# Investigating Fatalities in NYC Motor Vehicle Collisions by Hour of the Day

**Hanqi (Sunny) Shi**
**Swarthmore College**

## Research Question

We previously discovered that the hourly average number of crashes peaks from 4:00 to 5:00 p.m. every day, possibly due to the volume of traffic during rush hour (Explorer TDSP). In addition to the number of crashes, we should also aim at lowering the number of deadly crashes and decreasing the severity of injuries in collisions. This project asks whether the fatalities in collisions differ by time of the day and possible contributing factors to deadly collisions.

## Methods

The data used in this project comes from NYC Open Data. The dataset contains details of police reported motor vehicle collisions in New York City from 7/1/2012 to 10/20/2024. We first use Time of Day Analysis to visualize trends of injuries and fatalities in collisions throughout the day. We then select time periods with more fatal collisions and analyze associating factors using multiple line graphs and bar charts because these visualizations are clear and direct.

## Result: Injuries and Deaths by Hour of the Day

From 4:00 to 5:00 a.m. and from 11:00 p.m. to 12 a.m., we see the highest average number of persons injured or killed in collisions. Moreover, this number decreases in the morning hours and increases from noon to late night hours. The average number of fatalities in collisions is the highest from around 3 a.m. to 5 a.m.. This number remains roughly constant in day time and increases from 10:00 p.m. to 5:00 a.m. Collisions at night tend to cause more inrjuries and deaths than collisions during day time.

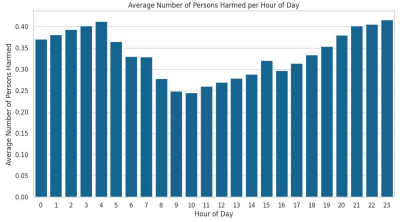


**Figure 1. Average number of persons injured or killed in collisions per hour of day**

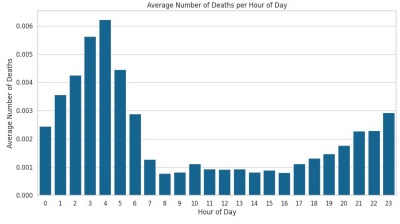


**Figure 2. Average number of persons killed in collisions per hour of day**

## Results: Factors Associated with Deadly Collisions

From 3:00 to 5:00 a.m., the average number of motorists killed in collisions is higher than that of pedestrians, whereas the reverse is true for the majority of the day. The second and third most common specified contributing factors to collisions from 3:00 to 5:00 a.m. are alcohol involvment and unsafe speed. Yet, neither is in the top nine specified contributing factors to collisions of the whole day (Explorer TDSP). This implies that more incidences of drunk-driving may be associated with more fatal collisions in a time of day.

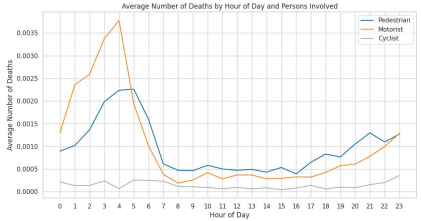


**Figure 3. Average number of deaths by hour of day and type of roadway user**

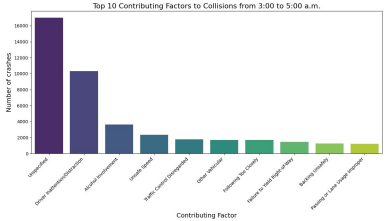


**Figure 4. Top ten contributing factors to collisions between 3:00 to 5:00 a.m..**

## Conclusion and Recommendations

Though on average there are less collisions at night, these crashes appear to be more deadly. Drunk driving at high speeds may be a leading cause in these deadly collisions, especially resulting motorist fatalities. Similarly, research in Pennsylvania shows alcohol-related collisions often occur at night and 94% of the persons killed in such crashes were in the vehicle (PA DUI Association).

We recommend the Department of Transportation to increase education on drunk driving targetted at drivers and staff at places serving alcohol at night. Additionally, cars can include designs to detect drunk drivers and stop them from driving. Cameras on the road can be programmed to identify dangerous driving patterns and alert the police.

**Acknowledgments** I want to thank Northeast Big Data Innovation Hub, National Student Data Corps, and the U.S. Department of Transportation Federal Highway Administration for giving me this opportunity to participate in the project and learn methods of data analysis, especialy time-series and geospatial analysis.

## References

Explorer Transportation Data Science Project Colab Notebook, NYC Open Data, PA DUI Association