**Introduction**

The dataset under examination is a comprehensive collection of historic NYPD shooting incident data, which encompasses an extensive range of details pertaining to each incident. Comprising a total of 27312 rows and 21 columns, this dataset presents about shootings that occurred in various locations across New York City. It allows us to delve into the time, location, and demographics surrounding these incidents.

Each part of this dataset offers a glimpse into the circumstances of individual shooting incidents, detailing information such as the date and time of occurrence, the borough in which the incident took place, the precinct, and specific locations, among other pertinent data. Additionally, it provides insights into the demographics of both the victims and the suspects, including age, gender, and race, enabling a comprehensive analysis of the profiles associated with these incidents.

Understanding the complexities and patterns of gun violence is essential for developing effective strategies for prevention and intervention. By exploring this dataset, researchers, policymakers, and law enforcement agencies can gain valuable insights into the factors contributing to these incidents and work towards a safer and more secure New York City.

**Background Information**

This dataset contains information about incidents where shootings took place in New York City. It provides details about when and where these shootings occurred, as well as information about the people involved. This data helps us understand more about these incidents, and it can be used to make the city safer.

**Date and Time**: When the shootings happened.

**Location**: Where in New York City the shootings occurred.

**Demographics**: Information about the people involved, like their age, gender, and race.

This data is useful for researchers, the police, and policymakers who want to find ways to prevent these shootings and make the city a safer place. By studying this data, we can better understand the reasons behind these incidents and work to reduce them.

**Aim/Hypothesis Statement:**

The aim of this study is to understand and analyze the patterns and factors associated with shootings in New York City. We want to find out why and where these incidents happen and who is involved.

Our hypothesis is that by examining this data, we can identify trends and insights that will help us develop strategies to prevent shootings and improve safety in the city. We believe that this information can lead to better policies and practices that reduce the occurrence of these incidents.

**Hypothesis**:

The analysis of historic shooting incident data in New York City will reveal distinct patterns and factors contributing to these incidents. We hypothesize that the time, location, demographics of individuals involved, and other contextual details will show correlations that can inform targeted crime prevention efforts. Furthermore, we anticipate that this research will highlight the importance of socio-economic factors in the occurrence of shootings and provide valuable insights for law enforcement and policymakers to develop effective strategies for reducing gun violence in the city. I predict if the incident is as a murder or not depend on other factors.

**Methods Evaluation:**

**Data Preprocessing**:

Data preprocessing was a fundamental step to clean and prepare the dataset for analysis. Ideally we perform tasks such as handling missing values, standardizing data formats, and removing duplicates. This ensured that the dataset was in an optimal condition for analysis, preventing errors and bias in our results.

For this dataset,

I had checked for null values and to my surprise data was so many missing values.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

I changed datatype of Occur\_Date from object to to datetime and added year, month, day

**Method Evaluation Cont. Data Mining Methods Description:**

**Finding hidden Patterns:**

A screen shot of a graph

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a graph

Description automatically generated

A screenshot of a computer

Description automatically generated

**A screenshot of a computer

Description automatically generated**

**A graph with blue lines

Description automatically generated**

**A graph with red and blue lines

Description automatically generated**

**Logistic Regression:**

It works by analyzing the data to create a special curve that shows the probability of one of the two choices happening. For example, it can tell you the probability of someone voting based on their age. If the probability is high, it means they are more likely to vote. If it's low, they are less likely.

So, logistic regression is like a detective that helps us understand the chances of something happening when we have two options, and it's really handy in all kinds of fields, like medicine, marketing, and more.

**Results Evaluation:**

We evaluate the results by examining the patterns, trends, and statistics related to shooting incidents. This could involve analyzing factors like the locations, times, and demographics of those involved in the shootings. By evaluating these results, you can draw conclusions, make informed decisions, and better understand the impact of these incidents on a specific area or community.

A screenshot of a computer

Description automatically generated

**Discussion Evaluation:**

First, we found that most of the incidents happened in certain boroughs, like Brooklyn and the Bronx. This tells us that these areas might need more attention when it comes to safety.

We also saw that there were different times when these shootings occurred. Some happened late at night, while others occurred during the day. This information can help law enforcement plan their resources better to address these incidents at different times.

The data also showed that there were different age groups and ethnicities involved in these incidents. This can help us understand who is most affected and where efforts should be focused to prevent these incidents from happening.

However, we should remember that this data only shows us what happened in the past. To make our neighborhoods safer, we need to use this information to create strategies and policies. This means more than just looking at numbers – it means taking action.

**Conclusion Evaluation:**

In conclusion, the data on shooting incidents in New York City provides valuable insights into the patterns and demographics of such incidents. We have observed that these incidents are not evenly distributed across the city; certain boroughs, such as Brooklyn and the Bronx, experience a higher frequency of shootings. This indicates the need for targeted efforts in these areas to enhance safety.

Moreover, the data shows variations in the timing of these incidents, with some occurring late at night and others during the day. This information can aid law enforcement agencies in allocating resources more effectively, depending on the time of day.

Additionally, the involvement of different age groups and ethnicities in these incidents highlights the importance of understanding the demographics of those affected. Tailored strategies and interventions can be developed to address the specific needs of these groups and work towards preventing future incidents.

However, it's essential to emphasize that this data reflects past occurrences. To create safer neighborhoods, it is crucial to translate this knowledge into action. This requires the development and implementation of policies and initiatives aimed at reducing gun violence, with a focus on community engagement, education, and awareness.

In summary, the data analysis is a significant step towards comprehending the history of shooting incidents in New York City. Yet, the true test lies in how we utilize this information to make our communities safer for all residents. Collaborative efforts, informed decision-making, and proactive measures are key to achieving this goal and preventing future incidents.

**Referencing Evaluation**

New York City shooting victims: Data analysis of the most at-risk ... (n.d.). https://infowiz.gmu.edu/server/api/core/bitstreams/2e8114a7-4c94-41e9-9322-966074c32e80/content

RPubs. (n.d.). https://rpubs.com/mhdmarar139/1034848

*Vital City: Gun violence in New York City: The Data*. Vital City: New York City policy, culture and data. (n.d.). https://www.vitalcitynyc.org/vital\_signs/gun-violence-in-new-york-city-the-data