

Analysis of Motor Vehicle Collision Crashes in New York City

Cicily Balachandar

May 1, 2022

Abstract

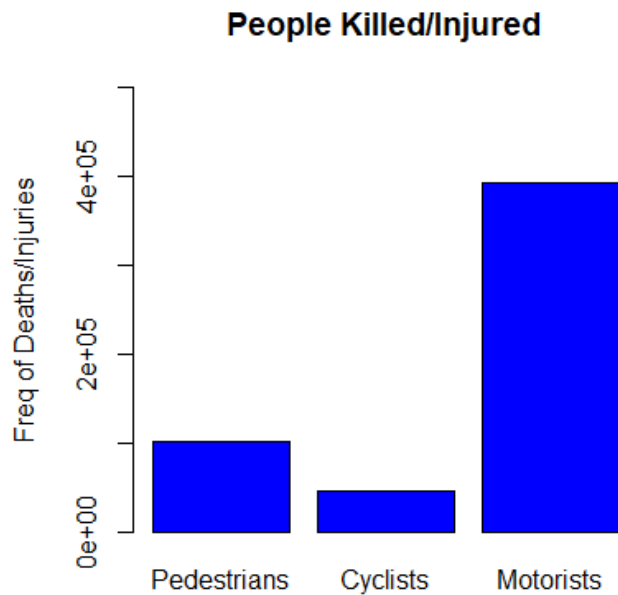
There seems to be some differences between the pedestrian, cyclist, and motorist groups. These differences are more pronounced in distinguishing between injuries and deaths, as well as at different times of the year.

In particular, it seems like pedestrians are more likely to be killed than motorists when involved in an accident, which is a contrast from motorists being more likely to sustain an injury compared to pedestrians.

Then, when comparing the groups by month, emerging trends are noticeable amongst the pedestrian and cyclist groups. More specifically, there seems to be higher frequencies of injuries and deaths in the pedestrian group during the colder months, while there are higher frequencies of injuries and deaths in the cyclist group during the warmer months.

