# NYPD Shooting Incidents

A report on shooting incident data provided by the NYPD.

Robert Forrest
Apr 14th 2024

# Overview

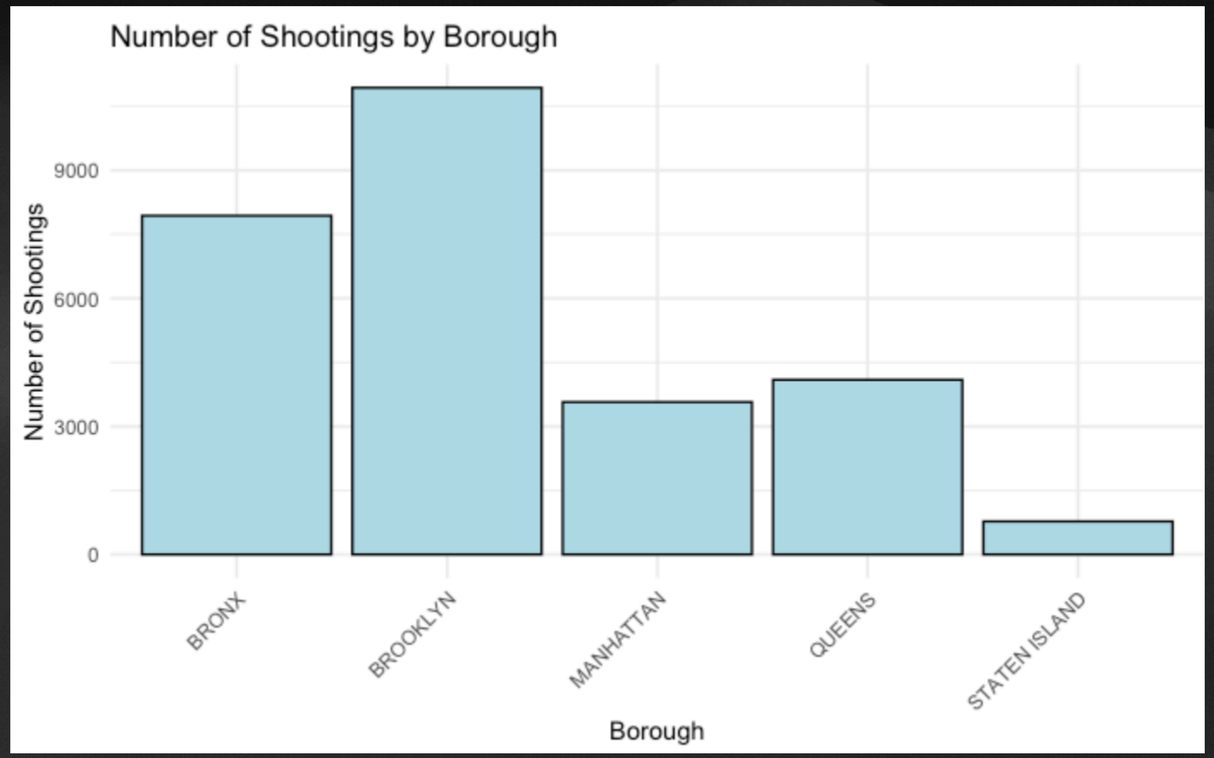
- Data hosted on the city of New York website.
- Contains data on shooting incidents from the different boroughs in New York.
- The purpose is to understand the trends in these shootings.

# Data Processing

- Made sure columns are the correct data types.
- Handled unknown and missing values.
- Created additional columns of interest.

