NYC Collision Data Report

Capstone by Moritz Schröder

Value of identifying factors in collisions

Traffic safety is of paramount importance to the NYPD and NYC city regulators. Identifying factors that impact the number and severity helps them in multiple ways:

- Increasing overall road safety and security for drivers as well as pedestrians
- Identifying more targeted focus areas for regulating and policing traffic with maximum efficacy when it comes to road safety
- Measuring the impact of changes in policing and regulating and the ways they effect overall safety

NYC Data set

The report is based upon a data set by the City of New York. It includes 1.048.575 collisions having occurred in NYC from April 2015 to August 2019. For each accident the data includes the following tags:

- Date & time of occurrence
- Borough & zip code it occurred in
- Number of persons killed & number injured
- Cause of accident
- Type of vehicle causing the accident

The data set was prepared by deleting incomplete entries and consolidating tags, thereby reducing the complexity of the data

Defining severity of accidents

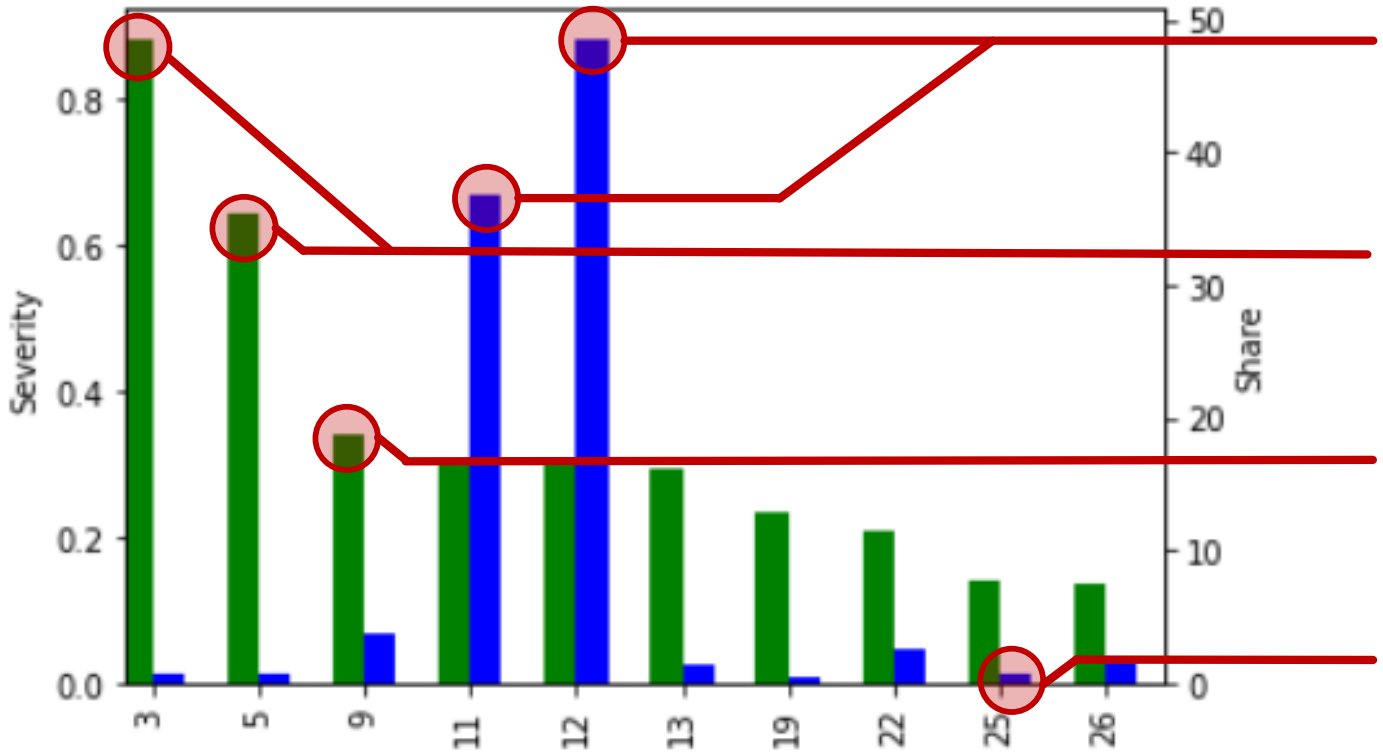
The fist step was defining a target variable to measure severity:

Overall Severity = # of injuries + 3x # of fatalities

Then the impact on severity and collision frequency was calculated for:

