NYPD

Shooting Analysis

Project Agenda

1. Introduction
2. Data Preprocessing
3. Time Trend Analysis
4. Geographic Analysis
5. Profiling Analysis
6. ARIMA Forecasting
7. Conclusion

Introduction

Data Source:

- The dataset contains 27,000+ shooting cases from 2006 to 2022.
- Each record includes location, timestamp, and people profiles.
- This data was manually collected by NYPD with potential bias.

Objective:

- Temporal Patterns Exploration
- Geographic Distribution Analysis
- Involved Individuals Profiling
- Future Cases Forecasting

Project Agenda

1. Project Introduction
2. Data Preprocessing
3. Time Trend Analysis
4. Geographic Analysis
5. Profiling Analysis
6. ARIMA Forecasting
7. Conclusion

Data Preprocessing

1. Selecting Attributes
2. Modifying Column Names
3. Adjusting Format
4. Handling Missing Values
5. Changing Data Types
6. Removing Duplicates

Project Agenda

1. Project Introduction
2. Data Preprocessing
3. Time Trend Analysis
4. Geographic Analysis
5. Profiling Analysis
6. ARIMA Forecasting
7. Conclusion

Exploring Time Distribution

- The highest year had over twice the cases of the lowest year
- Peak months and days were mainly concentrated in 2020

Group	Year	Counts
Highest	2006	2055
Highest	2021	2011
Highest	2008	1959
Highest	2020	1948
Highest	2011	1939
Highest	2010	1912
Average		1606
Lowest	2015	1434
Lowest	2013	1339
Lowest	2016	1208
Lowest	2017	970
Lowest	2019	967
Lowest	2018	958

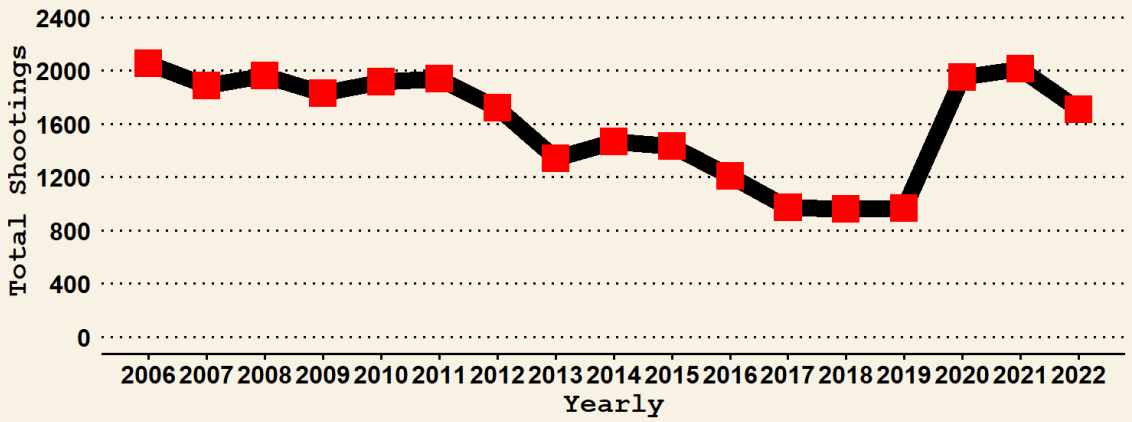
Group	Month	Counts
Highest	2020-7	325
Highest	2020-8	311
Highest	2020-6	265
Highest	2011-9	250
Highest	2012-7	247
Highest	2006-8	245
Average		134
Lowest	2017-3	56
Lowest	2018-3	56
Lowest	2014-2	52
Lowest	2020-2	51
Lowest	2017-2	48
Lowest	2018-2	41

Group	Date	Counts
Highest	07/05/2020	47
Highest	09/04/2011	31
Highest	07/26/2020	29
Highest	08/11/2007	26
Highest	08/27/2022	25
Highest	09/04/2006	25
Average		5
Lowest	12/08/2019	1
Lowest	08/29/2018	1
Lowest	11/25/2013	1
Lowest	04/04/2018	1
Lowest	06/06/2018	1
Lowest	02/06/2009	1

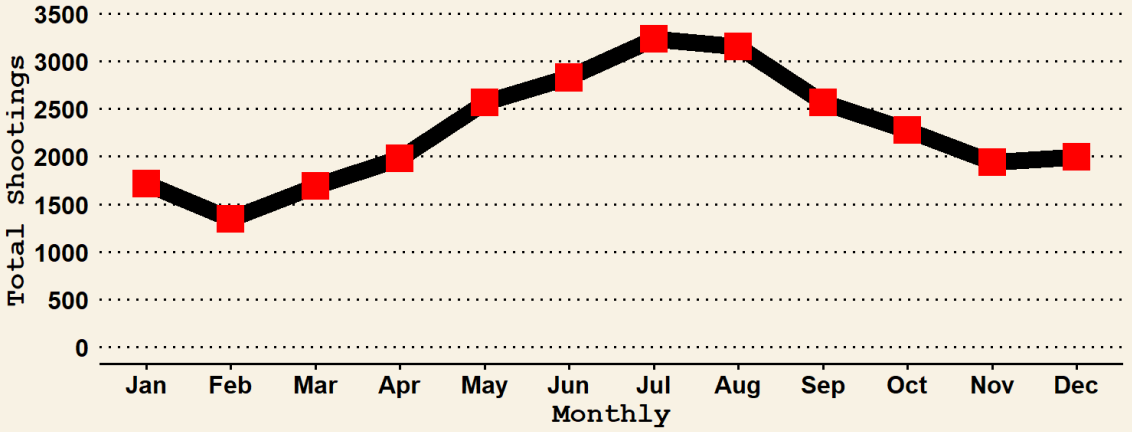
Comparison of Shootings Across Different Time Intervals

- Peaks and valleys are significant across various time intervals
- Summer, weekends, and midnights are obviously higher than others

