

NYPD Shooting Incidents Report

Introduction

This report analyzes shooting incidents in New York City based on historical NYPD data from the provided CSV file. The focus is on visualizing incident trends over time and examining victim age group distribution.

Importing Data

```
library(tidyverse)
```

```
url <- "https://data.cityofnewyork.us/api/views/833y-fsy8/rows.csv?accessType=DOWNLOAD"
```

Processing Data

The dataset contains information on shooting incidents, including but not limited to:

- Date and time of occurrence
- Borough and location details
- Victim demographics (age, race, sex)
- Whether the incident resulted in a fatality

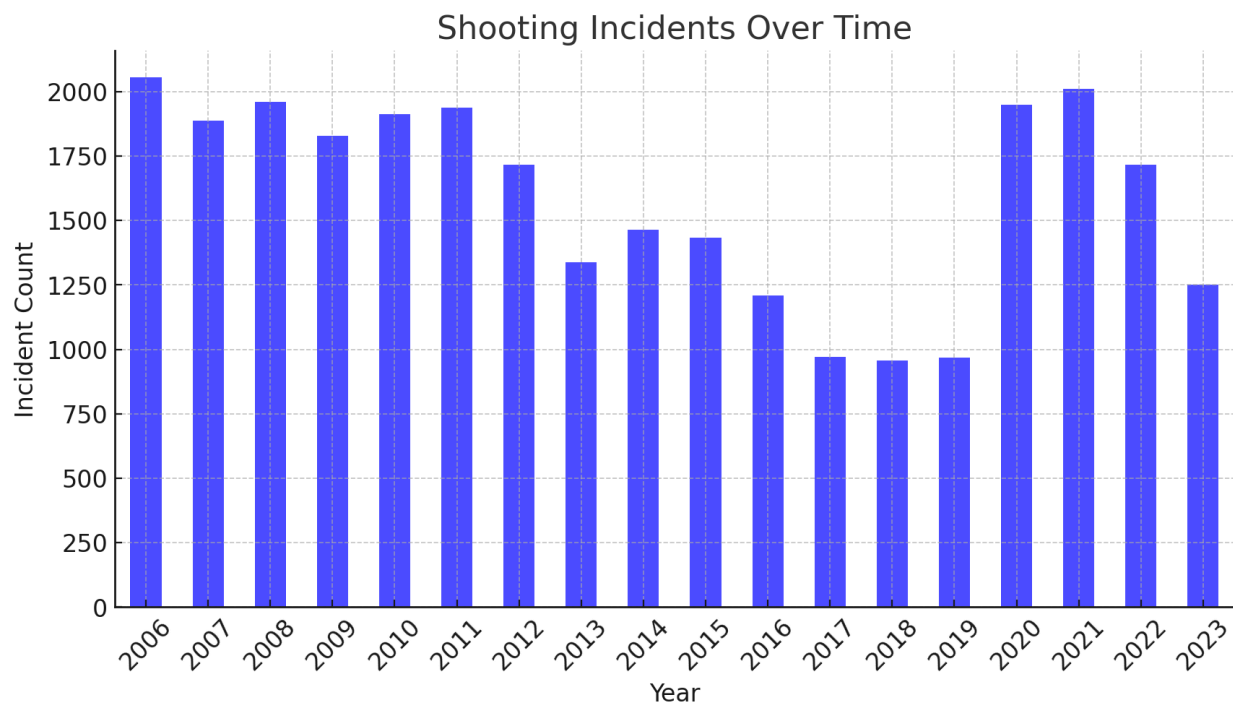
Screenshot of a relevant section in R code:

```
9 ## Incidents Over Time
10 ```{r}
11 data$OCCUR_DATE <- as.Date(data$OCCUR_DATE, format="%m/%d/%Y")
12 ggplot(data, aes(x = OCCUR_DATE)) +
13   geom_histogram(binwidth = 365, fill="blue", alpha=0.7) +
14   labs(title="Shooting Incidents Over Time", x="Year", y="Count")
15 ```
16
17 ## Borough-wise Distribution
18 ```{r}
19 ggplot(data, aes(x = BORO)) +
20   geom_bar(fill="red", alpha=0.7) +
21   labs(title="Shooting Incidents by Borough", x="Borough", y="Count")
22 ```
23
24 ## Victim Age Group Distribution
25 ```{r}
26 ggplot(data, aes(x = VIC_AGE_GROUP)) +