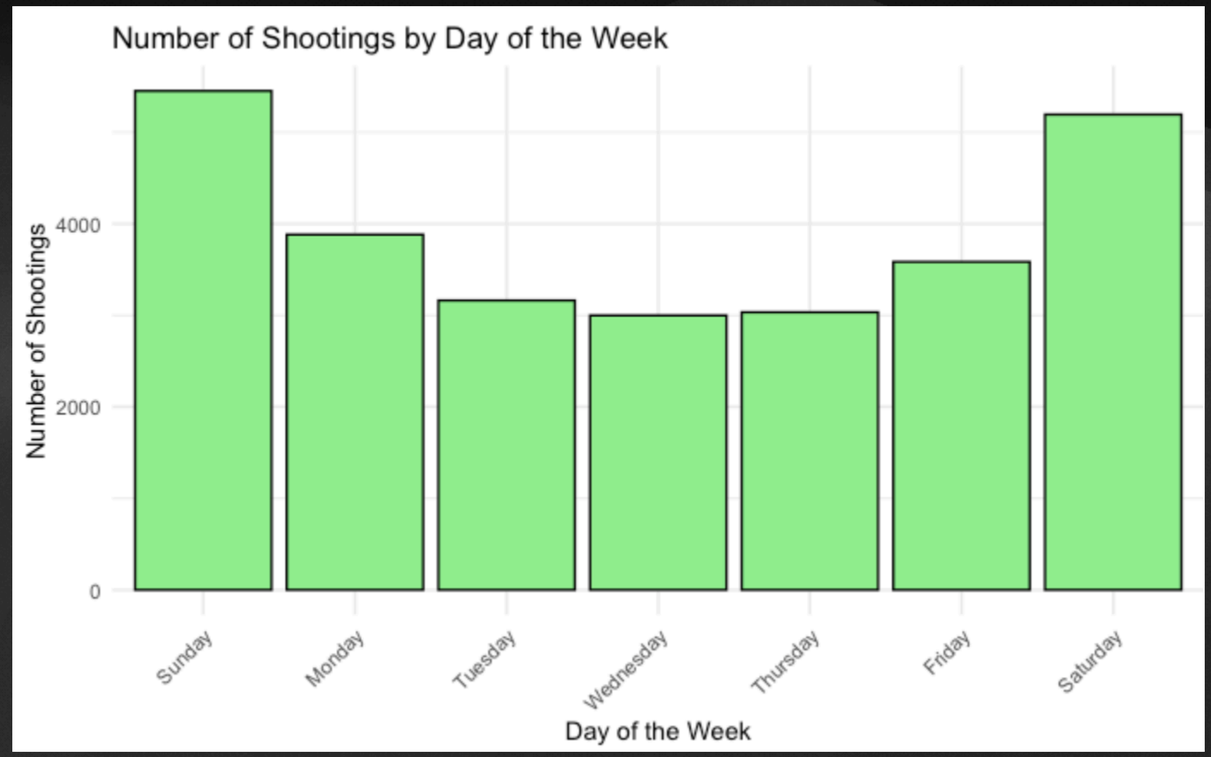
# Visualize Shootings by Borough

- Brooklyn has the highest amount of shootings while Staten Island has the lowest.