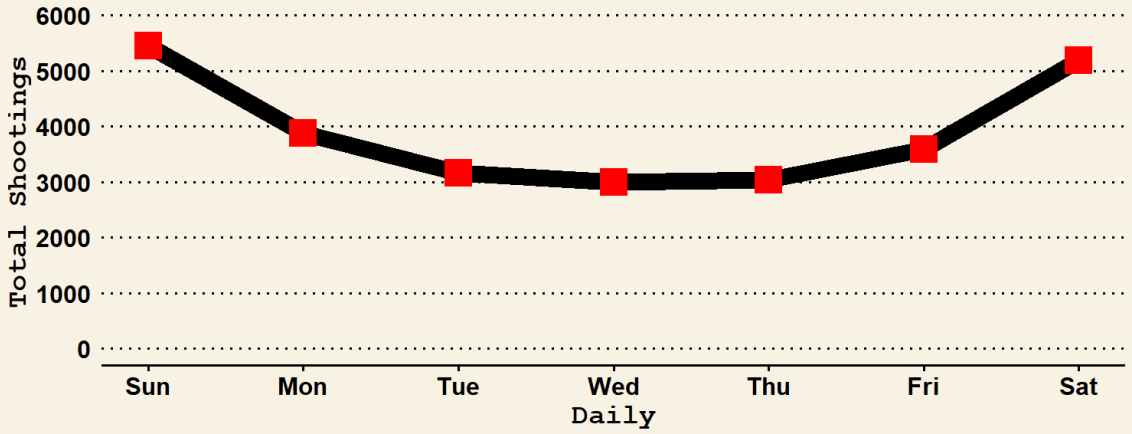
Shootings Across Yearly



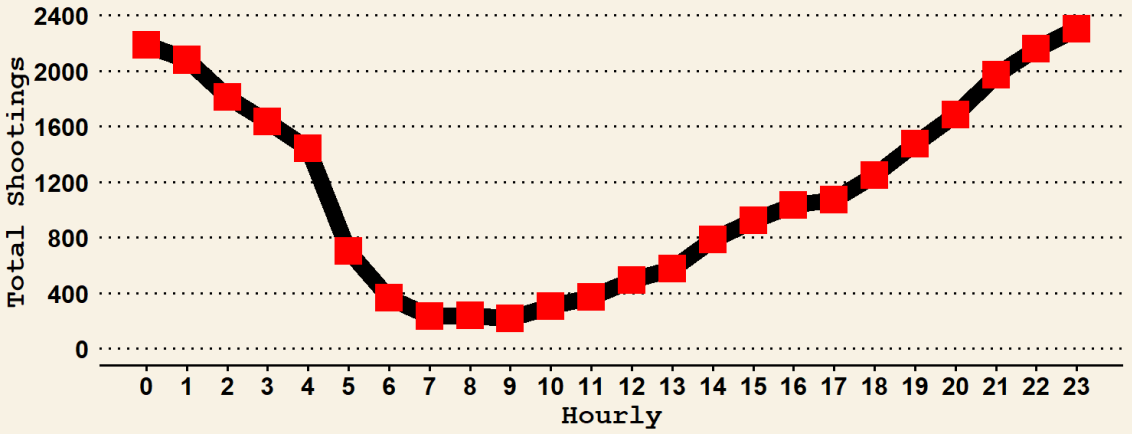
Shootings Across Monthly



Shootings Across Daily



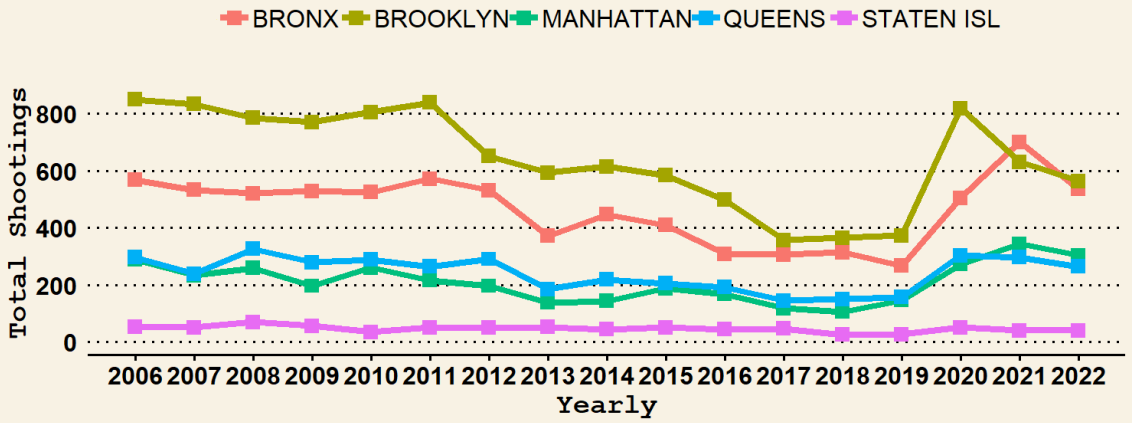
Shootings Across Hourly



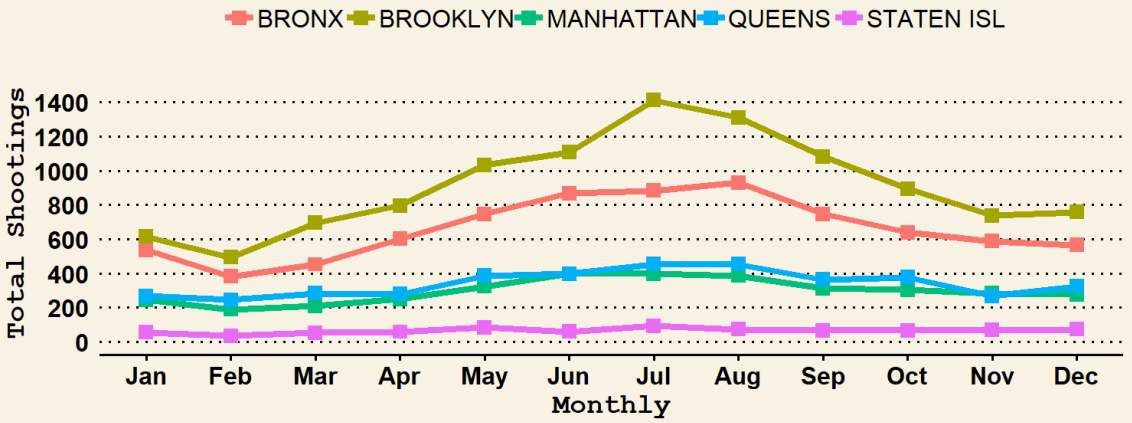
Comparison Across Different Time Intervals by Boroughs

- Boroughs Display Similar Trends Across Different Time Intervals
- Brooklyn and Bronx Show Larger Peaks and Valleys Disparities