27   geom_bar(fill="green", alpha=0.7) +
28   labs(title="Victim Age Group Distribution", x="Age Group", y="Count") +
29   theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Visualization and Analysis: Shooting Incidents Over Time

A histogram of shooting incidents over time shows fluctuations in crime trends. The analysis indicates:

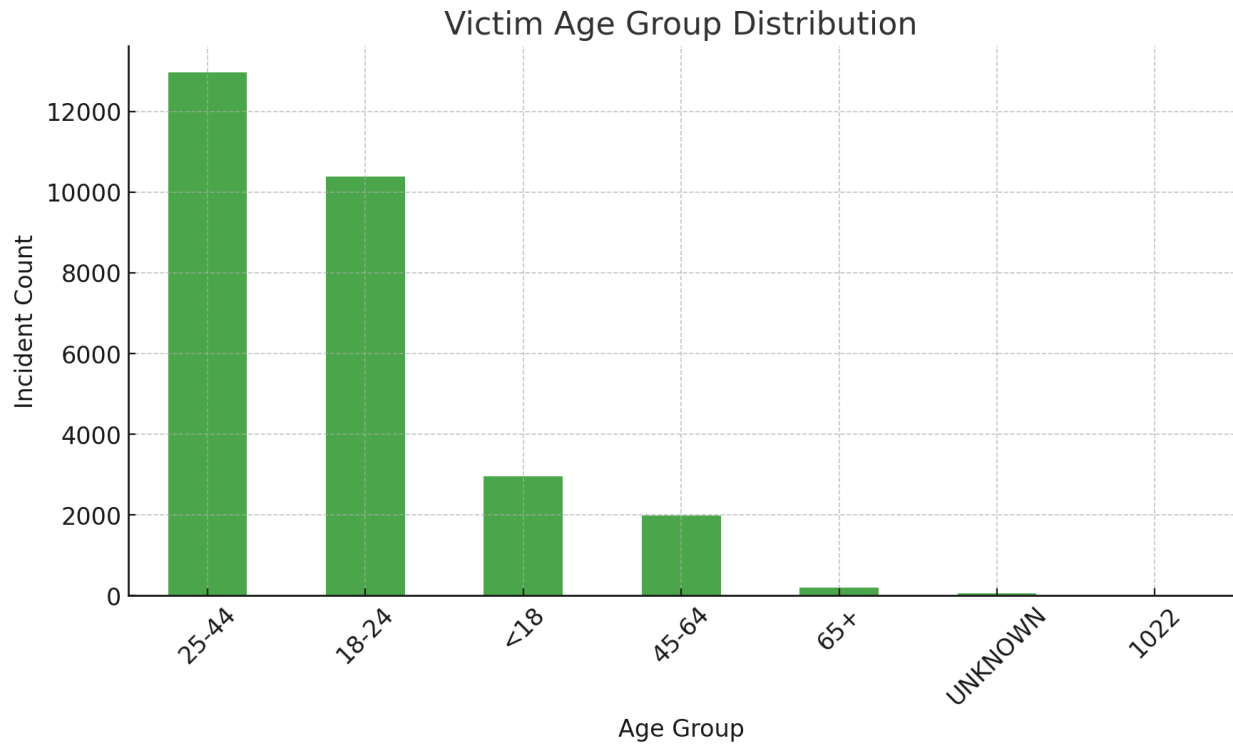
- A peak in incidents during certain years. The timing coincides with Covid pandemic start, further analysis is needed to confirm their relation.
- A declining trend in recent years, possibly due to policy changes or law enforcement strategies.
- Seasonal patterns where incidents increase in specific months may not be fully represented in this yearly graph



Visualization and Analysis: Victim Age Group Distribution

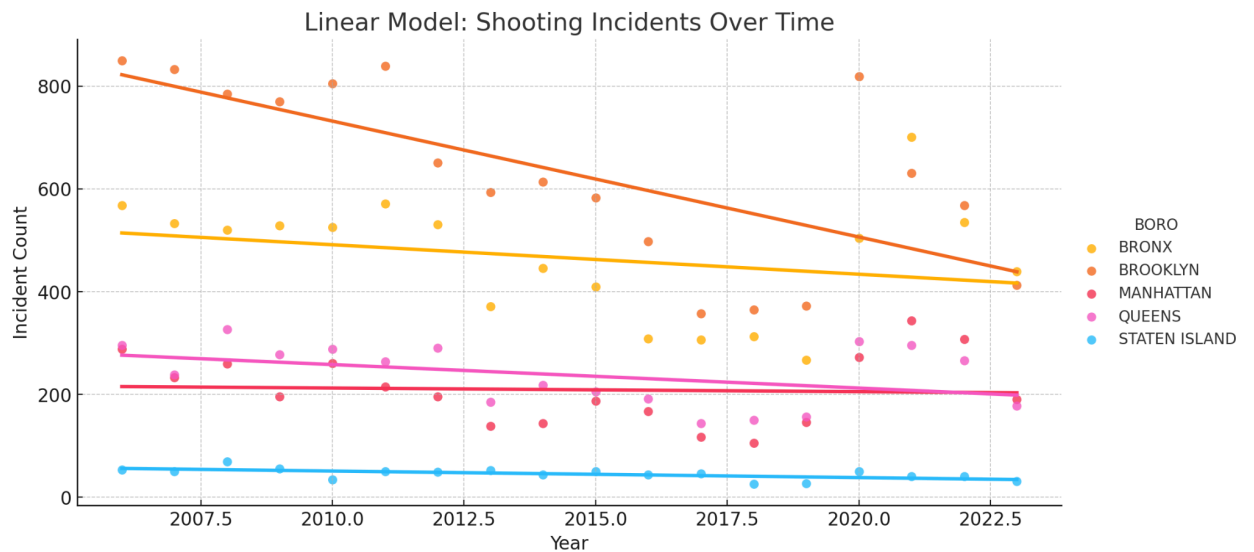
Analysis of victim age distribution highlights:

- The majority of victims belong to the 18-24 and 25-44 age groups.
- A smaller proportion of victims are under 18 or over 45.
- Age-related trends in victimization tend to correlate with socioeconomic factors and high-crime areas. Further analysis may be needed to prove this theory.



Model and Analysis: Shooting Incidents Over Time