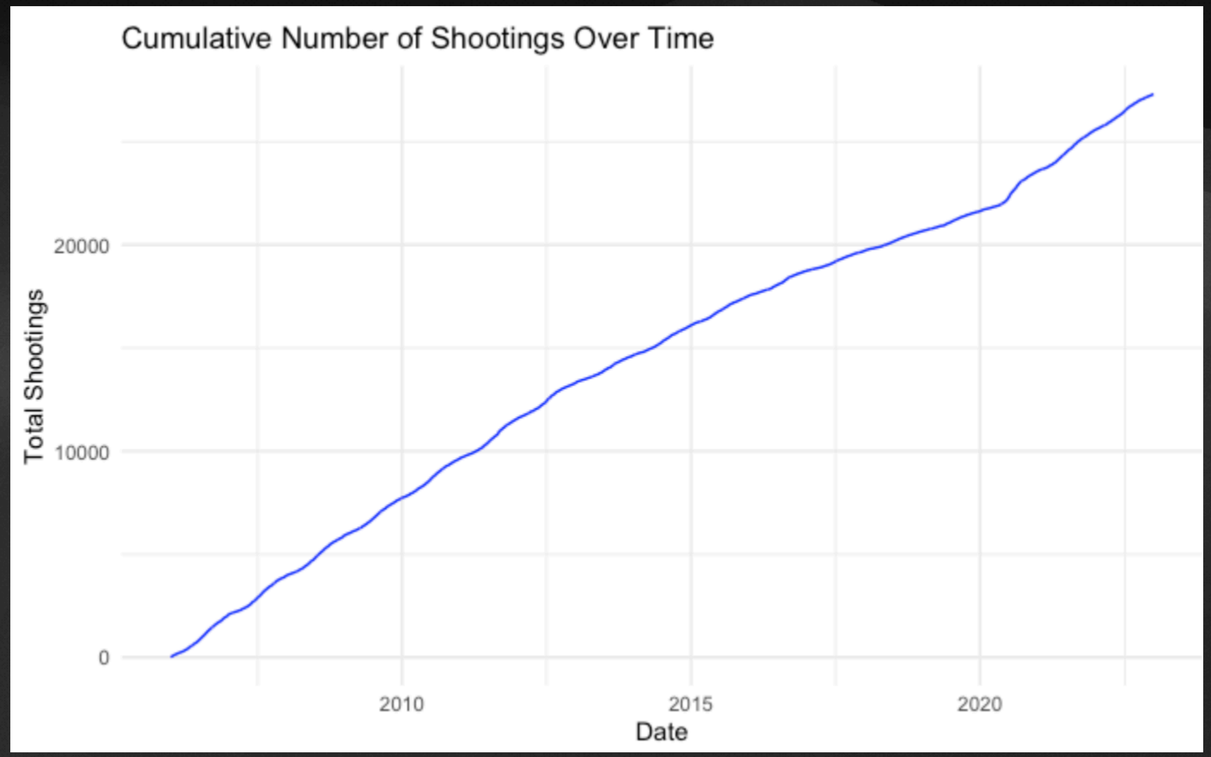
# Visualize Shootings by Day of Week

- Shootings typically occur on Saturday and Sunday.



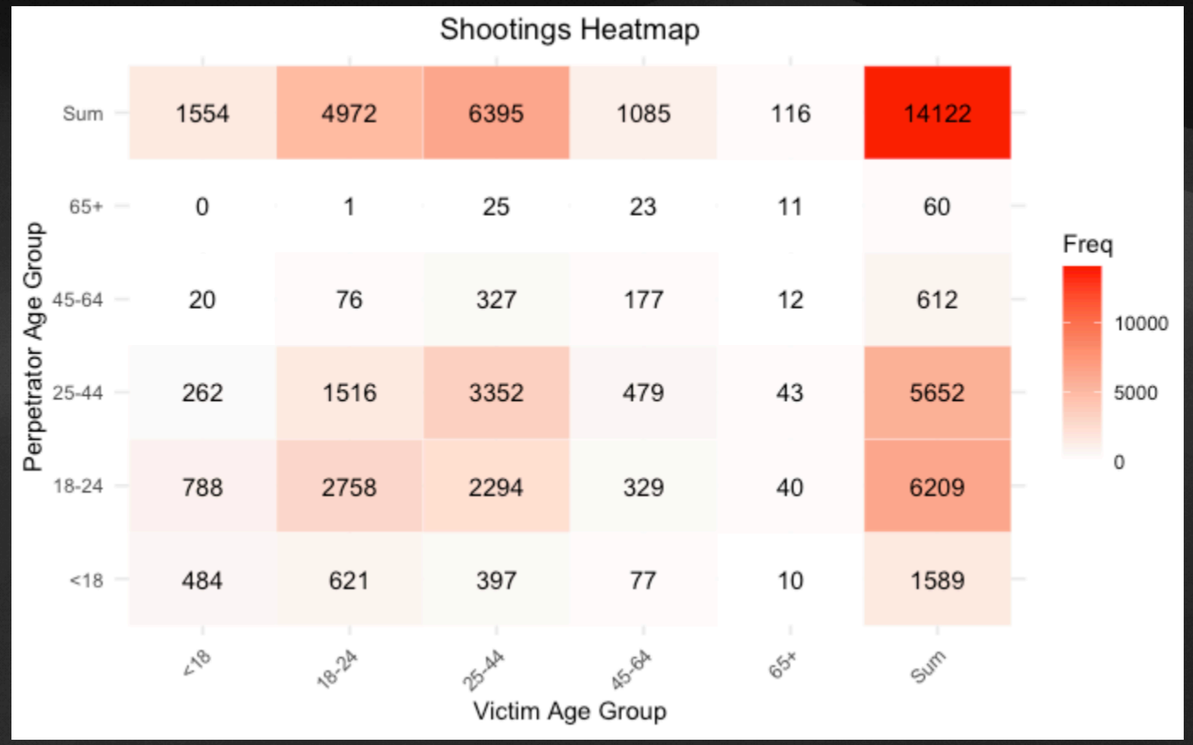
# Visualize Cumulative Shootings Over Time

- The rate of shootings appear to be constant over time, with a moment of increase around 2020.



