Project 4

Ruju Shah

Jash Shah

Vishwa Bhadiyadara

Malav Shah

—

NYPD SHOOTING DASHBOARD

Introduction

The NYPD Shooting Incident Dashboard is like a special website that shows information about shootings that happened in New York City. This report is all about understanding how this website works, what you can do with it, and how it shows the information. Imagine New York City, with all its hustle and bustle. Sometimes, unfortunate events like shootings happen there. This dashboard helps us understand where and when these shootings occurred, and who was involved. The website has different parts that you can click on to see different things. For example, you can choose a specific area in the city to see shootings only in that area. You can also see tables that list out details like the date, time, and location of each shooting. There are also pictures, called visualizations, that show the data in an easy-to-understand way. For instance, there are pie charts that show the age or gender of people involved in shootings. There are also maps that show where shootings happened in the city. Before all this information is shown on the website, the data needs to be prepared. This means making sure there are no missing pieces of information and organizing the data in a way that the website can understand. In simple terms, this report tells you how the NYPD Shooting Incident Dashboard helps us understand shootings in New York City by showing us clear pictures of the data. It's like a map that guides us through the city's safety issues, helping us make informed decisions for a safer future.

Dashboard Overview

The dashboard offers a user-friendly interface for exploring shooting incident data. It comprises several interactive components, including dropdown menus and various charts, allowing users to filter and visualize the data based on different criteria such as location, demographics, and time.

NYPD Shooting Incident Dashboard

All Locations

INCIDENT_KEY	OCCUR_DATE	OCCUR_TIME	BORO	PRECINCT	JURISDICTION_CODE	LOCATION_DESC	STATISTICAL_MURDER_FLAG	PERP_AGE_GROUP	PERP_SEX	PERP_RAC
211767651	4/5/2020 0:00	12/30/1899 03:19:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
235964022	11/5/2021 0:00	12/30/1899 22:10:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
174579221	2/9/2018 0:00	12/30/1899 00:12:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
156521519	9/7/2016 0:00	12/30/1899 22:48:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
229540933	6/14/2021 0:00	12/30/1899 22:19:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
85518939	7/1/2012 0:00	12/30/1899 01:30:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
186606593	8/19/2018 0:00	12/30/1899 01:30:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
84043104	4/4/2012 0:00	12/30/1899 11:48:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
74470614	9/5/2010 0:00	12/30/1899 03:36:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	
236950605	11/26/2021 0:00	12/30/1899 02:20:00	BROOKLYN	67	0	NOT AVAILABLE		NOT AVAILABLE	NOT AVAILABLE	

Data Preprocessing

Before visualizing the shooting incident data in the dashboard, several preprocessing steps were undertaken to clean and prepare the dataset for analysis. We have made use of excel and power BI for preprocessing. These steps ensure that the data is accurate, consistent, and suitable for visualization purposes.

Handling Missing Values

The dataset was inspected for missing or null values. In this dataset, missing values were represented as NaN 4in certain columns such as 'PERP_AGE_GROUP', 'PERP_SEX', 'PERP_RACE', 'VIC_AGE_GROUP', 'VIC_SEX', and 'VIC_RACE'. These missing values were replaced with appropriate placeholders or handled as needed based on the context of each column.

Converting Data Types

The 'OCCUR_DATE' and 'OCCUR_TIME' columns were converted from string format to datetime objects to facilitate temporal analysis. This conversion enables easier manipulation and extraction of date and time components for visualization purposes.

Data Transformation

- **Date and Time Extraction**: Date and time components were extracted from the 'OCCUR_DATE' and 'OCCUR_TIME' columns, respectively. This transformation allows for granular analysis of shooting incidents based on temporal factors such as hour of the day or day of the week.
- **Feature Engineering**: Created two new columns one which was whether it was weekend or weekday for doing analysis by that and second was time category which showed what time of the day, most shootings are likely to happen

Dashboard Components

The NYPD Shooting Incident Dashboard comprises several interactive components, each designed to provide users with valuable insights into shooting incidents reported by the NYPD. These components include:

Location Dropdown

Allows users to filter shooting incident data based on different boroughs of New York City.