- The linear model suggests a trend in shooting incidents over time across different boroughs. Some boroughs exhibit an increasing trend in shooting incidents, while others show a decline.



```

9
10 # Aggregate Shooting Incidents by Year and Borough
11 incidents_by_year <- data %>%
12   mutate(year = year(OCCUR_DATE)) %>%
13   group_by(year, BORO) %>%
14   summarise(incident_count = n(), .groups='drop')
15
16 # Encode Borough as Factor
17 incidents_by_year$BORO <- as.factor(incidents_by_year$BORO)
18
19 # Fit Linear Model
20 lm_model <- lm(incident_count ~ year + BORO, data = incidents_by_year)
21
22 # Summary of the Model
23 summary(lm_model)
24
25 # Plot Predictions
26 ggplot(incidents_by_year, aes(x = year, y = incident_count, color = BORO)) +
27   geom_point() +
28   geom_smooth(method = "lm", se = FALSE) +
29   labs(title="Linear Model: Shooting Incidents Over Time", x="Year", y="Incident Count")

```

Bias Identification

Addressing bias requires transparency in data collection and model development, ensuring fair representation and interpretation. Some possible bias in the data and analysis:

- Reporting Bias: Certain crimes may be underreported in specific communities.
- Sampling Bias: Data may disproportionately represent certain demographic groups.
- Law Enforcement Practices: Differences in policing strategies across boroughs could influence recorded incidents.