NYPD SHOOTING ANALYSIS

CU Boulder - Data Science As A Field

Cassandra Jones

4/25/2025

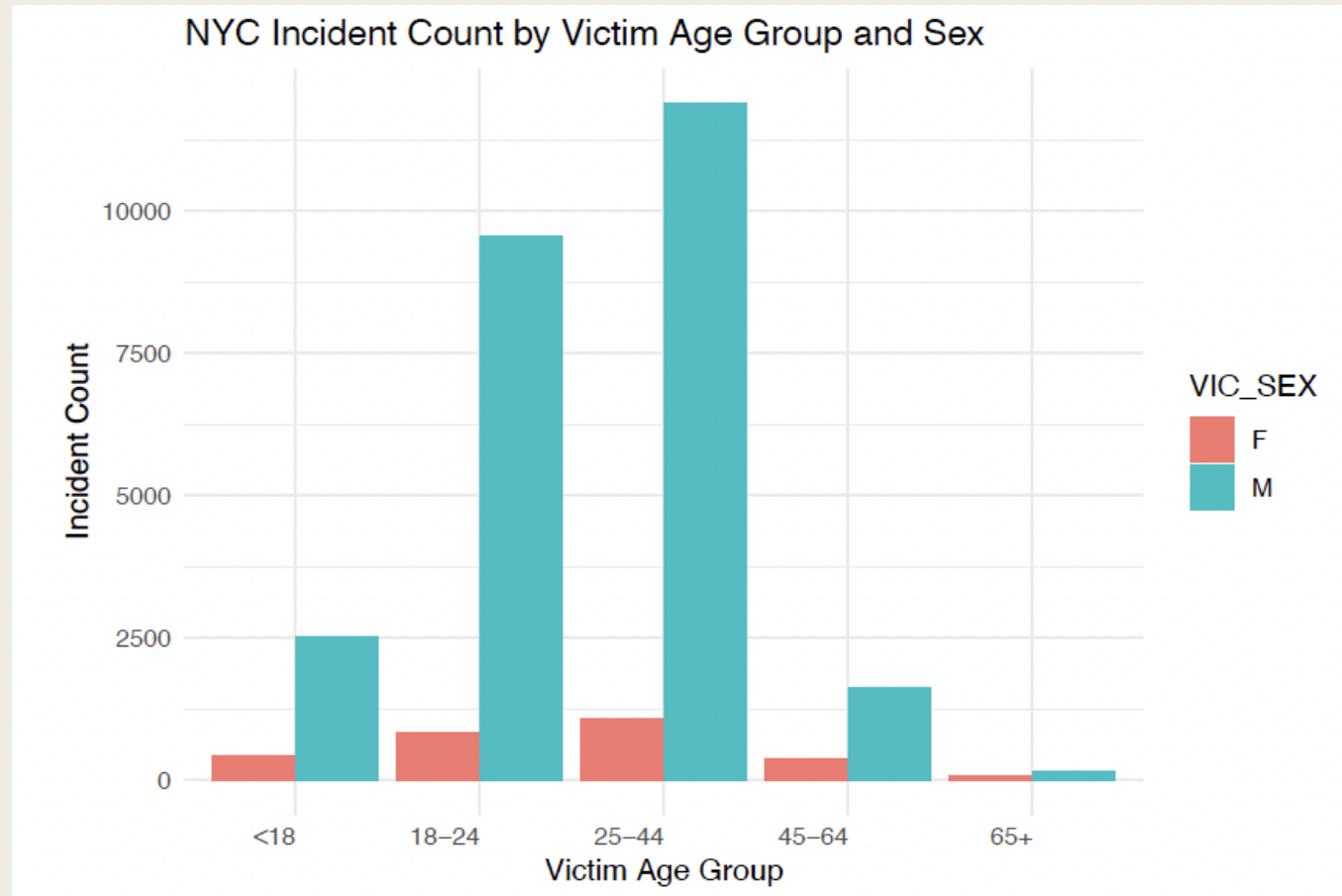## Data Source

Shooting incidents occurred in New York City during the year 2006.

