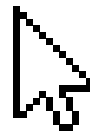




New York City: A Study on Motor Vehicle Collisions



David Wang



01 What are we studying?

02 What material and methods are we using?

03 Data Exploration

04 Data Modeling

05 Concluding Thoughts



01

What are we studying?



What are we studying?

- We will be studying the Motor Vehicle Collision dataset from NYC Open Data to gain insights into collisions that occur on NYC roadways.
 - The dataset is a collection of NYPD Collision Reports that is anonymous for data analysts to studied.
- By combining data about zoning districts and weather data, we will aim to get insights into the patterns and behaviors of collisions that occur in NYC.



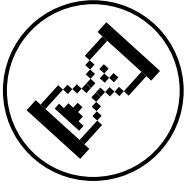
02

What materials and methods
for this study?



What methods will we be employing?

- We will be web scraping to gather our weather data.
- In order to conduct our study on our dataset, we will be utilizing data visualization libraries to graph our data to better understand trends that exist within our dataset.
- These visualizations will provide insight compare to just looking at columns and rows of data.
- We will utilizing models to develop a model to predict if a collision would result in a injury by selecting features that provides the best results.



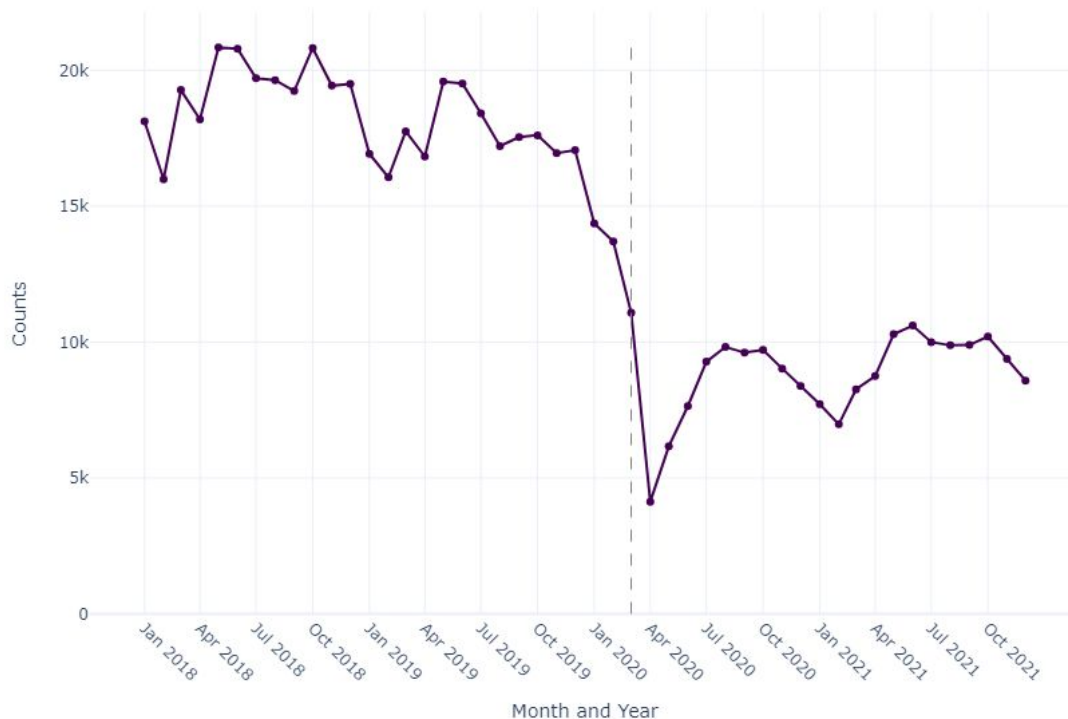
03

Data Exploration



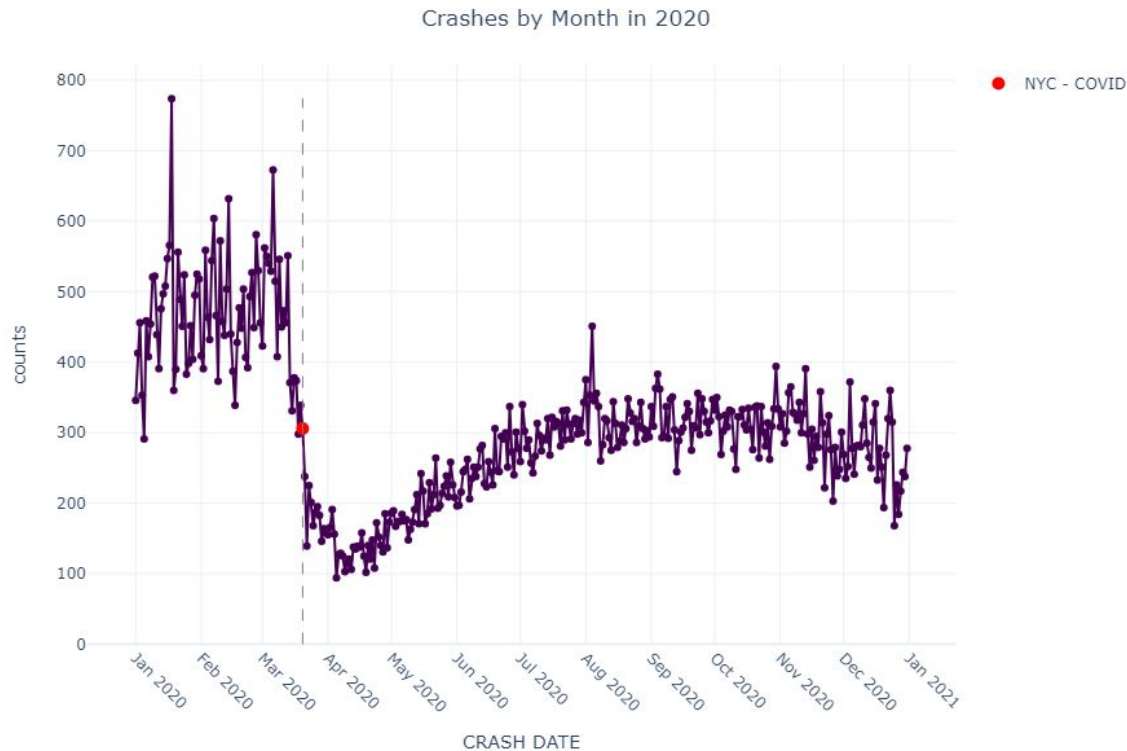
Is there any notable trend?