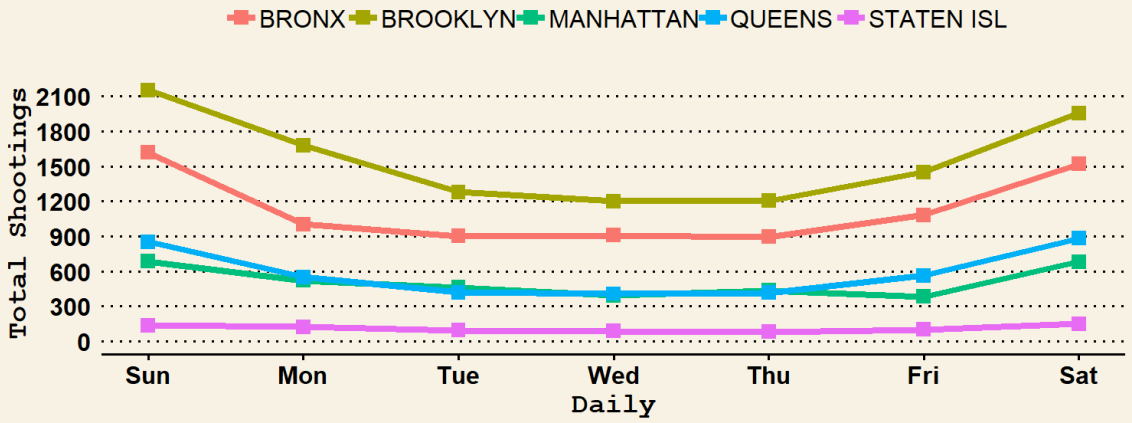
Shootings Across Yearly



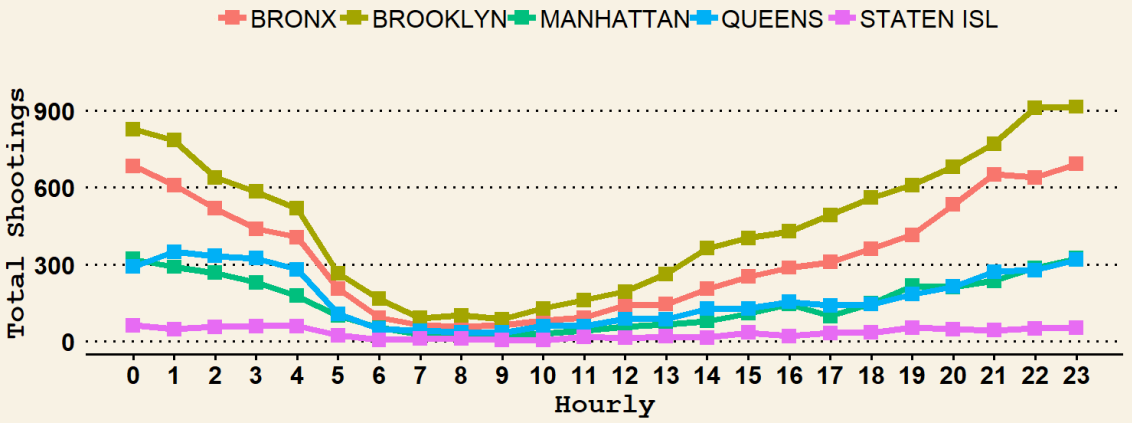
Shootings Across Monthly



Shootings Across Daily



Shootings Across Hourly

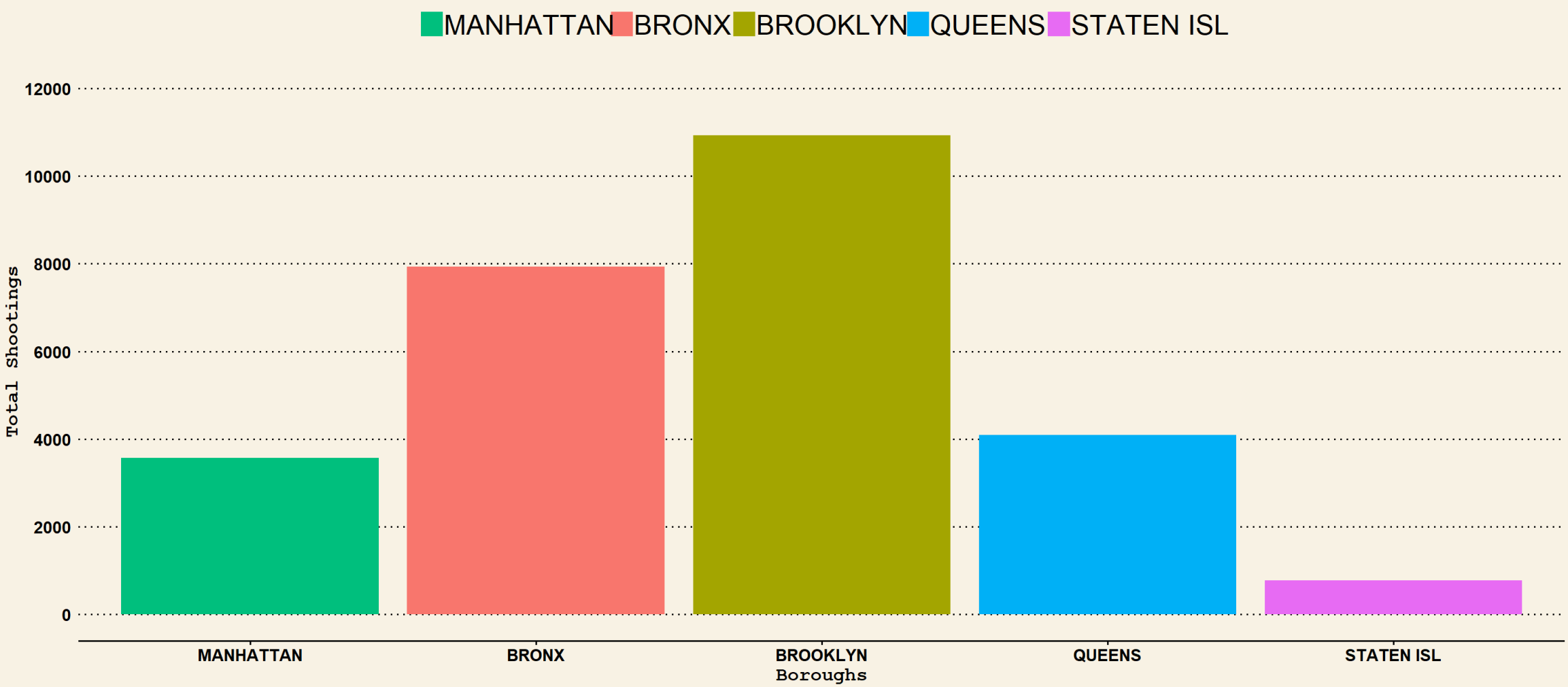


Project Agenda

1. Project Introduction
2. Data Preprocessing
3. Time Trend Analysis
4. Geographic Analysis
5. Profiling Analysis
6. ARIMA Forecasting
7. Conclusion

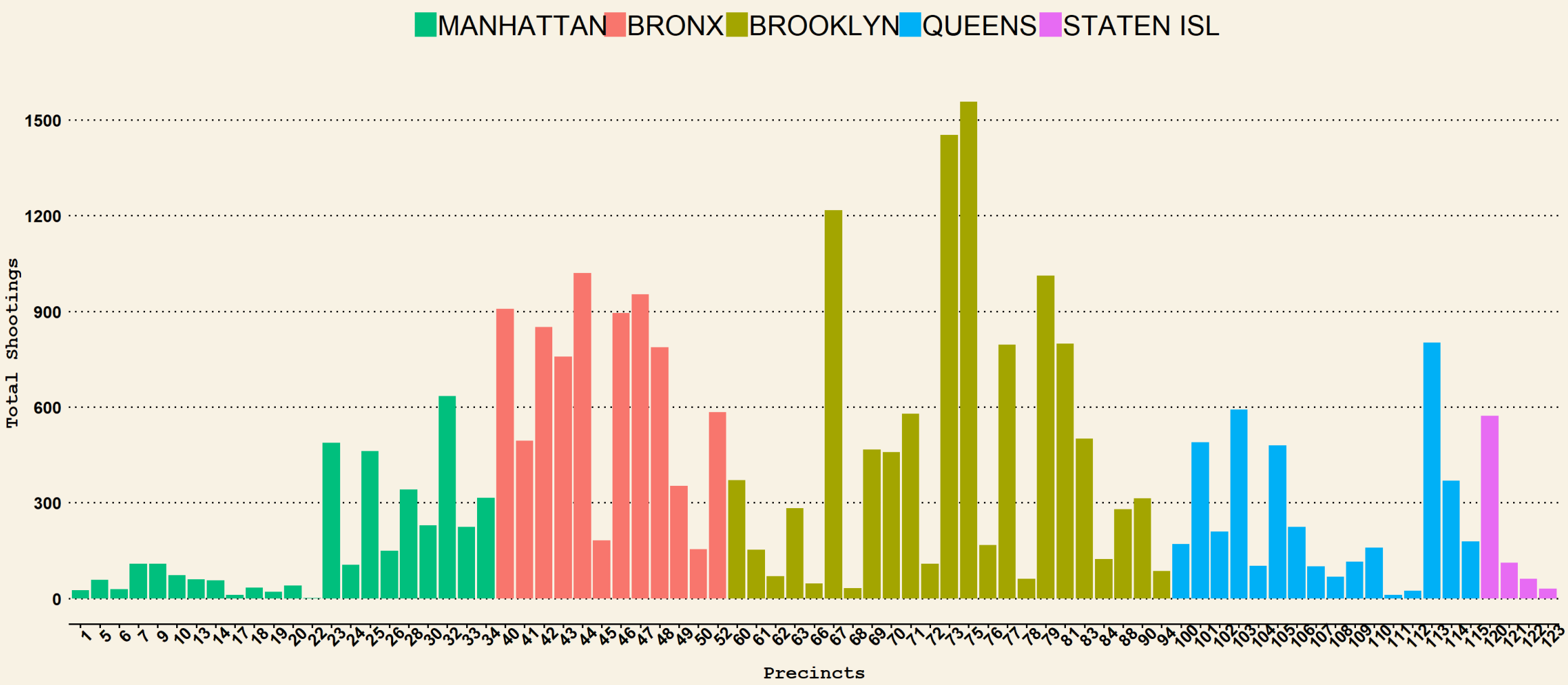
Comparison of Shootings Across Boroughs (2006~2022)

- Brooklyn: Nearly 11,000 shootings, over 40% of total
- Bronx: Close to 8,000 incidents, about 30% of total



Comparison of Shootings Across Precincts (2006~2022)