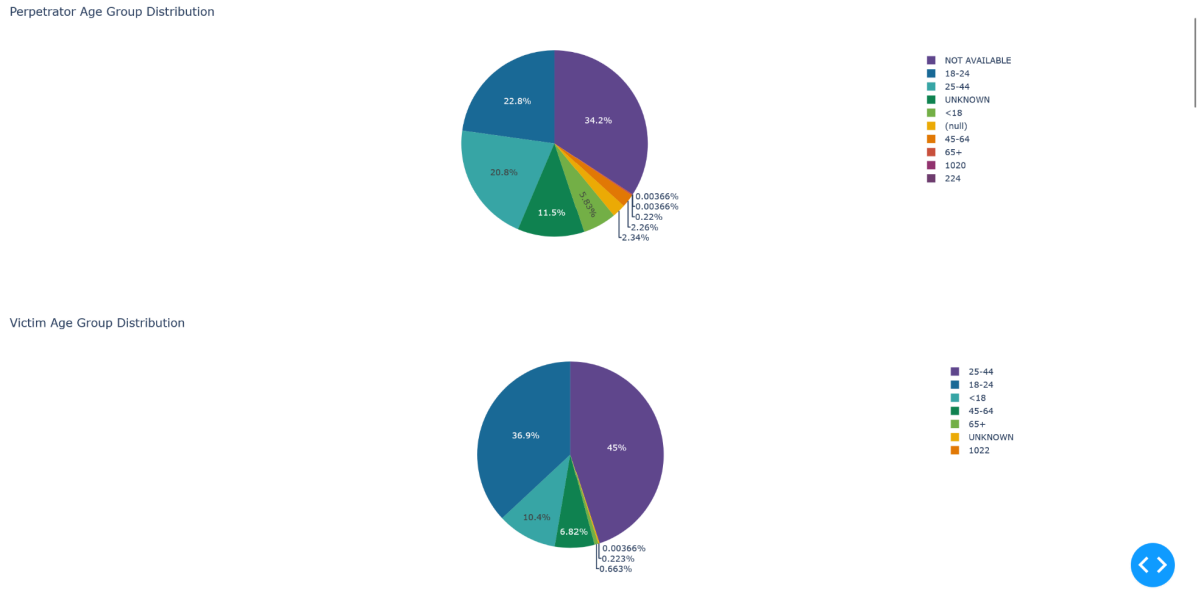
Table Container

Displays a tabular representation of shooting incident data based on the selected location, providing details such as the date, time, location, and other relevant information about each incident.

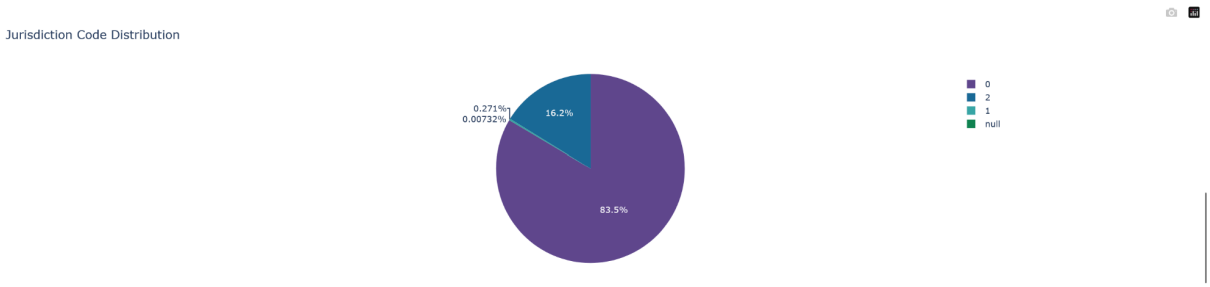
Pie Charts

Perpetrator Age Pie Chart: Visualizes the distribution of perpetrator age groups involved in shooting incidents.
Victim Age Pie Chart: Illustrates the distribution of victim age groups in shooting incidents.

As you can see below the age group who was mostly involved in shooting were 18-24 years and people of that similar age group where the victims of shooting

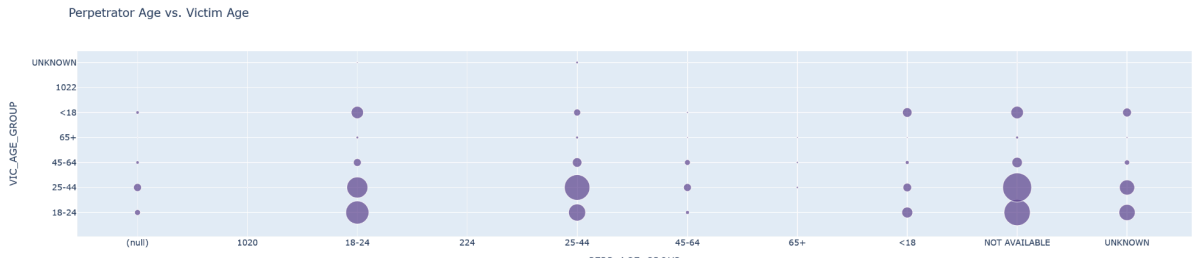


Jurisdiction Donut Chart: Displays the distribution of jurisdiction codes associated with shooting incidents.