Crashes by Month (Jan 2018 - Dec 2021)



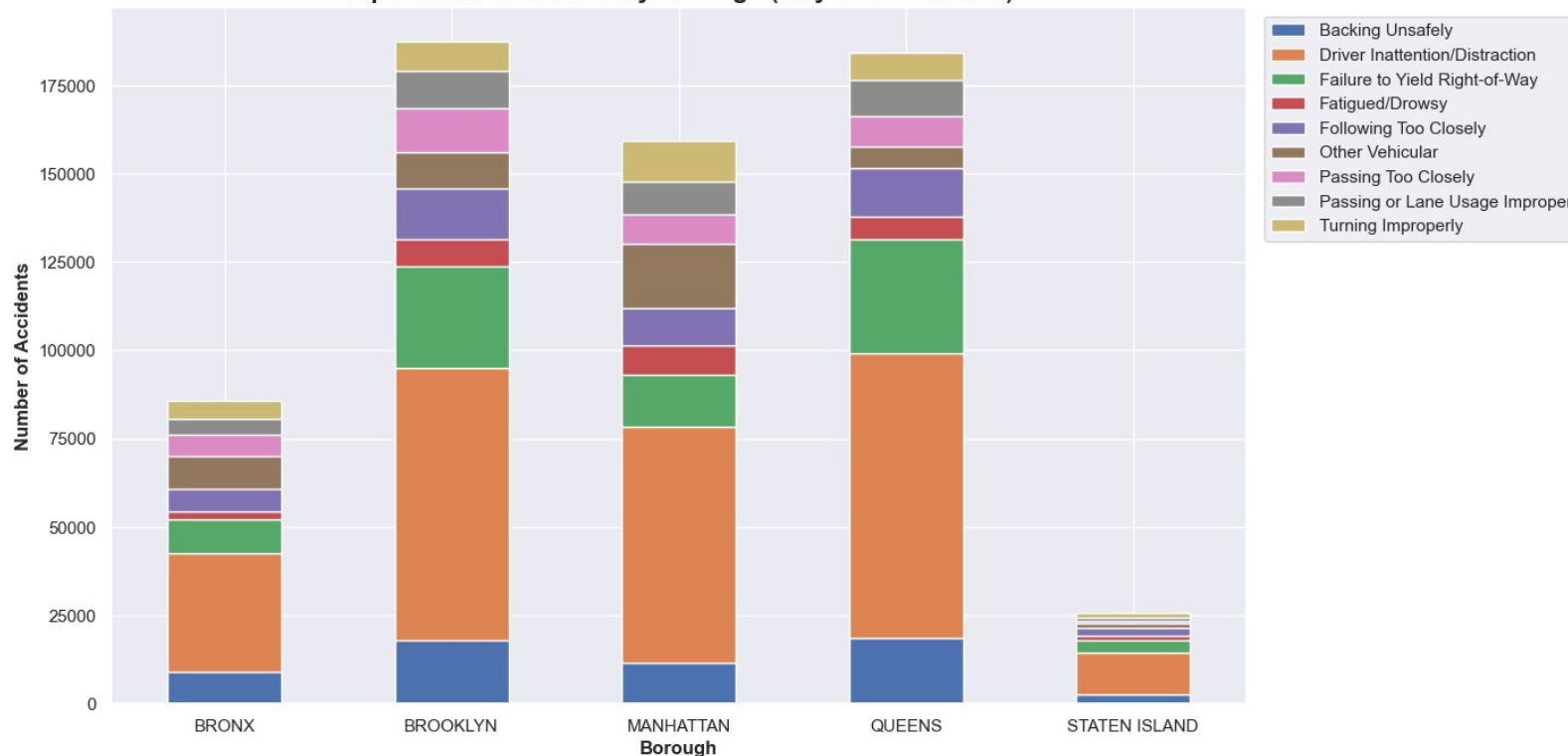


Did COVID 19 affect the number of collisions to occur?