Overall, it might be interesting to visually see the comparison between pedestrians, cyclists, and motorists, in terms of how they are affected by being involved in accidents.

Figure 1. Total Number People Killed or Injured for Each Group



In Figure 1, the frequency of death and injury in the motorists is much higher compared to the other two groups. Pedestrians have a higher frequency compared to cyclists. In Figure 2 below, this same observation holds. However, in Figure 3, more pedestrians are killed in comparison to the other two groups. This makes sense since pedestrians have little to no protection compared to people in vehicles.

Figure 2. Frequency of People Injured for Each Group

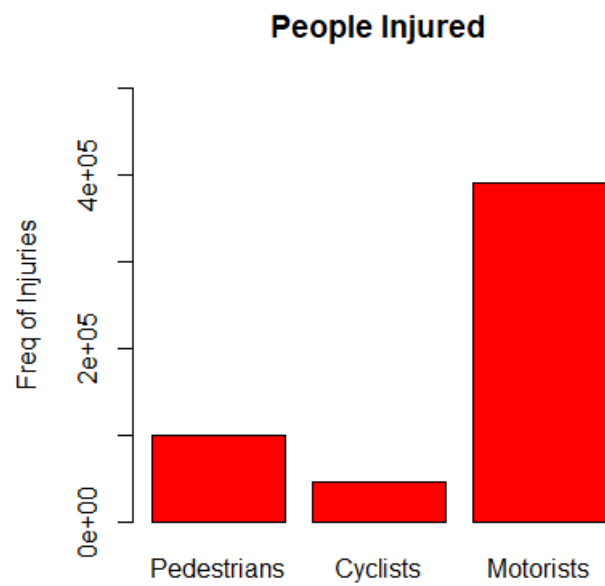
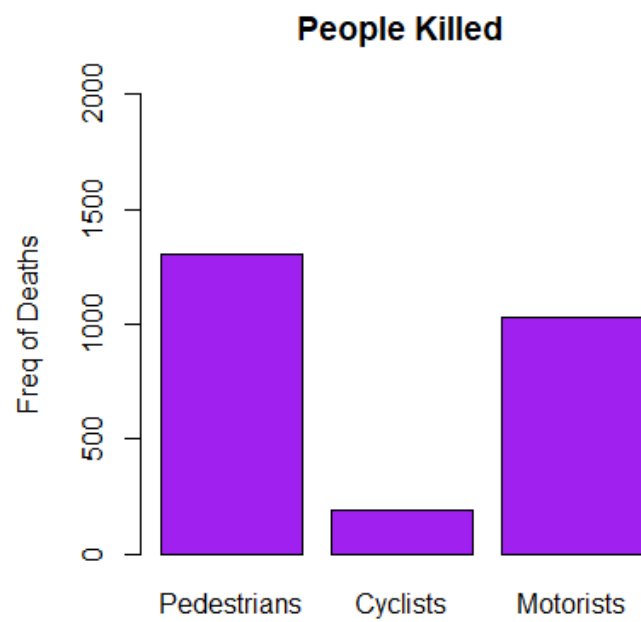
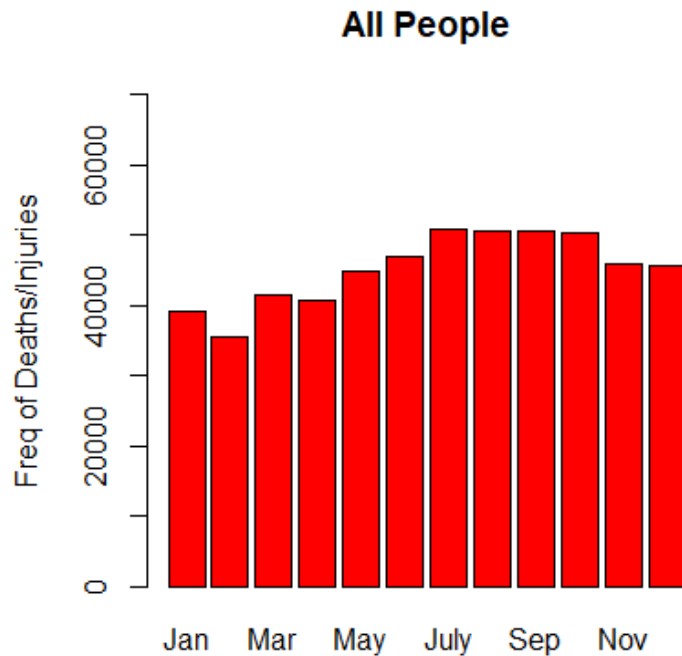


