium for the Study of Terrorism and Responses to Terrorism (START), University of Maryland. (2018). The Global Terrorism Database (GTD) [Data file]. Retrieved from <https://www.start.umd.edu/gtd>"
2. Practical Terrorism Prevention, Executive Summary by Brian A. Jackson, Ashley L. Rhoades, Jordan R. Reimer, Natasha Lander, Katherine Costello, Sina Beaghley. (2019). Retrieved from [https://www.rand.org/pubs/research\\_reports/RR2647z2.html](https://www.rand.org/pubs/research_reports/RR2647z2.html)
3. (START Consortium) Global Terrorism Database retrieved from <https://www.kaggle.com/START-UMD/gtd>
4. City of New York: [NYPD Arrests Data \(Historic\)](#). (2019). Retrieved from <https://data.cityofnewyork.us/api/views/8h9b-rp9u/rows.csv?accessType=DOWNLOAD>
5. New Yorkers Call for More Surveillance Cameras. Will Doig. (2017). Retrieved from <https://nextcity.org/daily/entry/new-york-surveillance-cameras-police-safety>
6. <https://catalog.data.gov/dataset>
7. <https://www.mi5.gov.uk/terrorist-targets>
8. United Nations Office on Drugs and Crime. Retrieved from [https://www.unodc.org/images/odccp/terrorism\\_weapons\\_conventional.html](https://www.unodc.org/images/odccp/terrorism_weapons_conventional.html)