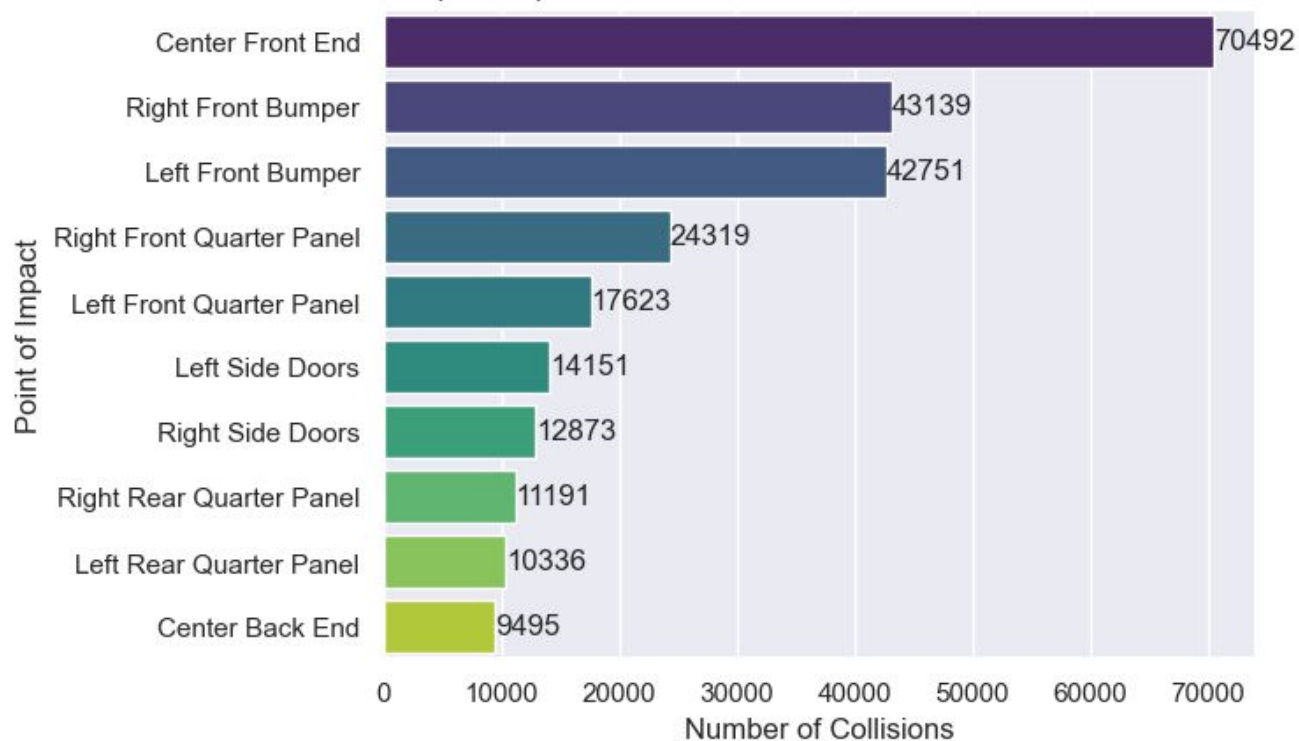


Top 9 Accident Causes by Borough (July 2012 - Present)