- **Vehicle type**: Showing if the type of vehicle causing *(not equivalent to "involved in")* the accident had an impact on its severity and if there was vehicle types especially prone to (severe) accidents
- **Cause**: Decipher which causes resulted in the most/the most severe accidents
- **Time**: Showing how frequency and severity of collisions changed throughout the day
- **Borough**: Finding out which parts of NYC had the most/most severe accidents

Differentiation by vehicle type



3 – Bikes
5 – Motorcycles
9 – Taxis

11 – Sedan
12 – SUV
13 – Bus

19 – Van
22 – Pick-up Truck
25 – Tractor
26 – Box Truck

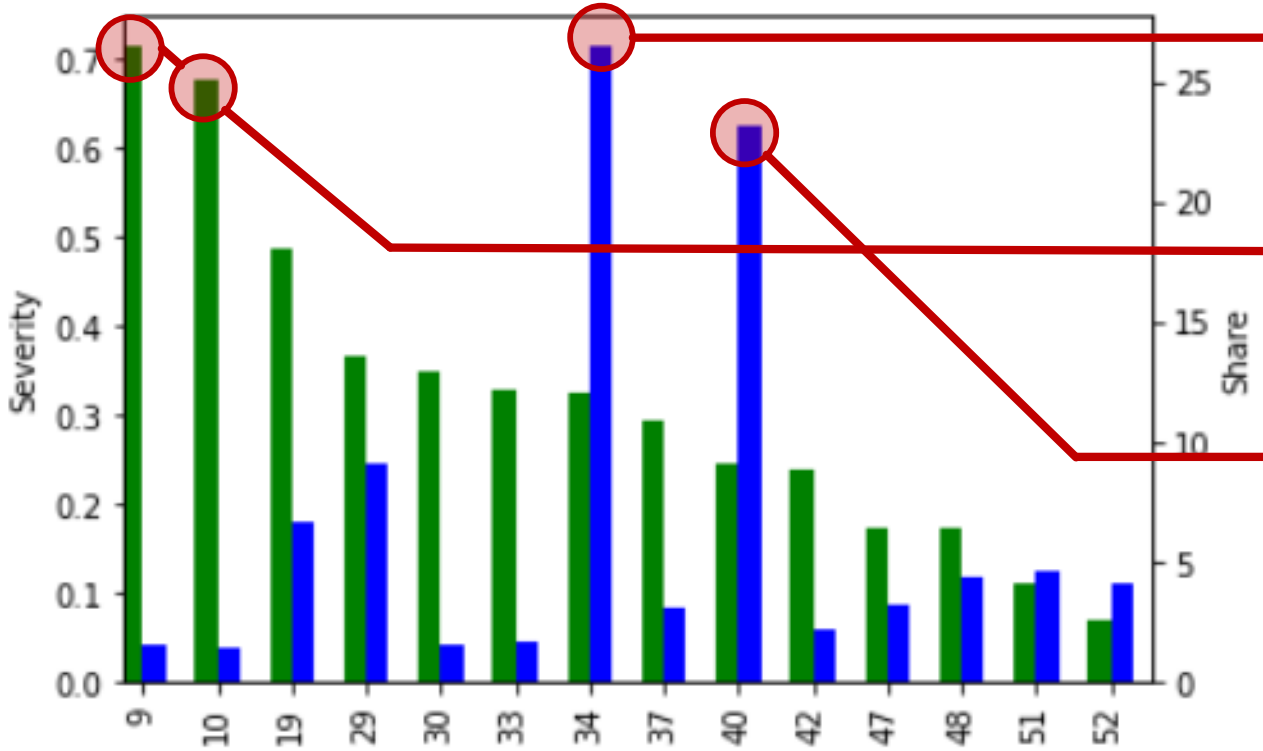
SUVs (48.5%) and **Sedans (36.8%)** make up over **85%** of the collisions in the dataset.

Bikes and motorcycles averaged the highest severity score with **0.88** and **0.64** respectively

Taxis ranked almost **10% higher** on the average severity compared to **SUVs and Sedans** despite being similar vehicles

Tractors cause an astonishing **0.8%** of collisions (in NYC!)