Figure 3. Frequency of Deaths in Each Group



Next, grouping the data by month for each of the groups can be useful.

Figure 4. Frequency of Deaths/Injuries in All Groups by Month



In Figure 4 above, there does not seem to be much of a difference in the total frequencies of deaths and injuries for each month. Then, when looking at Figure 5 and Figure 6, there seems to be a much more noticeable difference in the pedestrian and cyclist groups.

In particular, there seems to be a decrease in injuries and deaths between March and September for pedestrians, while there is an increase in injuries and deaths in the same time period for cyclists. Considering the weather conditions, for pedestrians, they might be more impacted by accidents during colder months due to poor driving conditions, like snow and ice. For cyclists, this pattern might be because people are less likely to use their bikes during the winter, so there will most likely be higher numbers of injuries and accidents in the warmer months for this group since more people tend to bike in warmer conditions.

Figure 5. Frequency of Deaths/Injuries in Pedestrians by Month

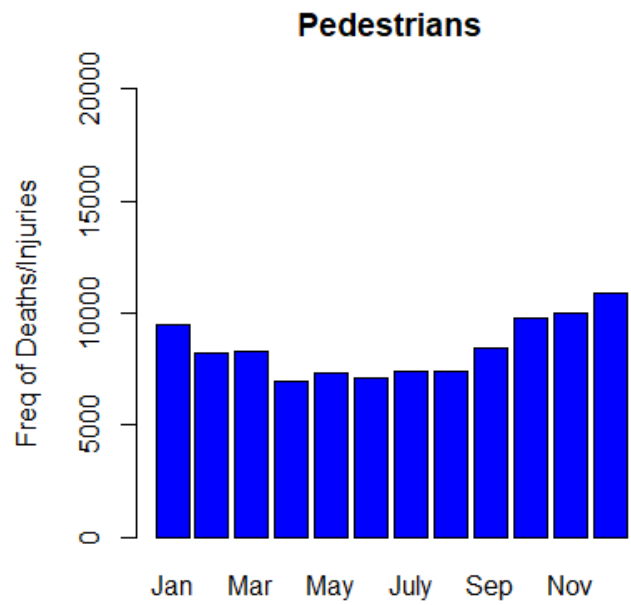
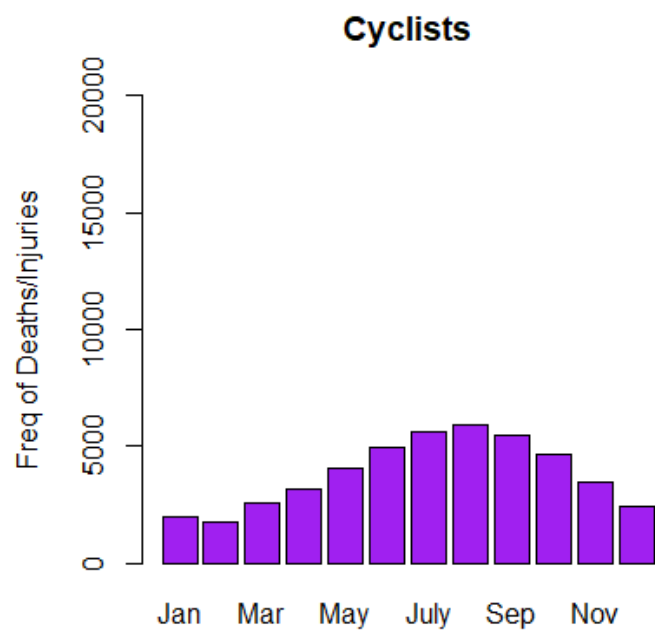
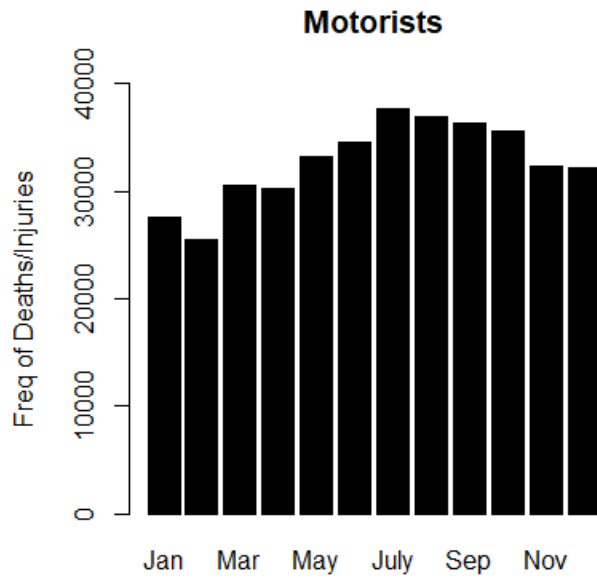


Figure 6. Frequency of Deaths/Injuries in Cyclists by Month



In Figure 7, the frequency of deaths and injuries are high at all months for motorists.

Figure 7. Frequency of Deaths/Injuries in Motorists by Month



Furthermore, the different contributing factors of collision crashes can be considered when comparing the three groups. The three main contributing factors focused on here will be distracted or inattentive driving, alcohol, and fatigue/drowsiness. In Figures 8-11, it seems like distracted driving has the highest frequency of injuries and deaths, out of the other two contributing factors, for each of the groups.

The alcohol and fatigue categories seem to be around the same. Although, there is a slight noticeable difference between these factors in the motorist group. Overall, distracted driving seems to be one of the more concerning contributing factors, among each group.