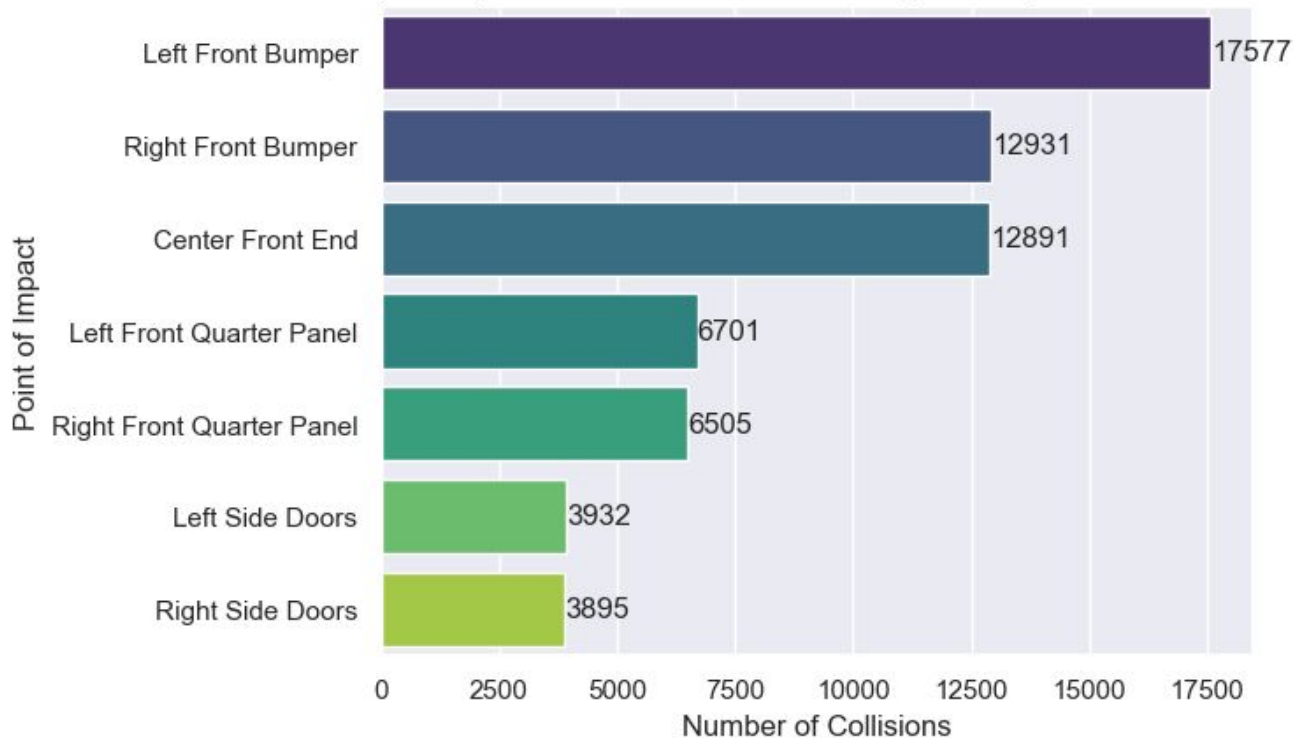


Top 10 Impact Points for Driver Inattention/Distraction Collisions



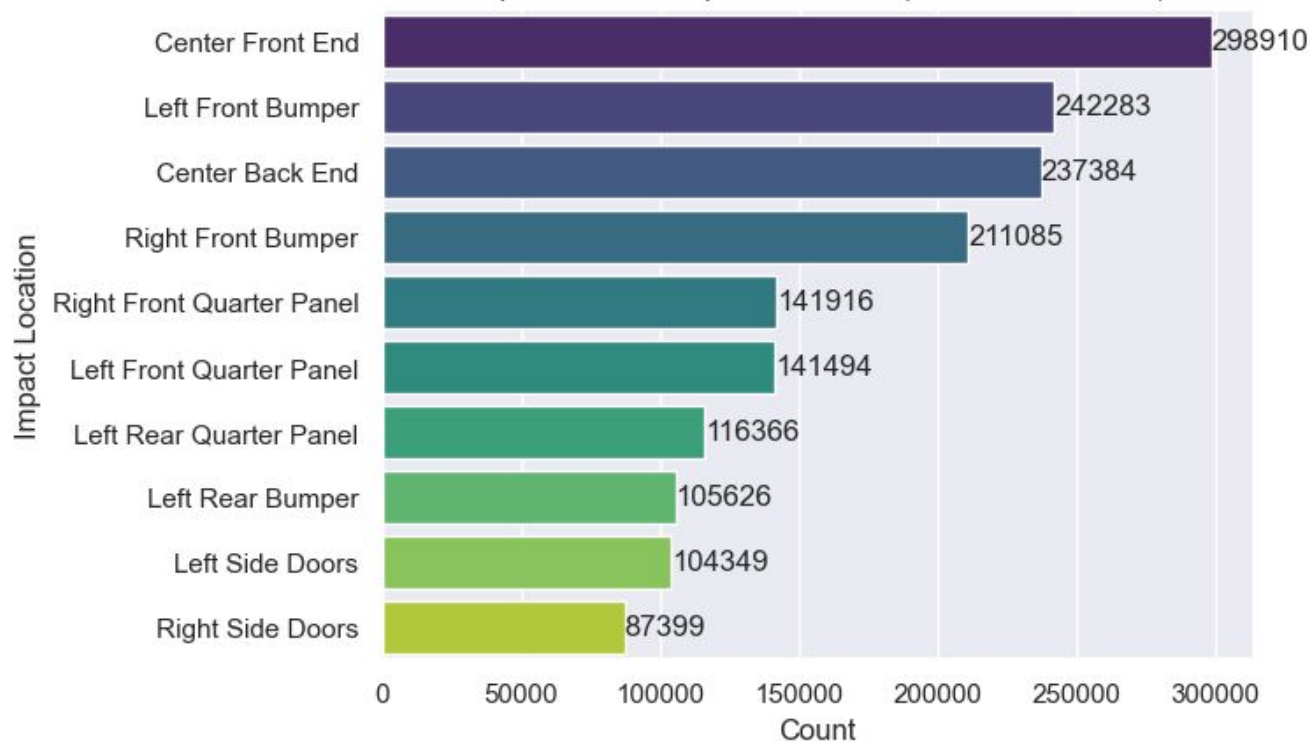