Differentiation by vehicle type



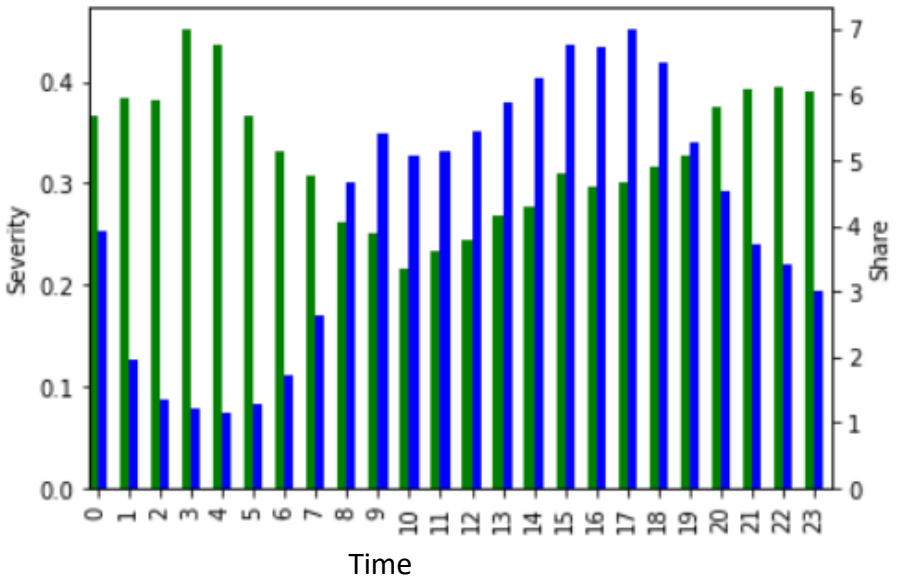
9 – Traffic Control Disregard
10 – Unsafe Speed
19 – Failure to Yield
29 – Followin too closely
30 – Reaction to Uninvolved Vehicle
33 – Driver Inexperience
34 – Distraction
37 – Other Vehicular
40 – Unspecified
42 – Turning Improperly
47 – Unsafe Lane Change
48 – Improper Passing
51 – Backing Unsafely
52 – Passing too Closely

Driver Inattention causes **26.5%** of all accidents

Traffic Control Disregard and **Unsafe Speed** cause the by far most severe accidents

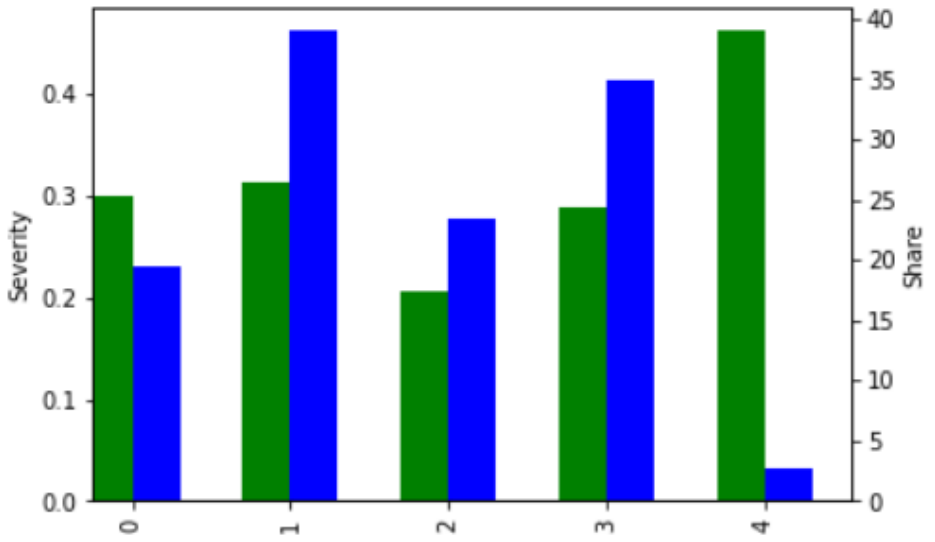
For **23.3%** of all collisions no cause is noted, resulting in a less telling dataset

Differentiation by time & place



The severity peaks in the **early morning hours** steadily reducing until **noon**

The amount of incidents corresponds strongly to **commuting/business hours**



0 – Bronx
1 – Brooklyn
2 – Manhattan
3 – Queens
4 – Staten Island

Incidents in **Staten Island** are on average more than twice as severe as in **Manhattan**

Brooklyn accounts for almost **40%** of all incidents tracked in NYC

Summary

The four major takeaways are:

- **Driver inattention causes more than 1 in 4 accidents in NYC**
- **Taxis cause 3.8% of collisions which are 10% more severe than comparable vehicles**
- **“Traffic control disregard” and “unsafe speed” cause accidents that are twice as severe than average**
- **Incidents in Manhattan are half as severe as ones in Staten Island**

The report has shown that the data provided is an extremely valuable resource. It has displayed a variety of causes and effects impacting the traffic system of NYC. However, there is far more information to be retrieved but this report helps inf finding out where to intensify the analysis