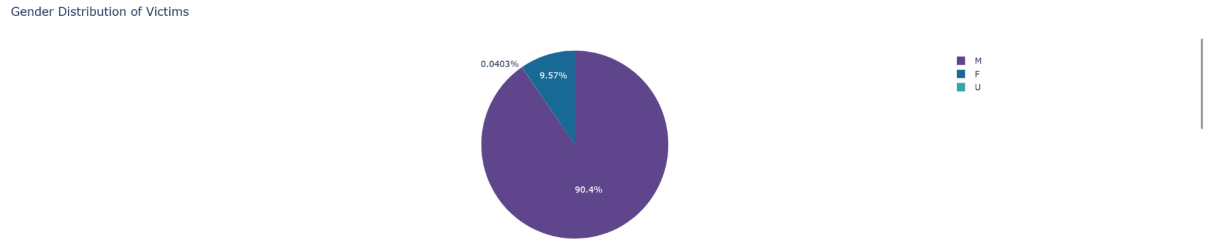
Bubble Chart

The Perpetrator-Victim Bubble Chart depicts the relationship between perpetrator age groups and victim age groups in shooting incidents, with the size of each bubble representing the frequency of incidents involving specific age group combinations.



Gender Pie Chart

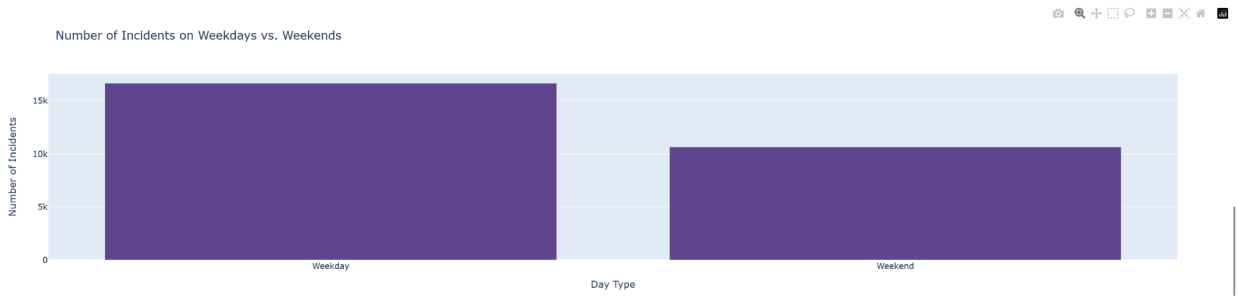
Presents the gender distribution of victims involved in shooting incidents.



As we can see above, the majority of victims were males

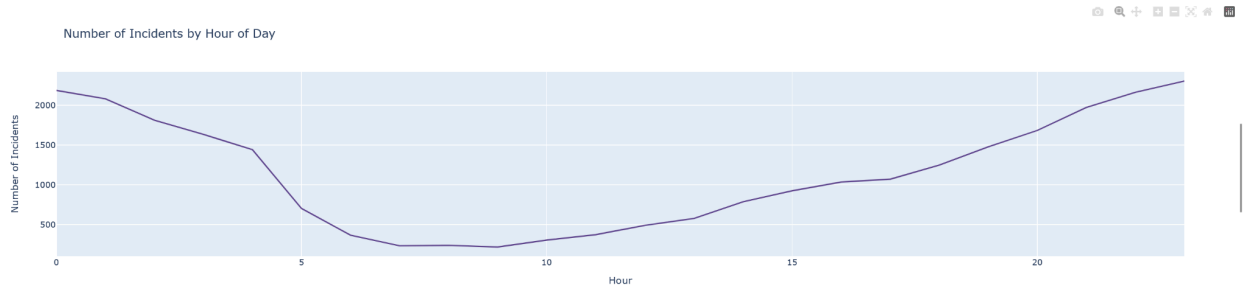
Bar Chart

Weekday-Weekend Bar Chart: Shows the number of shooting incidents that occurred on weekdays versus weekends.