Top 10 Impact Points for Failure to Yield Right-of-Way Related Collisions



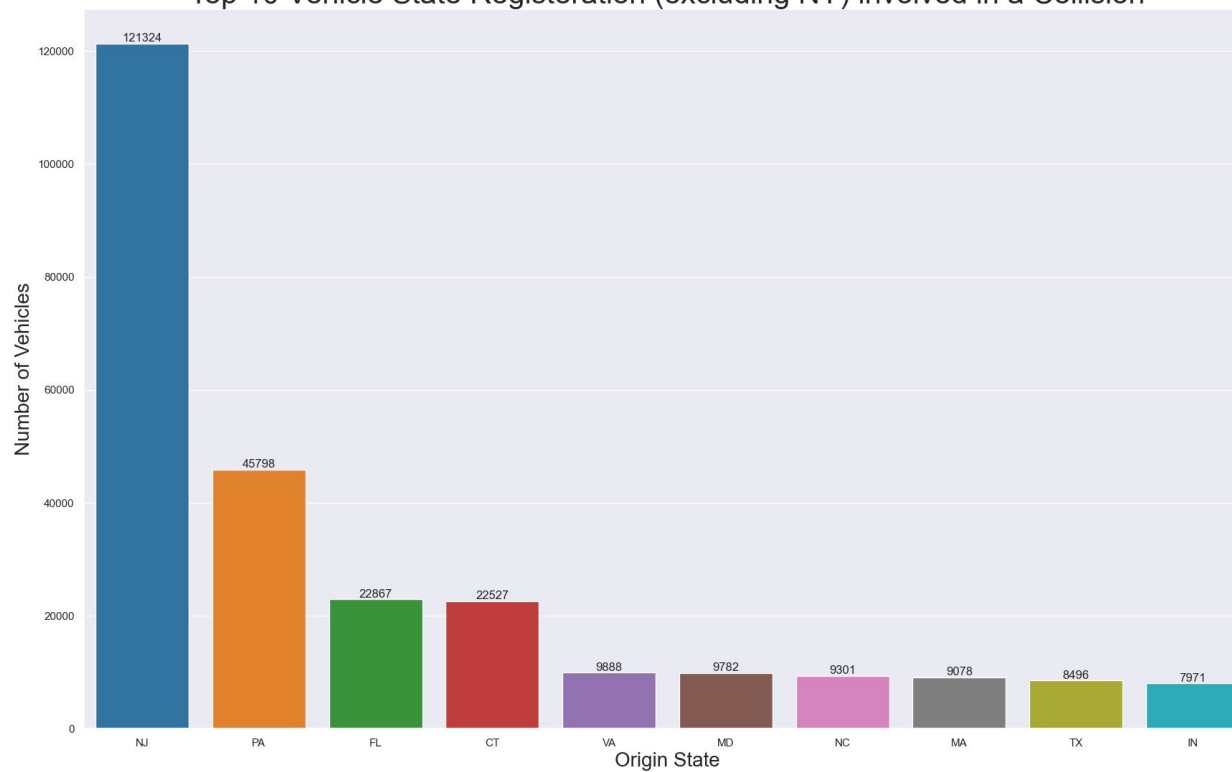


Top 10 Point of Impact Locations (June 2012-Present)



