Beyond the Numbers: Analyzing NYC's Traffic Collision Patterns (2012-2024)

Duc Thanh (Timothy) Nguyen

U.S Department of Transportation Federal Highway Administration | Northeast Big Data Innovation Hub (NEBDHub) | National Student Data Corps (NSDC)

Introduction

Unveiling NYC's Traffic Safety Patterns: Our comprehensive analysis of crash data from 2012-2024 reveals striking trends in collision causes, from distracted driving to right-of-way conflicts. By examining over a decade of data, we've identified critical patterns in accident timing and behavior that could transform how we approach road safety in New York City. This groundbreaking Transportation Data Science Project harnesses the power of big data to create safer streets for all New Yorkers.

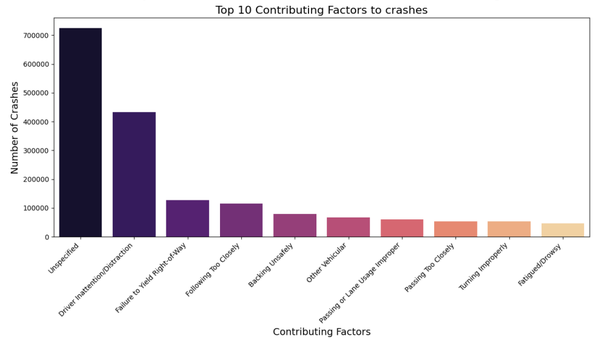

Figure 1. Top Contributing Factors to crashes

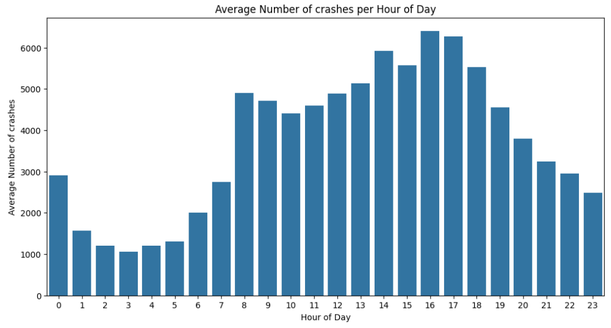

Figure 2. Average number of crashes per hour of day

Method

Our analysis of NYC traffic safety combined rigorous data preparation and ethical considerations with advanced time series and geospatial techniques to understand crash patterns. The methodology included cleaning and standardizing the NYC Motor Vehicle Collisions dataset, followed by detailed statistical analysis of crash patterns across time and location. Through interactive heatmaps and visualizations, we identified high-risk areas and temporal trends while considering socioeconomic factors that influence crash patterns.