So as we can see, the maximum number of shootings happens on weekday

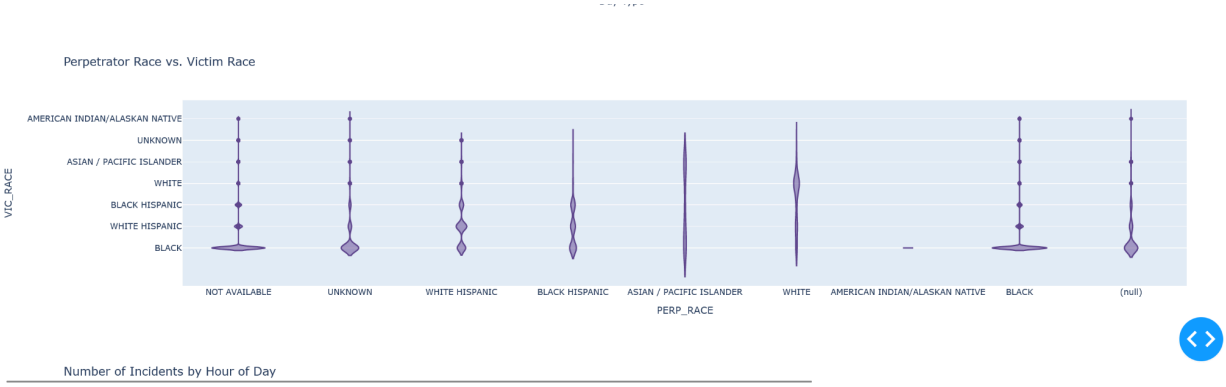
Hourly Line Chart: Visualizes the distribution of shooting incidents by hour of the day.



As we can see in the image, the timing reported of most of the shootings is in midnight

Violin Plot

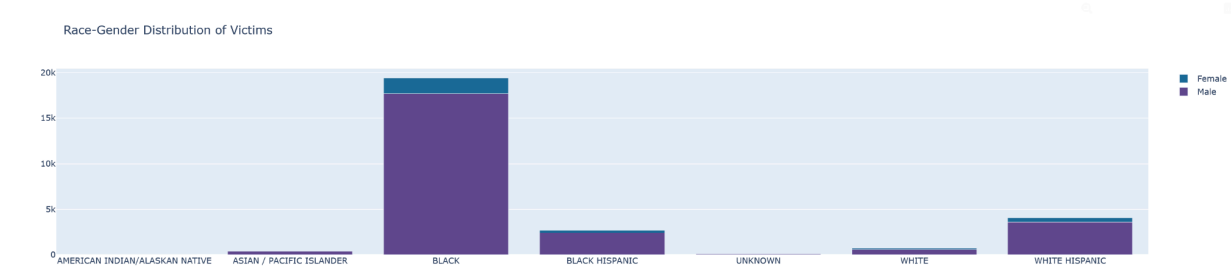
The Race Violin Plot illustrates the distribution of perpetrator race versus victim race in shooting incidents.



As we can see from the data that most of the victims are black and the perpetrator race is also black

Stacked Bar Chart

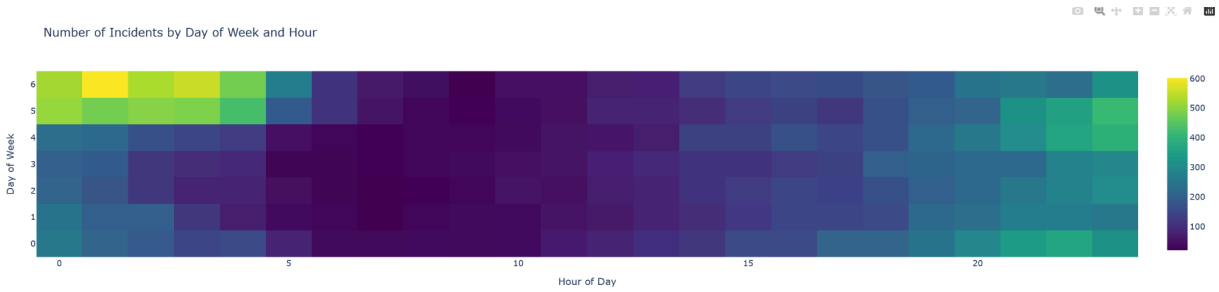
The Race-Gender Stacked Bar Chart provides insights into the distribution of victims based on race and gender.