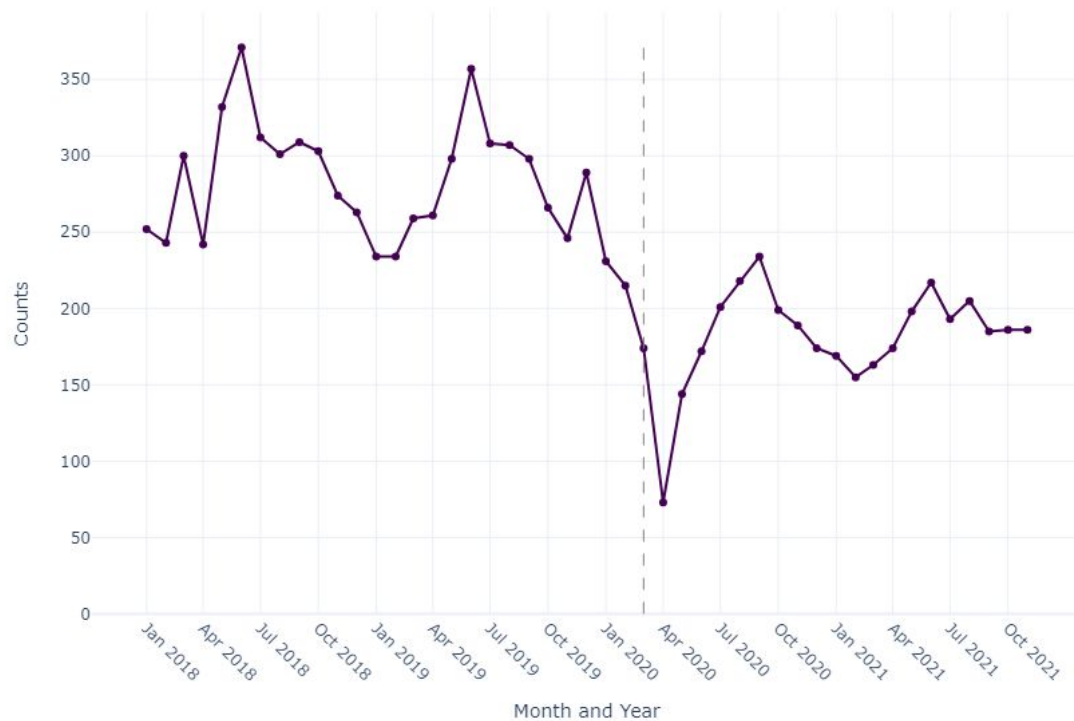


Top 10 Vehicle State Registration (excluding NY) involved in a Collision



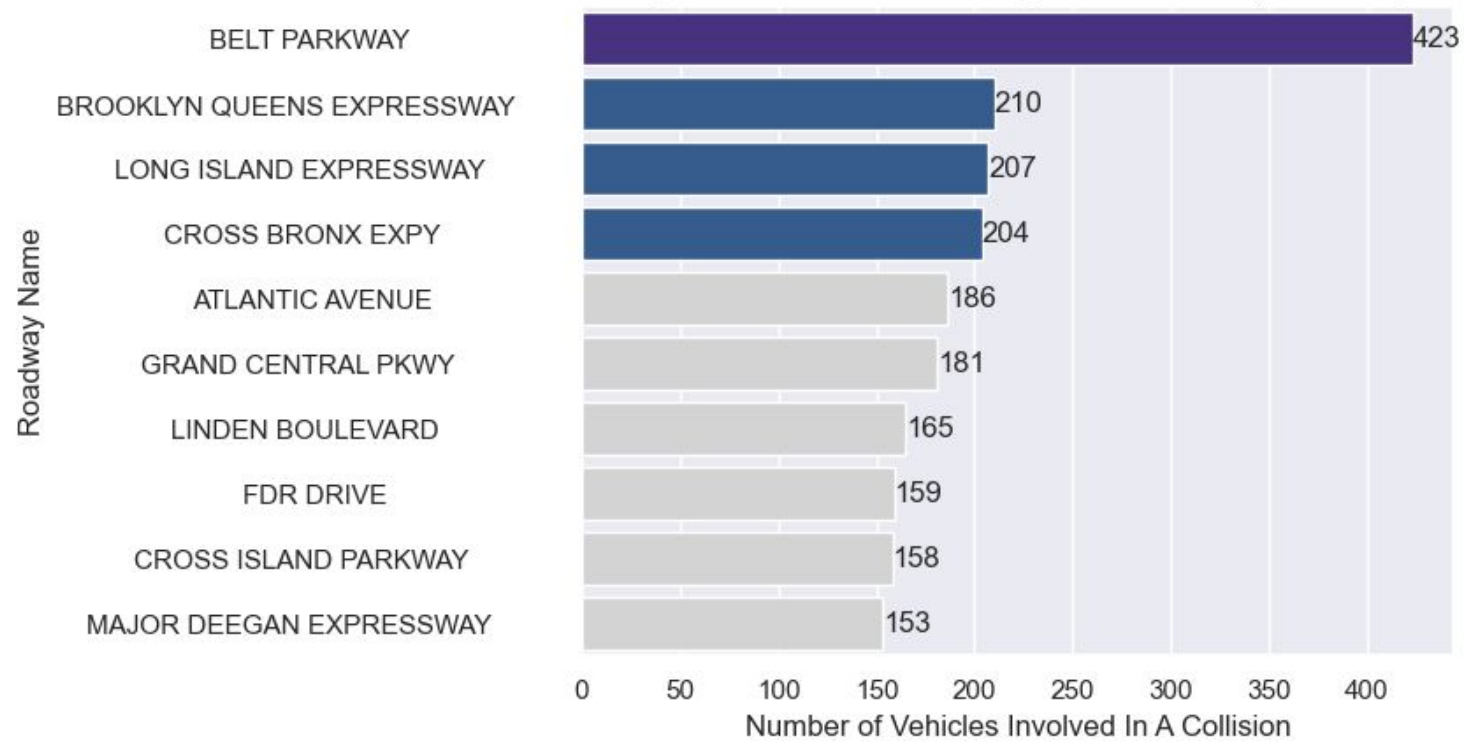


Floridian Registered Vehicles Collisions by Month (Jan 2018 - Nov 2021)





Top 10 Streets with Florida Registered Vehicles (2016-2023)





Collision Distribution Across NYC Zoning Districts (June 1st, 2012 - Dec 2nd, 2023)

Zone	Number of Collisions Reported
Residential	489115
Commercial	252229
Park	181189
Mixed Manufacturing/Residential	119237

Collision Distribution Across NYC Zoning Districts in 2022

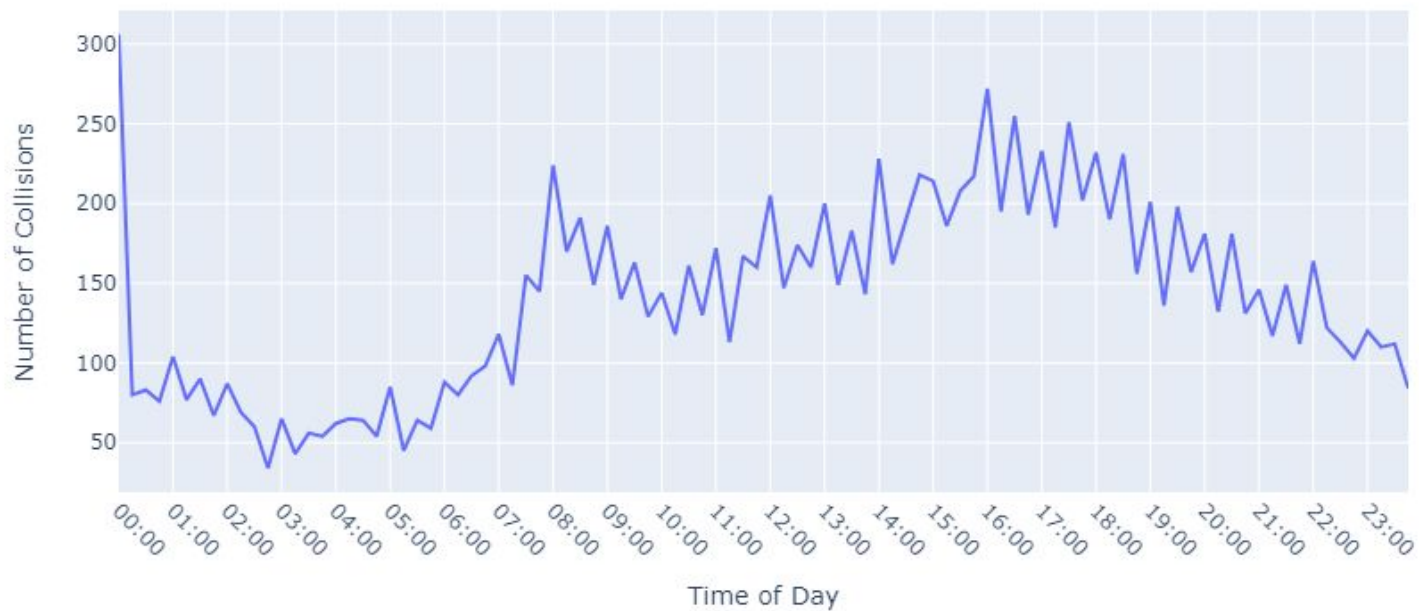
Zone	Number of Collisions Reported
Residential	15303
Commercial	6162
Park	5742
Mixed Manufacturing/Residential	3738