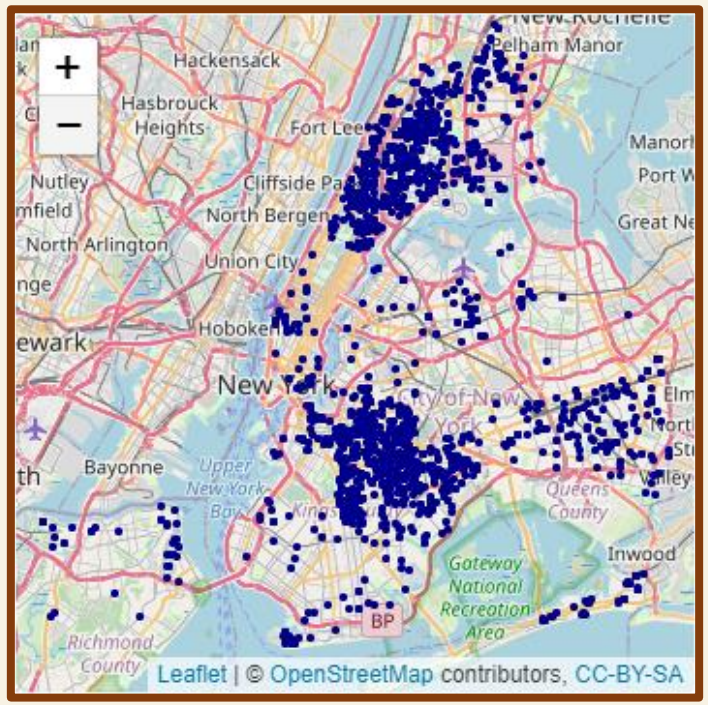
- Brooklyn precincts (73' and 75') lead in shooting numbers
- Bronx precincts consistently show high counts



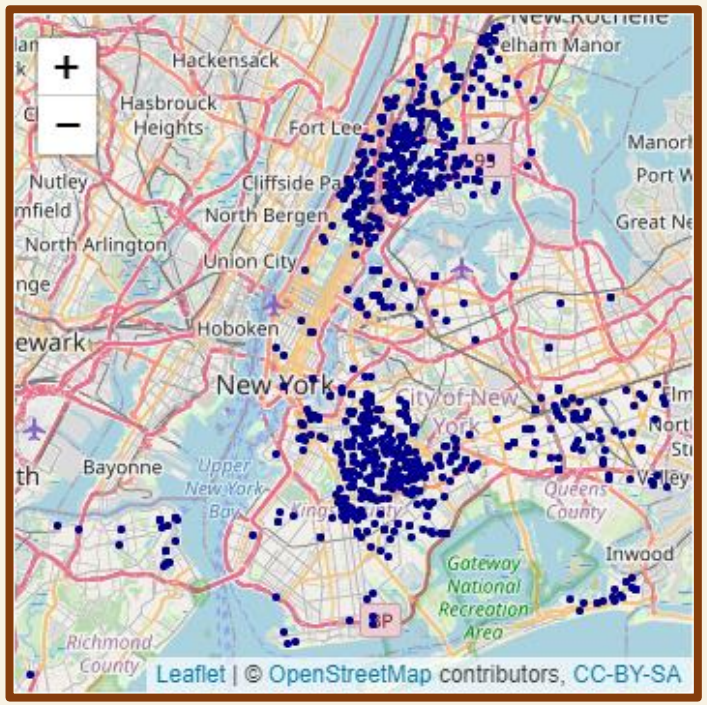
Comparing the shooting maps of 2006, 2019, and 2022

- Shootings in 2019 were notably sparser compared to 2022
- The Bronx's shooting distribution worsened significantly in 2022