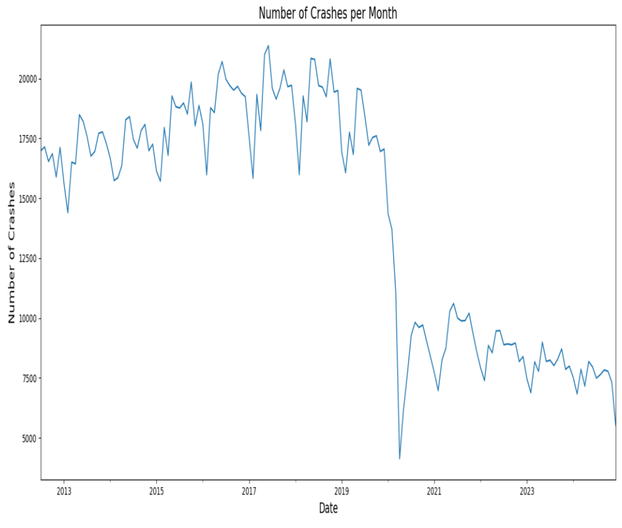

Figure 3. Number of crashes per month

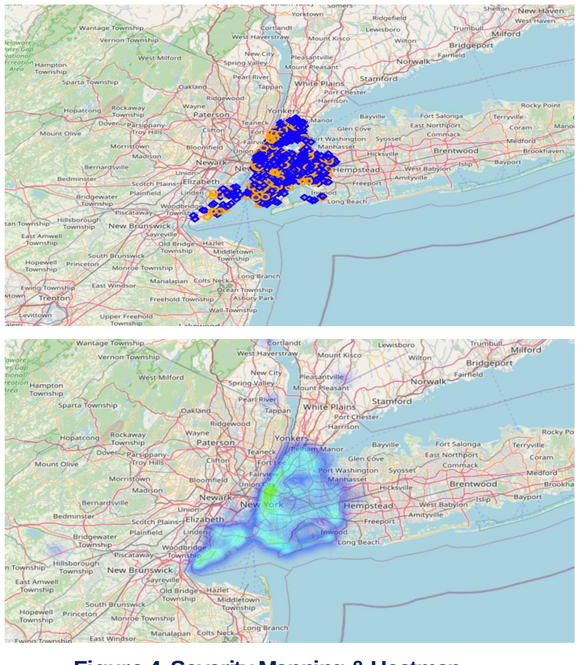

Figure 4. Severity Mapping & Heatmap

Conclusion

- Temporal patterns show crashes peak during evening rush hours (3-5 PM), with a significant decrease during COVID-19 that led to sustained lower crash levels, indicating lasting changes in traffic patterns.
- Spatial analysis reveals crash hotspots in high-density areas, with Brooklyn recording the highest number of crashes, followed by Queens and Manhattan, though this correlates with population density and borough size.
- Socioeconomic factors play a crucial role in crash reporting and frequency, suggesting a need for more equitable distribution of traffic safety resources and improved reporting systems across all neighborhoods, particularly in under-resourced communities.

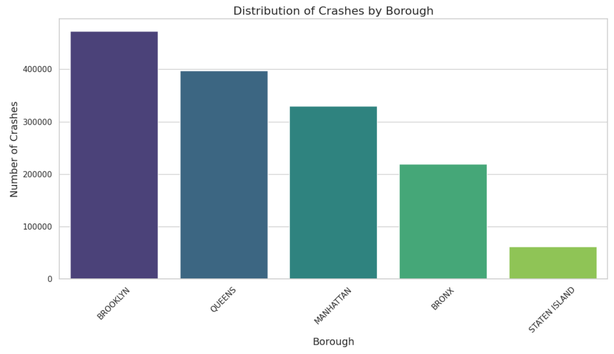

Figure 5. Distribution of crashes by Borough

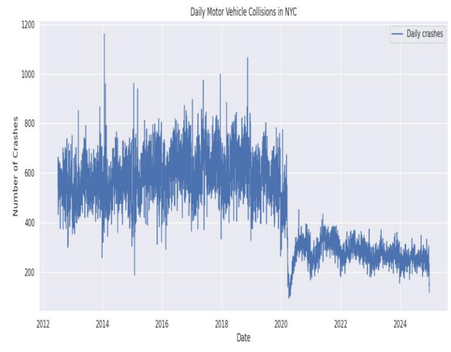

Figure 6. Motor Vehicle Collison Trend

Future Direction

- Integrate real-time weather data, traffic volume statistics, and road construction information to develop a more comprehensive understanding of crash risk factors and their interactions.
- Develop predictive machine learning models that could forecast high-risk periods and locations, enabling proactive deployment of traffic safety resources and emergency response teams.
- Expand the socioeconomic analysis by incorporating demographic data, income levels, and infrastructure quality metrics to ensure equitable distribution of traffic safety improvements across all communities.

Acknowledgments

I extend my sincere gratitude to the U.S. Department of Transportation Federal Highway Administration, Northeast Big Data Innovation Hub, National Student Data Corps, and NYC Open Data for providing the essential resources and datasets that enabled this comprehensive traffic safety analysis.

References

Police Department (NYPD)

NYC Open Data