# Visualize Ages of Victims and Perpetrators

- The 18-24 age group caused the most shootings and the 25-44 age group were victimized most often.



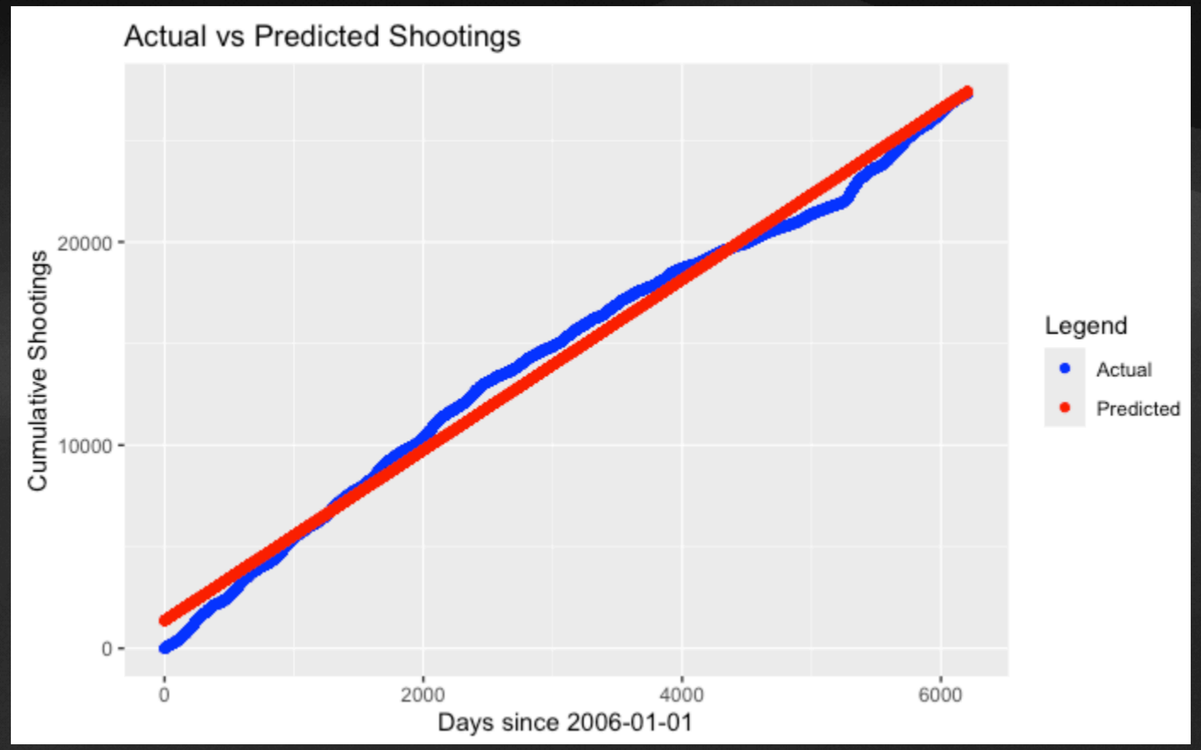
# Modeling Cumulative Shootings

```
Call:
lm(formula = CUMULATIVE_SHOOTINGS ~ DAYS_SINCE_DATA_START, data = event_counts_for_model %>%
    filter(PERP_AGE_GROUP != "UNKNOWN" & VIC_AGE_GROUP != "UNKNOWN"))

Residuals:
     Min       1Q   Median       3Q      Max
-1410.30  -629.52   -75.33   717.14  1291.99

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)          1.360e+03  1.282e+01     106   <2e-16 ***
DAYS_SINCE_DATA_START 4.195e+00  3.684e-03    1139   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 766.7 on 12020 degrees of freedom
Multiple R-squared:  0.9908,    Adjusted R-squared:  0.9908
F-statistic: 1.297e+06 on 1 and 12020 DF,  p-value: < 2.2e-16
```



# Conclusions

- Shootings noticeably vary by borough and day of the week.
- Younger people are more likely to be willing or unwilling participants in shootings.
- Data only originating from NYPD, introduces bias.