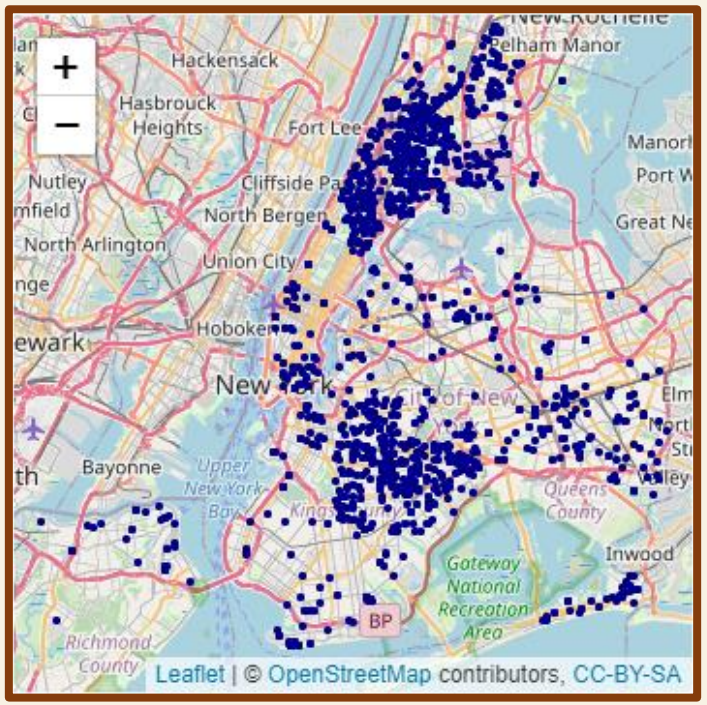
Year 2006



Year 2019



Year 2022

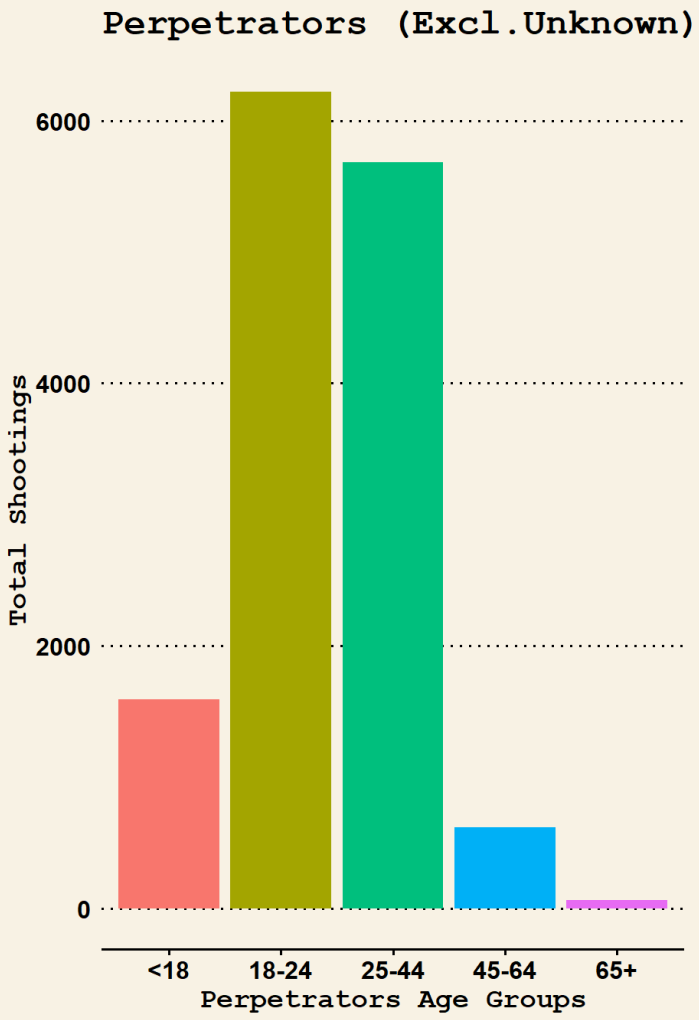


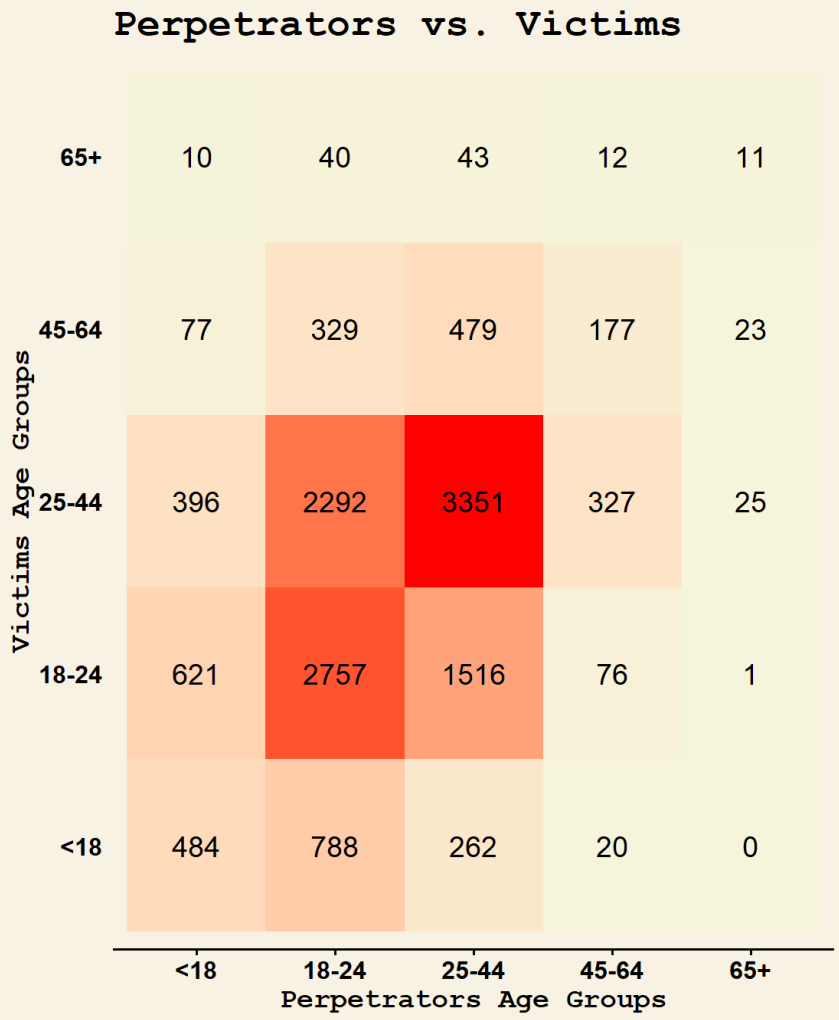
Project Agenda

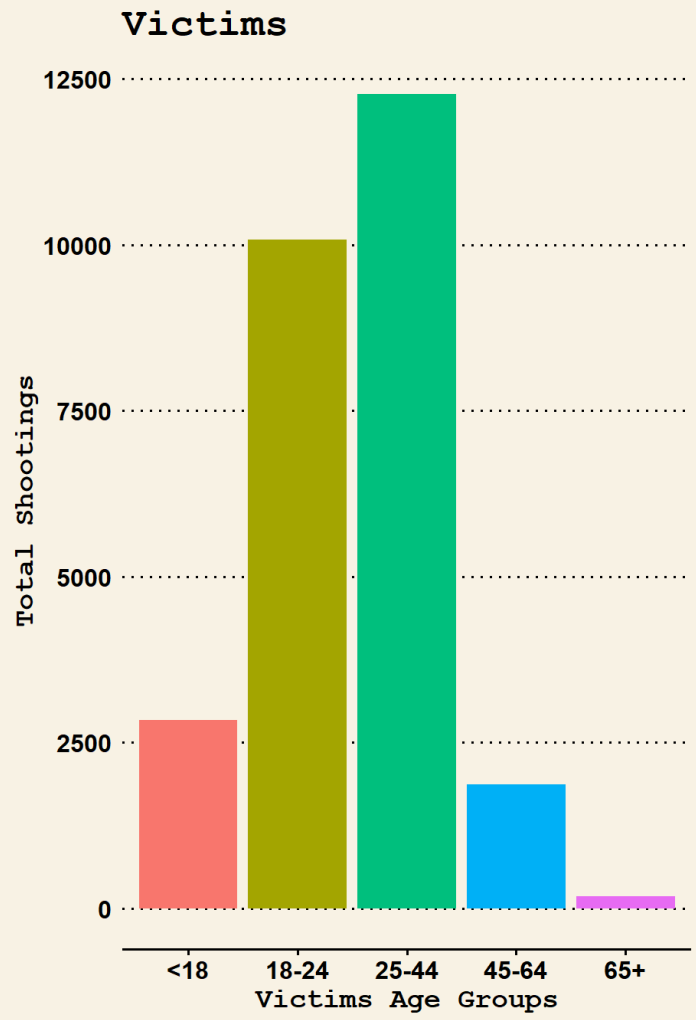
1. Project Introduction
2. Data Preprocessing
3. Time Trend Analysis
4. Geographic Analysis
5. Profiling Analysis
6. ARIMA Forecasting
7. Conclusion

Comparison of Perpetrators and Victims Across Age Groups

- '18-24' and '25-44' groups contribute to 70% of shootings
- Age group <18 is involved in more shootings than the >45 groups