https://catalog.data.gov/dataset/nypd-shooting-incident-data-historic
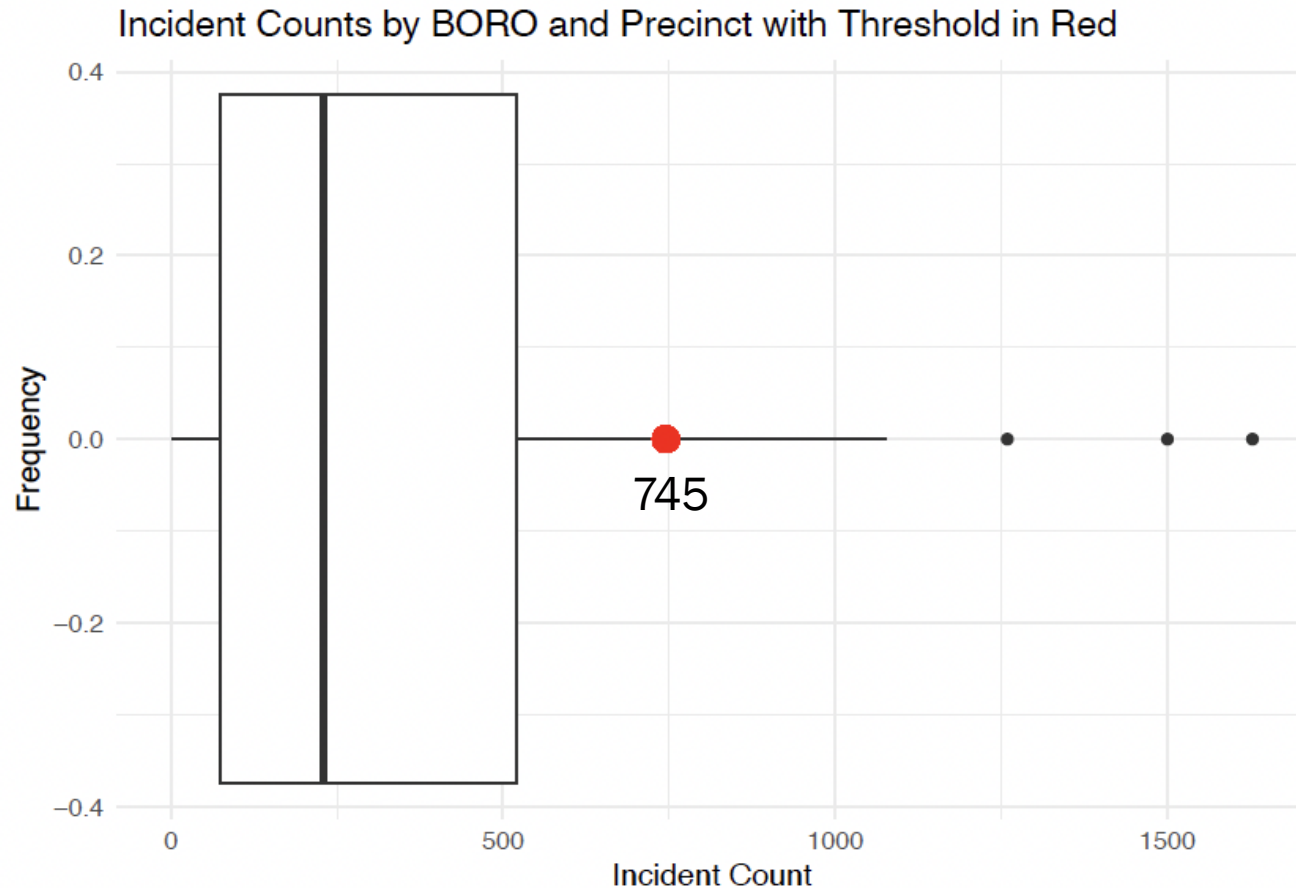
Goal

To explore the characteristics the the incidents and create a predictive model to determine whether an incident was a murder.

- Male victims outnumbered female victims across all age groups.
- With18-24 and 25-44 age groups experiencing the highest number of incidents

NYC Incident Count by Victim Age Group and Sex

Incident Counts by BORO and Precinct with Threshold in Red

745

- mean number of incidents to be 370.93, with a standard deviation of 374.15

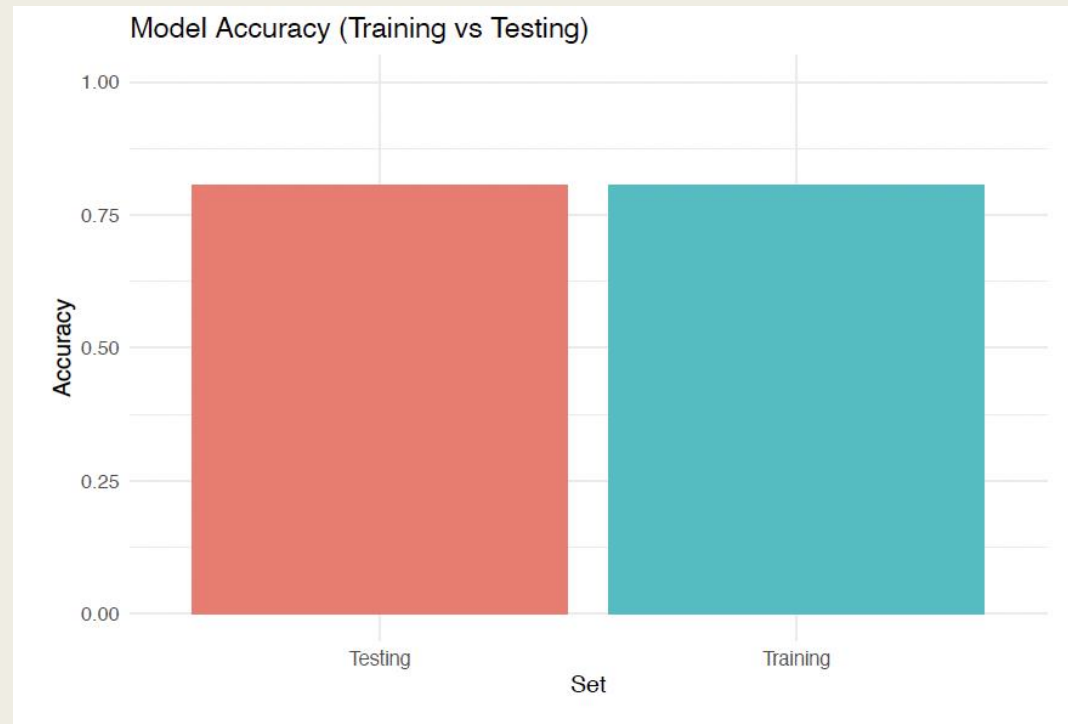- BORO and precinct with Incidents above 745 (370.93 + 374.15) can be identified as potential unsafe place

Potential Bias: Only considered incident counts without factoring in the population size

Logistic Model – to predict whether an incident was a murder

Variables: BORO, Precinct, victim age group, victim race, victim sex, latitude, longitude, month, hour, day of week

Training score: 0.8063
Test score: 0.8062

THANK YOU FOR REVIEWING!