Figure 8. Frequency of People Injured for Different Contributing Factors

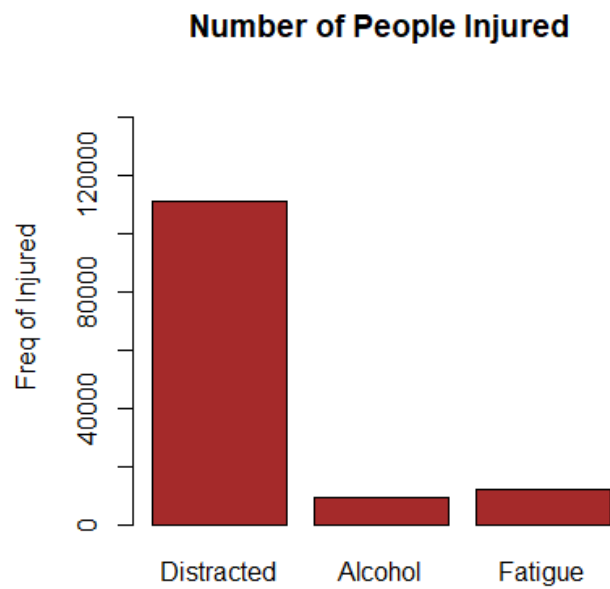


Figure 9. Frequency of Pedestrians Injured for Different Contributing Factors

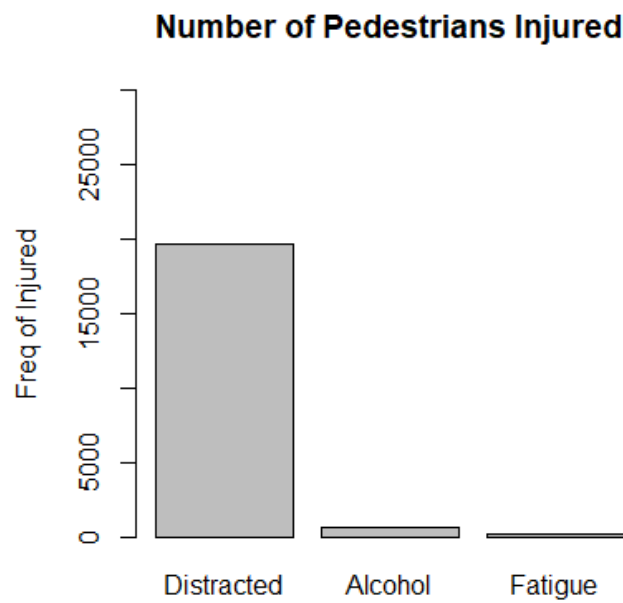


Figure 10. Frequency of Cyclists Injured for Different Contributing Factors

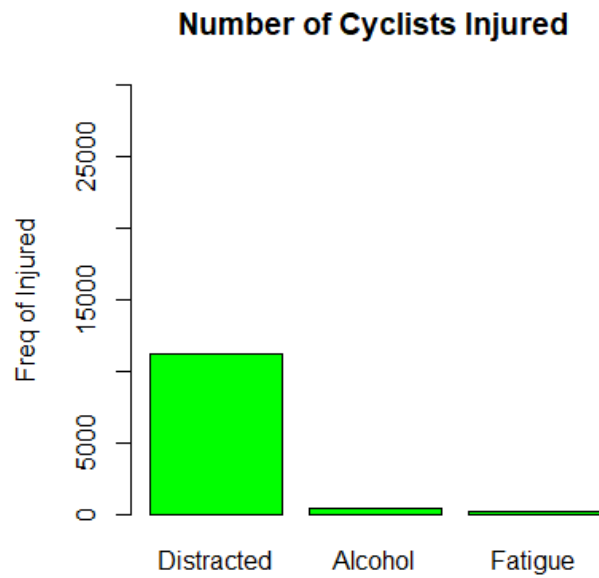
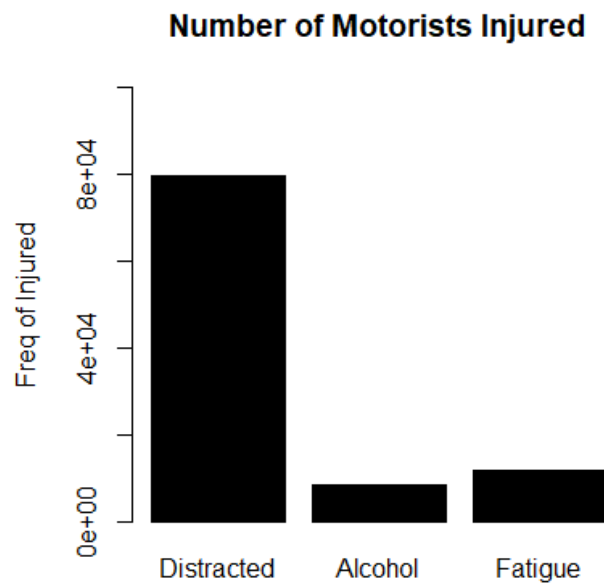


Figure 11. Frequency of Motorists Injured for Different Contributing Factors



In general, it is important to investigate these questions to see trends. More specifically, more information relating to the time of year is helpful when analyzing the data and it will be easy to see what needs to be done during months of higher injury and death rates to prevent accidents.

Appendix A

The dataset only includes columns relating to injuries and deaths. In addition, the time, date, and Contributing Factor 1 column were included in the dataset. In terms of data management, different subsets were created to combine the injury and death columns for each of Persons, Pedestrians, Cyclists, and Motorists groups, and also to group the data by month.

References

- Blog, R. A. (2020, April 17). *A comprehensive introduction to handling date & time in R: R-bloggers*. R. Retrieved May 1, 2022, from <https://www.r-bloggers.com/2020/04/a-comprehensive-introduction-to-handling-date-time-in-r/#:~:text=Summary%20%20%20Function%20%20%20,%20Current%20Time%20%203%20more%20rows%20>
- Dates and Times in R*. Dates and times in R. (n.d.). Retrieved May 1, 2022, from <https://www.stat.berkeley.edu/~s133/dates.html#:~:text=R%20provides%20several%20options%20for%20dealing%20with%20date,dates%20and%20times%20with%20control%20for%20time%20zones>.

