

Motor Vehicle Collisions in New York City

New York City is one of the most populous cities in the United States, with over 8 million people living there. This population increases the number of vehicles that are on the road in any given time. For my analysis I wanted to see how nature affected vehicle collisions in NYC. More specifically I wanted to see if the seasons and/or time of day had a significant impact on collisions, and if we could make predictions based on this analysis.

First, I wanted to see how COVID-19 affected total vehicle collisions. By using a bar graph, I can see that in 2020 the amount of vehicle collisions decreased. I then separated the data in pie charts to see the distribution of collisions based on borough, season, and day/night time. Brooklyn had the highest number of total collisions. The collisions were evenly spread between the seasons, while most accidents occurred during the daytime (defined as 6AM-6PM). Looking at a boxplot that showed how many collisions happened by hour, we can see that most accidents occur between 9AM-5PM. Looking at a line plot that specifies who was injured or killed, we can see that a motorist was injured most often, while pedestrians were killed the most during these collisions. Comparing injured vs killed in a scatterplot, we can see Staten Island clustered on the low ends, while Brooklyn was scattered amongst the higher numbers. A histogram of the hours the collisions occurred showed that most collisions happened between 4PM-5PM.

After doing this analysis, I have concluded that the seasons have no major effect of collision frequency. Most collisions happened during the daytime. It appears that population has biggest effect on motor vehicle collisions. This is shown during 2020, when there were less people on the road. Total collisions by borough from highest to lowest are as follows: Brooklyn, Queens, Manhattan, Bronx, Staten Island. This is the same order of population per borough. It is important to look at other factors, such as road geometry, driver distraction, driving while impaired, faulty vehicle, etc. But based on the data I have analyzed I would hypothesize that population density has a strong correlation with motor vehicle collisions. The safest time to avoid a collision is at 3AM, while you have the highest chance of being involved in a collision at 4PM. Avoid Brooklyn if you want to minimize your chances of being in a collision and only stay in Staten Island. Of course, for most people this is not feasible, but at the very least anyone traveling at a time where collisions are most frequent should practice a higher level of alertness. Looking ahead, we could use this data to look at what can be done to decrease the overall number of collisions and make it safer for everyone. An analysis of other major cities could be done to see if they have the same results, and if not, what makes the difference between the two cities.

References

1. Wikipedia contributors. (n.d.). *Demographics of New York City*. In Wikipedia. Retrieved December 9, 2024, from https://en.wikipedia.org/wiki/Demographics_of_New_York_City?utm
2. U.S. Department of Transportation. (n.d.). *Motor vehicle collisions (crashes)*. Data.gov. Retrieved December 9, 2024, from <https://catalog.data.gov/dataset/motor-vehicle-collisions-crashes>
3. Illescas, E. (2024). Custom code for analyzing motor collision data. Unpublished.