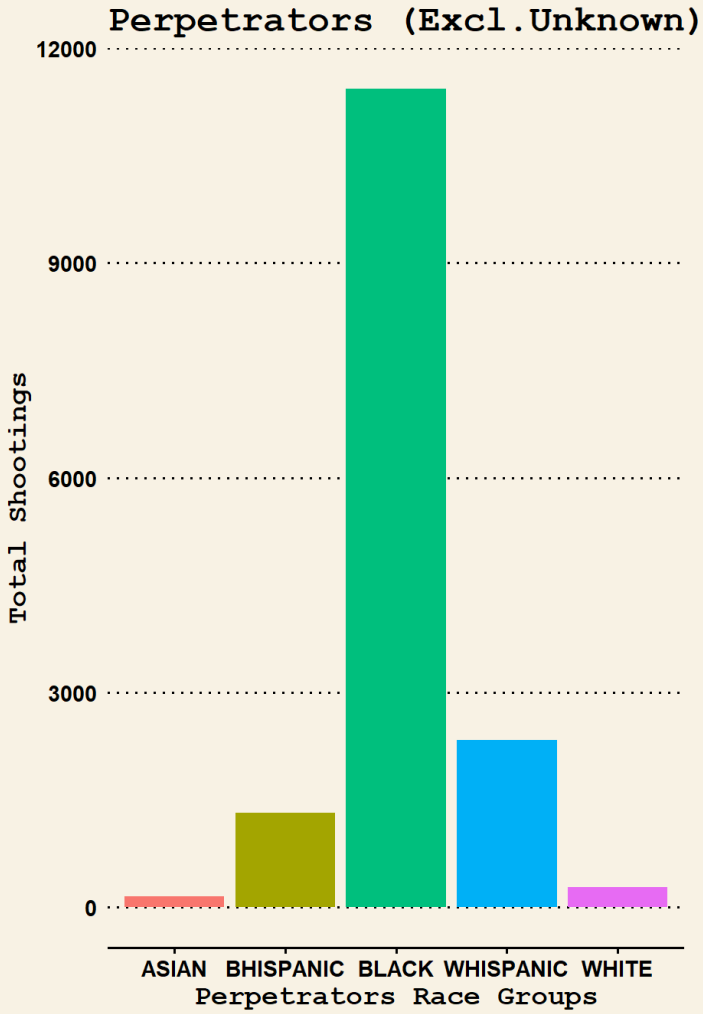


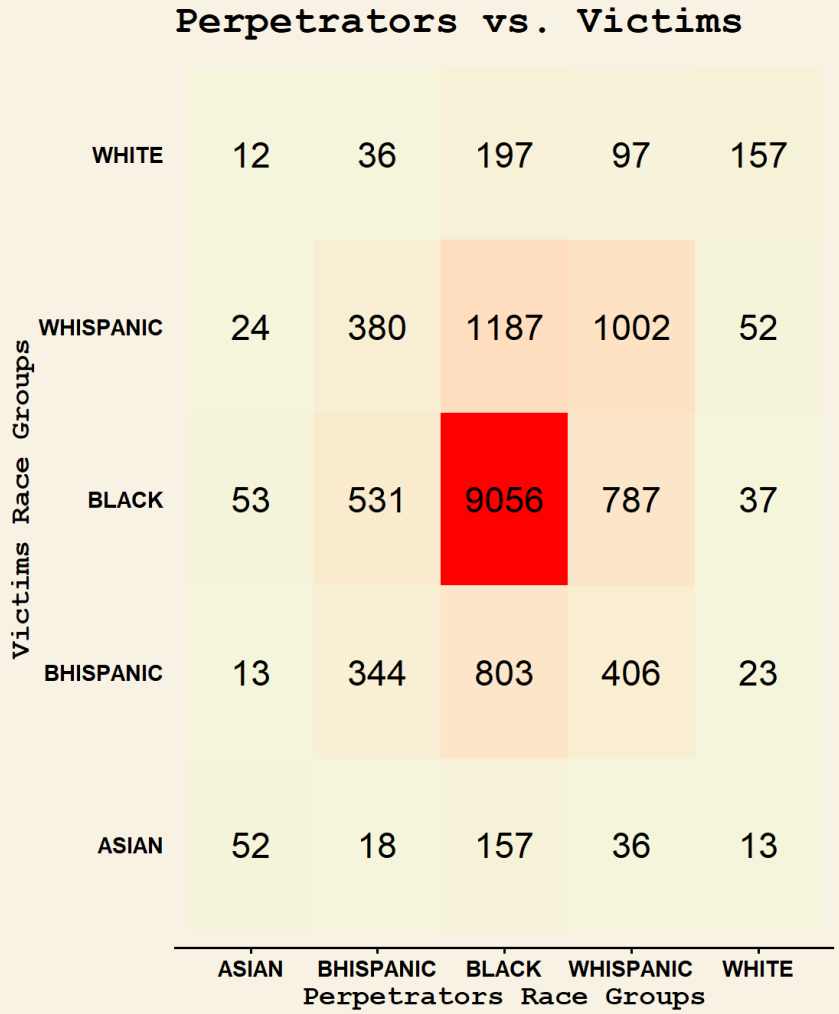


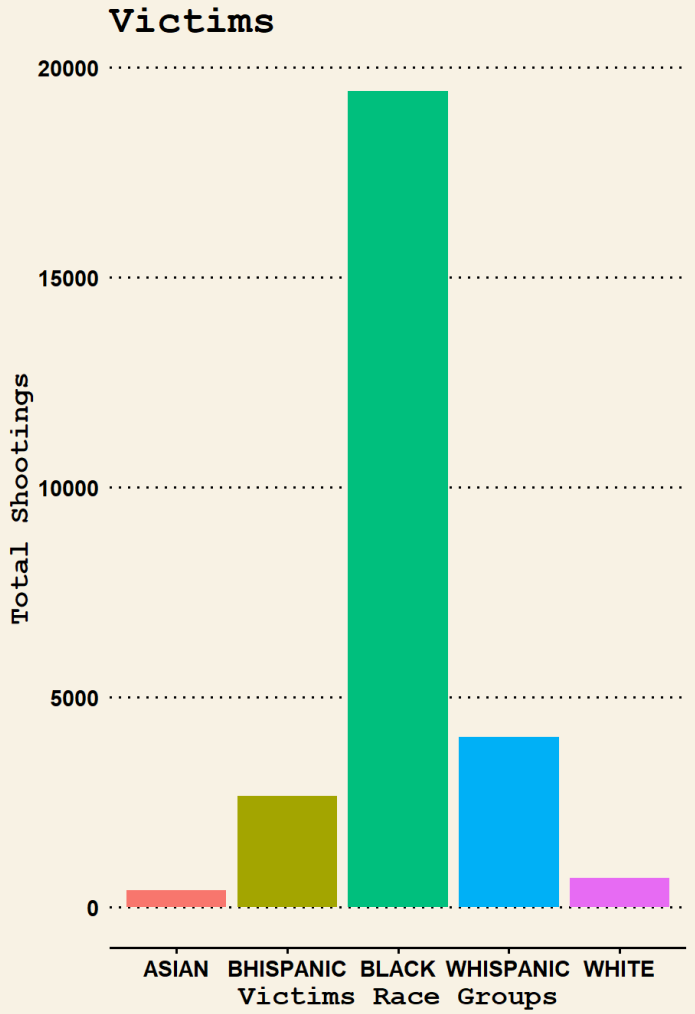


Comparison of Perpetrators and Victims Across Race Groups

- Black takes up about 70% of shootings in both Perps and Vics
- White and Black Hispanic together exceed 20% in both groups