Injuries Reported at Midnight in 2022

Collision Reported Injury	Count
No Injuries	271
Injury	186

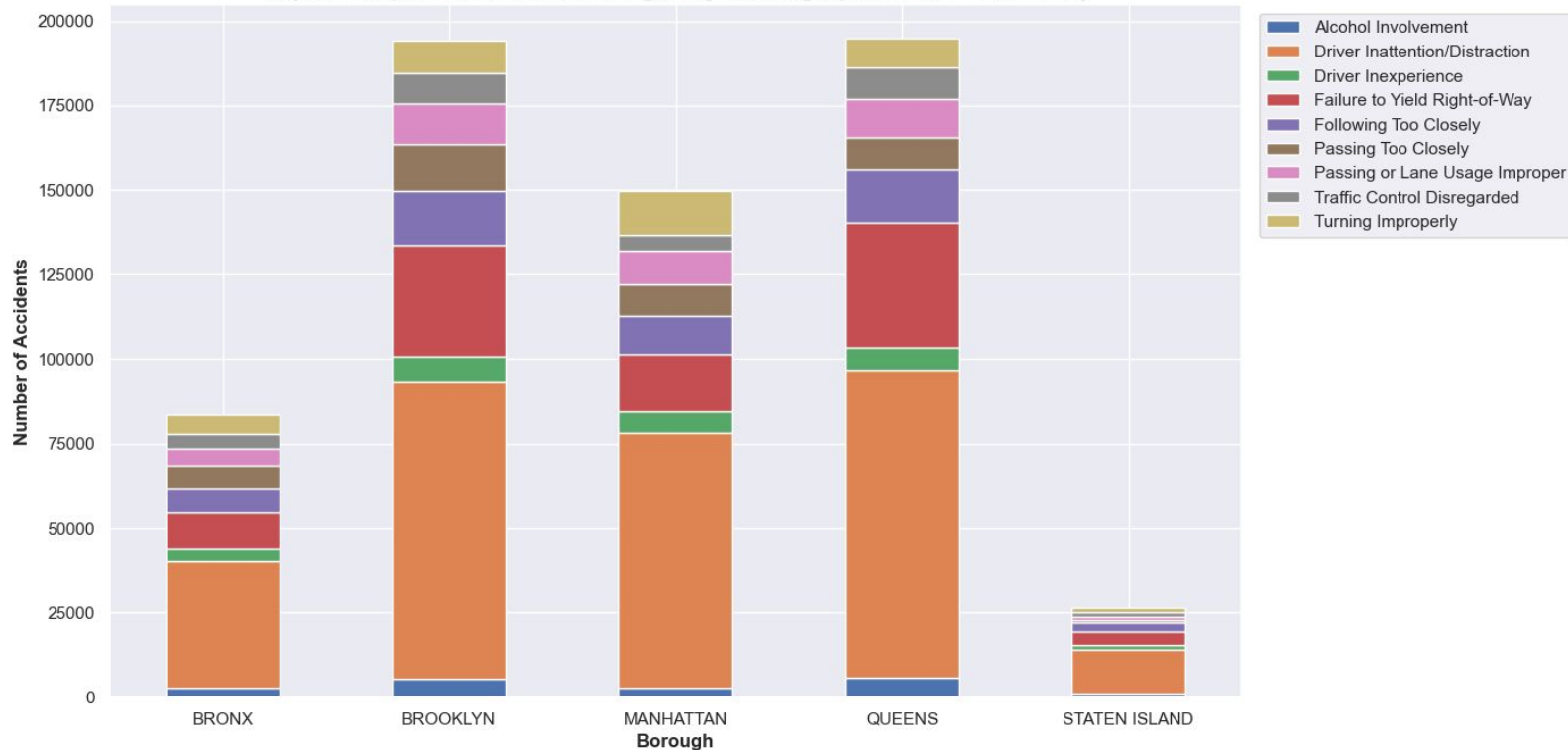


Collisions in Residential Areas by Time of Day (2022)





Top 9 Accident Causes At Midnight By Borough (Jan 2022 - Dec 2022)