First most victims are of race black, second most victims race is white hispanic as seen

Heatmap

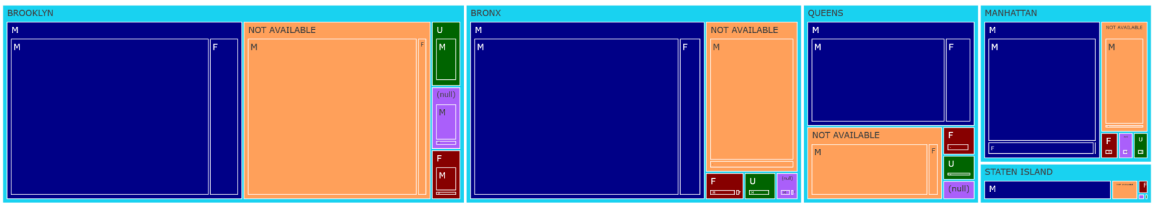
The Day-Hour Heatmap displays the number of shooting incidents based on the day of the week and hour of the day.



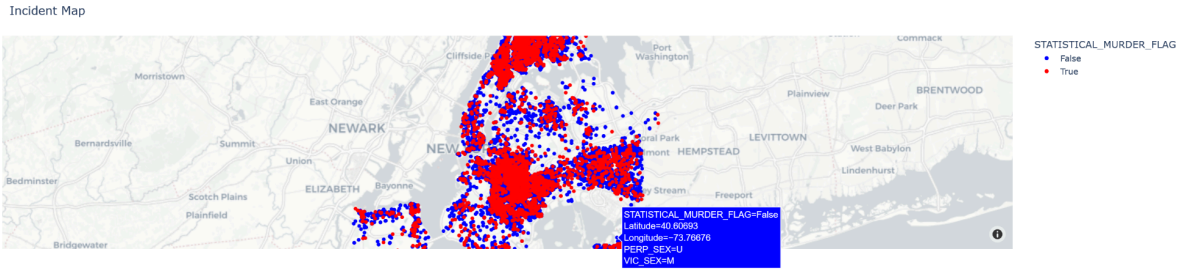
As we can see clearly in the heatmap and mentioned before, most incidents happen at 1 am(near to midnight)

Treemap and Incident Map

- **Treemap**: Presents a hierarchical view of shooting incidents based on borough, perpetrator sex, and victim sex.



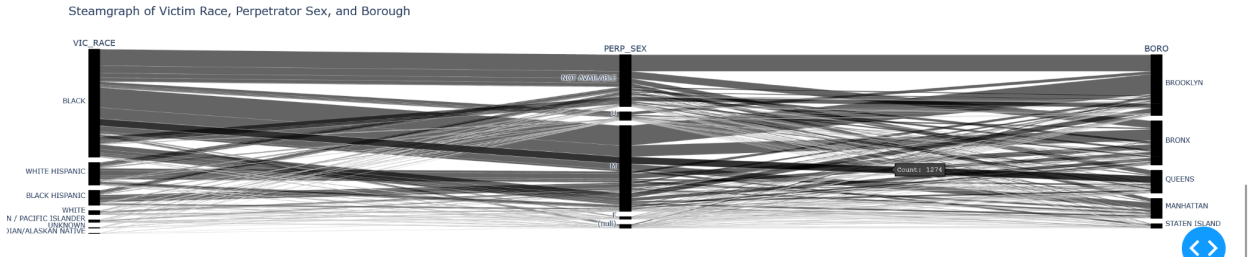
- **Incident Map**: Visualizes the geographic distribution of shooting incidents across New York City.



As we can see here, the red dots represent that the person died and the blue represent the person did not die

Steamgraph

The Steamgraph offers a graphical representation of the relationship between victim race, perpetrator sex, and borough in shooting incidents.



Conclusion

The NYPD Shooting Incident Dashboard provides valuable insights into the characteristics and patterns of shooting incidents in New York City. By leveraging interactive visualizations and filtering options, users can gain a deeper understanding of the data and make informed decisions regarding crime prevention and law enforcement strategies.

Future Enhancement

- Integration of real-time data updates.
- Addition of predictive analytics features to forecast future trends.
- Incorporation of advanced geospatial analysis for hotspot identification.