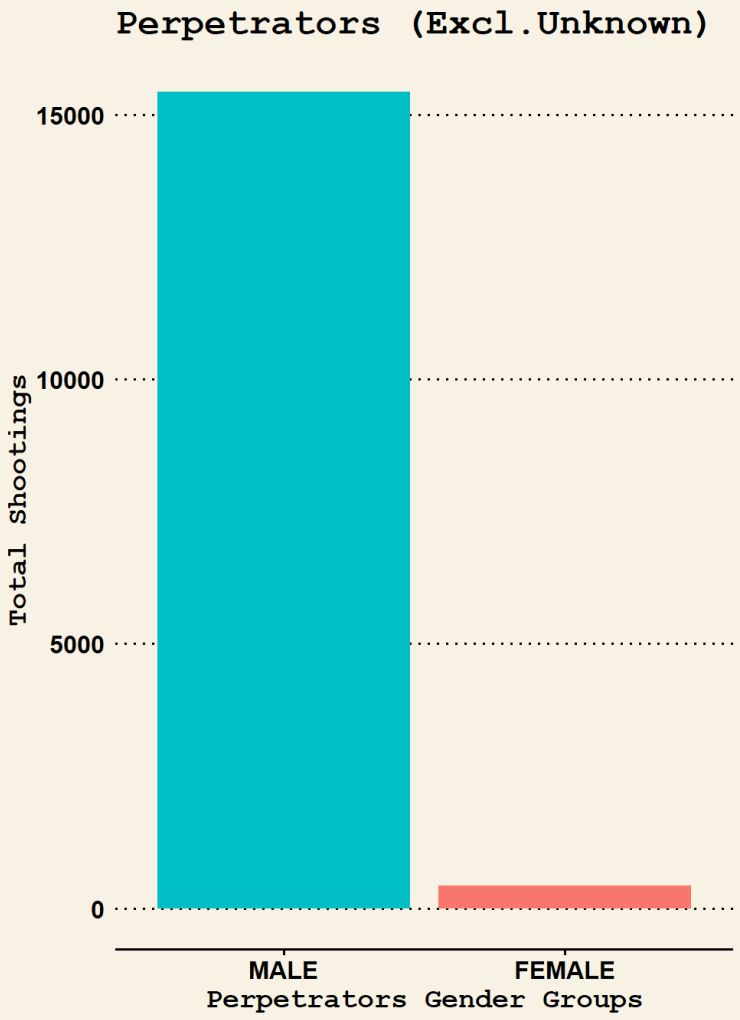


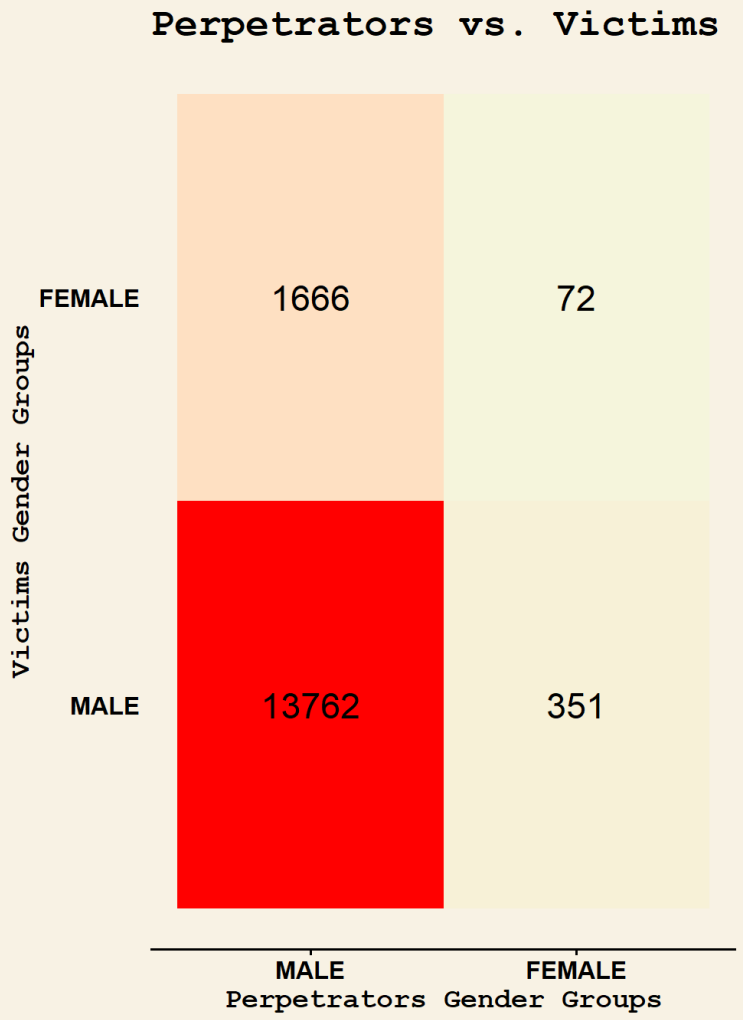


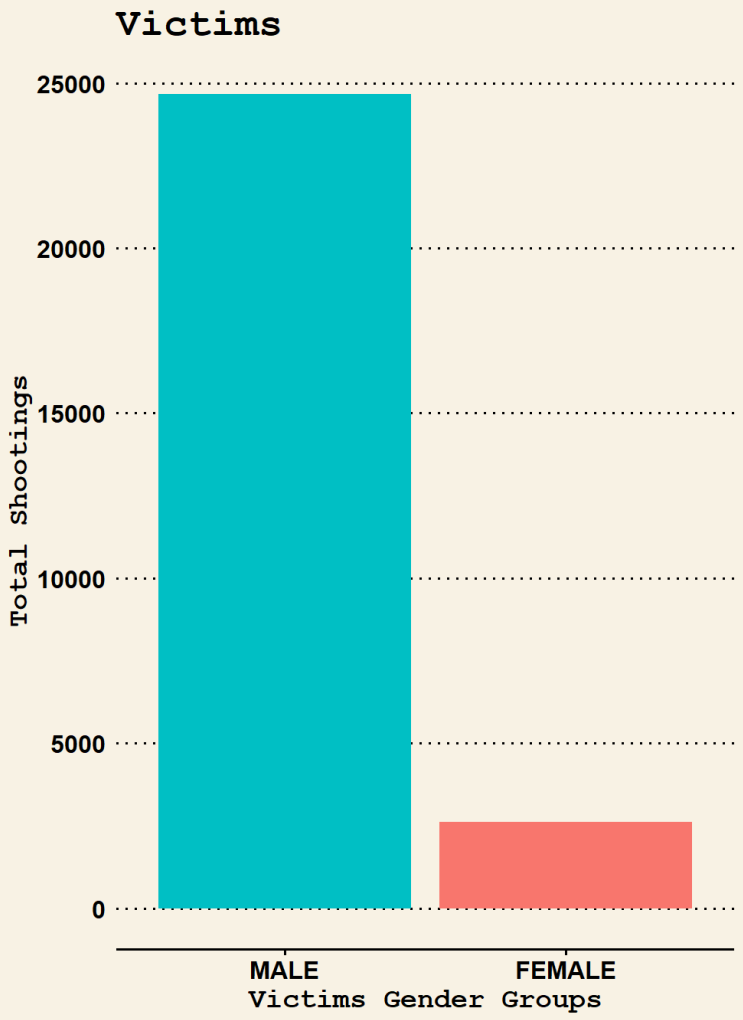


Comparison of Perpetrators and Victims Across Genders

- Over 97% of Perpetrators are male
- More than 89% of Victims are male





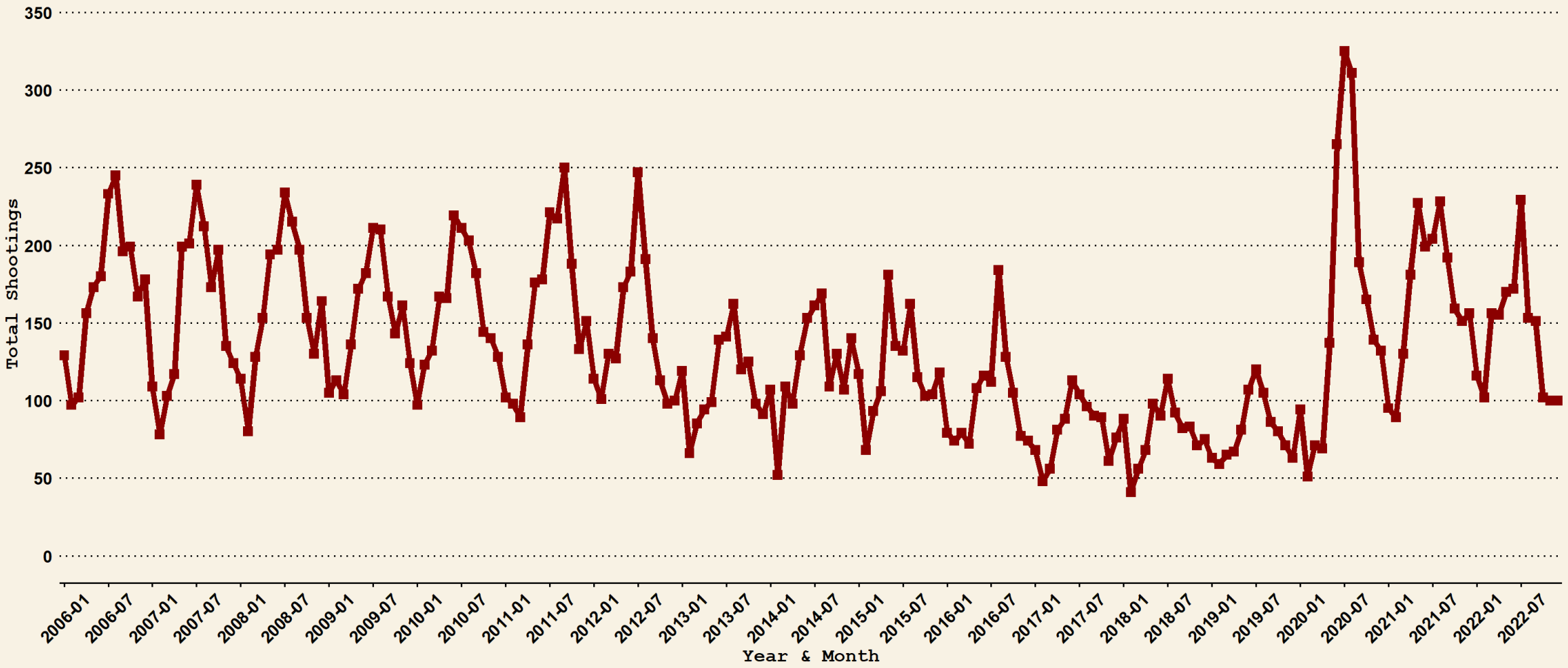


Project Agenda

1. Project Introduction
2. Data Preprocessing
3. Time Trend Analysis
4. Geographic Analysis
5. Profiling Analysis
6. ARIMA Forecasting
7. Conclusion