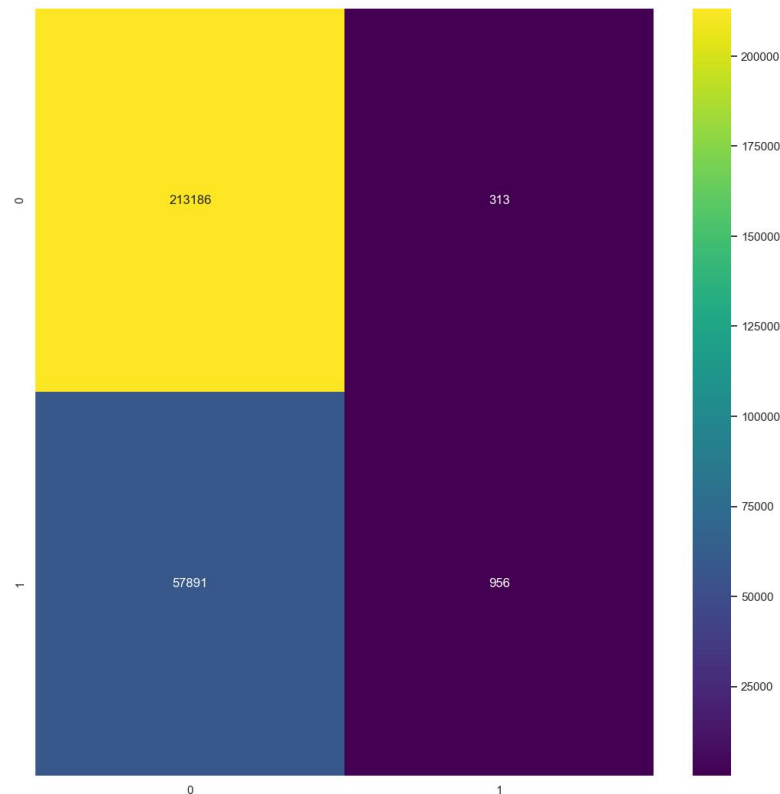
04

Data Modeling Results



Baseline Random Forest Model

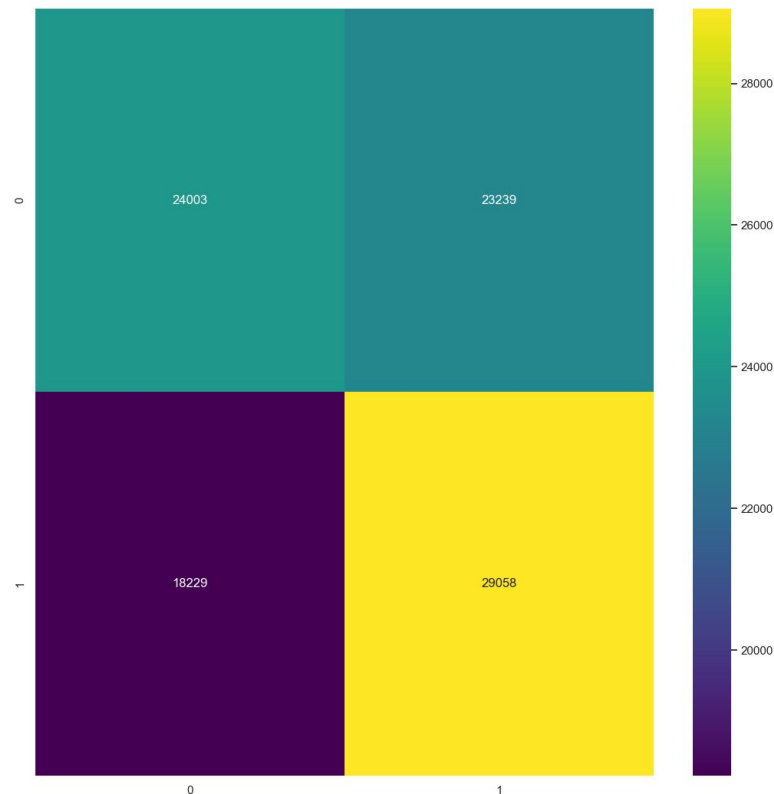
	precision	recall	f1-score	support
No	0.79	1.00	0.88	213499
Yes	0.75	0.02	0.03	58847
accuracy			0.79	272346
macro avg	0.77	0.51	0.46	272346
weighted avg	0.78	0.79	0.70	272346





Undersampled Random Forest Model

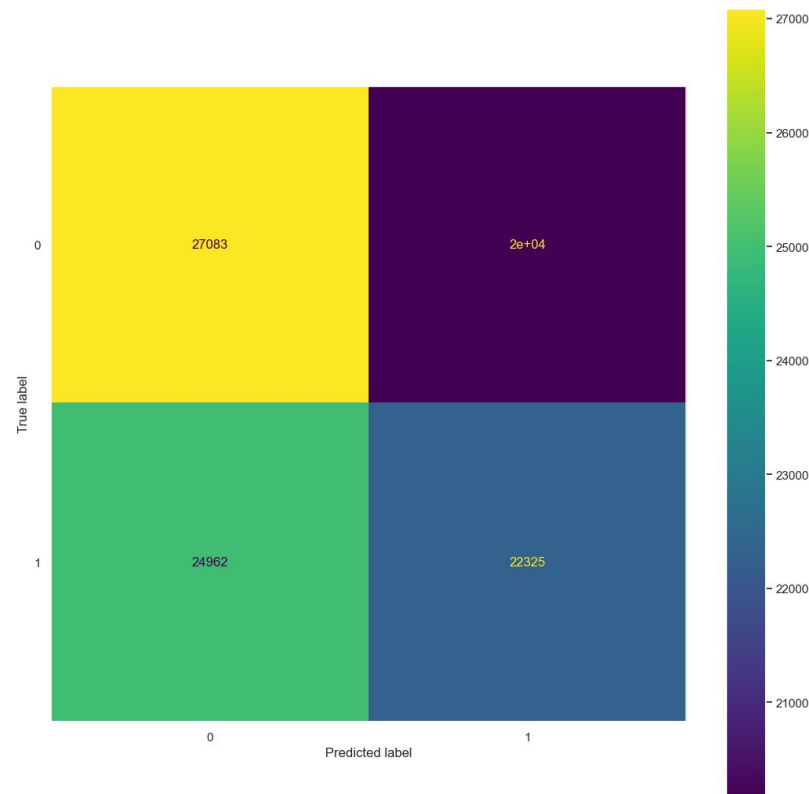
	precision	recall	f1-score	support
No	0.57	0.51	0.54	47242
Yes	0.56	0.61	0.58	47287
accuracy			0.56	94529
macro avg	0.56	0.56	0.56	94529
weighted avg	0.56	0.56	0.56	94529





Undersampled K-Neighbors Model

	precision	recall	f1-score	support
No	0.52	0.57	0.55	47242
Yes	0.53	0.47	0.50	47287
accuracy			0.52	94529
macro avg	0.52	0.52	0.52	94529
weighted avg	0.52	0.52	0.52	94529





05

Concluding Thoughts



Concluding Insights

- COVID-19 Quarantine protocols had a drastic impact on the decreasing number of collisions that occurred on NYC roadways.
- The fewer people on the road, the less chance someone will be involved in a collision.
- Driver distractions were the leading cause of collisions that occurred. Those types of collisions often had the point of impact at the center of the front end of a vehicle.
- Driver's inability to yield the right of way was another lead cause of collisions. Those types of collisions often had the point of impact at the left front bumper.
- New Jersey registered vehicles are the most common out-of-state registration that is mostly involved in accidents on NYC roadways.
- A majority of collisions occur every half hour between 8 AM EST and 10 PM EST.