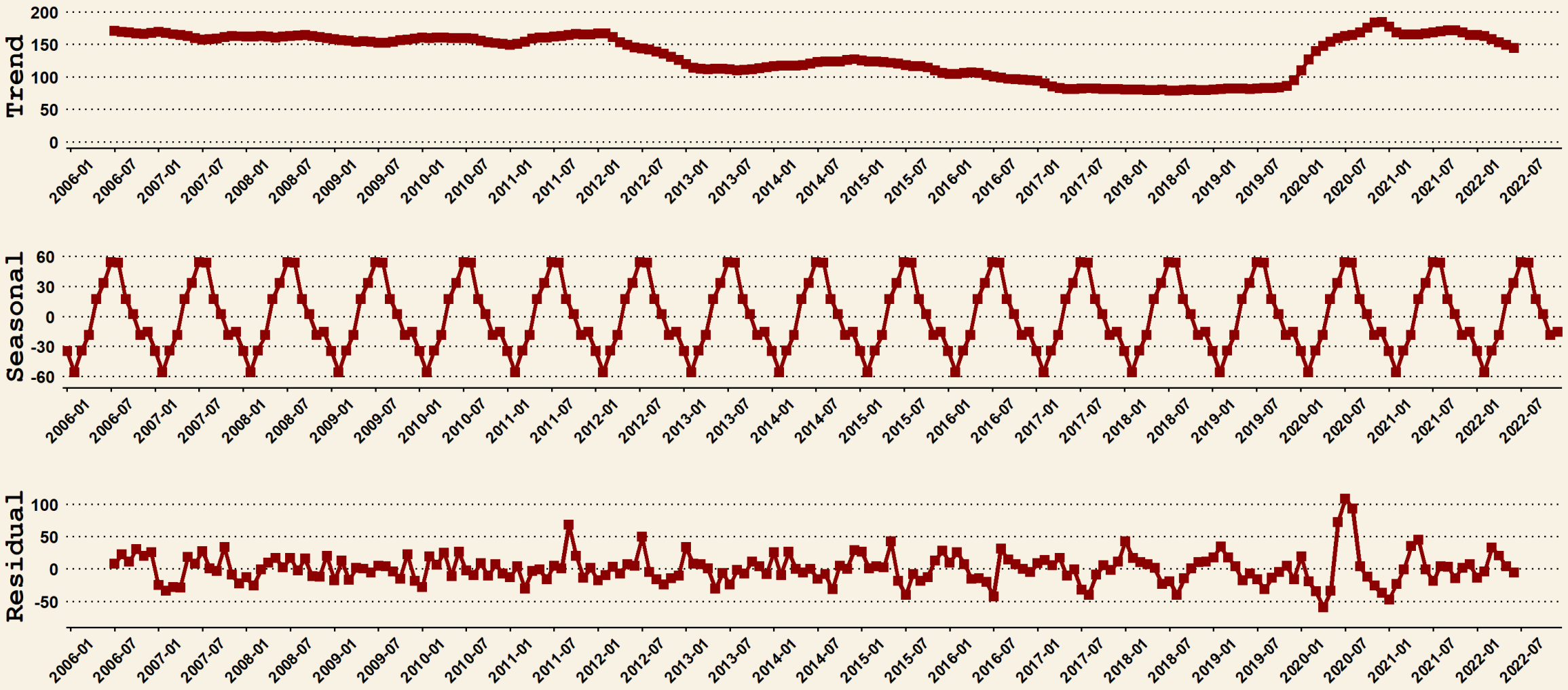
Time Series Monitor by Year & Month (2006~2022)

- Overall trend is lower in the middle and higher on the sides
- There is a noticeable seasonal fluctuation within each year



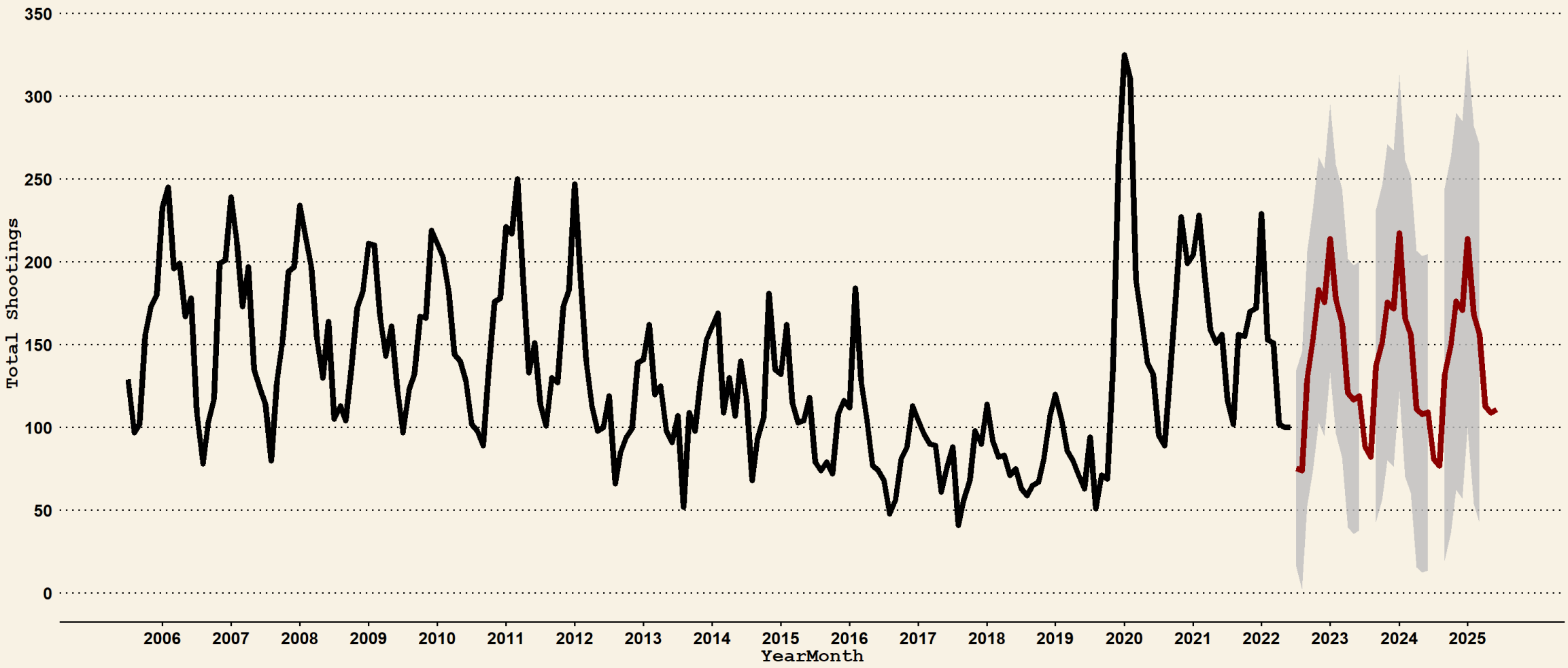
Time Series Decomposition

- Monthly counts decompose into Trend, Seasonal, and Residual
- High residual suggests randomness or external influence



Time Series Forecasting for Next 3-Years Using ARIMA

- The average trend will stay consistent with the level of 2022
- The range of fluctuations could reach the peak in 2020



Project Agenda

1. Project Introduction
2. Data Preprocessing
3. Time Trend Analysis
4. Geographic Analysis
5. Profiling Analysis
6. ARIMA Forecasting
7. Conclusion

Conclusion

- Through Timeline Trend Analysis: we explored the overall trend and found significant peaks during summer, weekends, and midnight
- Through Geographic Analysis: we investigated the spatial patterns of case distribution by boroughs and precincts
- Through Suspects & Victims Analysis: we identified distinct perpetrator features and a trend towards involving more teenagers
- Through Time Series Analysis: we identified evident seasonal patterns and forecasted shooting counts for the next 3 years

Conclusion

Report Bias

Bias in Data Source:
Potential bias in NYPD data recording overly focuses on particular boroughs and race groups

Bias in Data Analysis:
We utilize more dimensions to explore correlations within cases, rather than solely concentrating on perpetrator profiles.

Further Actions

1. Suggest to adjust patrol schedules and fund allocation based on Geographic Analysis.
2. Suggest to conduct comprehensive inquiries into organized crime based on Profiling Analysis.
3. Suggest to prepare preventative measures for significant events by utilizing Time Series Analysis.

Thanks

for Watching