Investigating motor vehicle collisions in Toronto: Bad weather not a major cause of collisions*

Jierui Miao

2024-01-25

The impact of severe weather on traffic crashes is an ongoing concern, especially during the winter months when snow falls and spring rains are frequent, and vehicle collisions sometimes result in serious public and personal property damage, injuries and fatalities. The purpose of this report is to investigate motor vehicle collisions resulting in injuries and fatalities in Toronto from 2006 to 2022. The study showed no significant association between inclement weather and collision risk. This finding suggests that inclement weather is not a major contributing factor to crashes, and that when discussing collision avoidance, consideration should be given to the fact that improving roadway infrastructure, enhancing education for new drivers, and fostering a culture of safe driving can further reduce collision risk.

Table of contents

*Code and data are available at: https://github.com/jerry-maker-765/vehicle-collisions.git

Introduction

The impact of inclement weather on traffic crashes is an ongoing concern in the Toronto area and in many northern regions, especially during the winter months when snowfall and spring rains are frequent. Previous studies have shown that winter precipitation is an important environmental factor in increasing the frequency of motor vehicle collisions, resulting in personal injury and property damage. (Mills et al. 2019) Rain can make roadways slippery, making vehicles less manageable, and poor visibility in inclement weather can affect a driver's ability to analyze roadway conditions. And for a densely populated and heavily traveled area like the Toronto area, the frequency of road traffic accidents can lead to dangerous trips and longer commutes for everyone, as well as financial losses and injuries. The Global Burden of Disease Study, conducted by the World Health Organization (WHO), Harvard University and the World Bank, showed that in 1990, traffic accidents were assessed to be the world's ninth leading health problem. The study predicts that by 2020, road traffic crashes will rise to third place among the leading causes of death and disability globally (Khaleghei Ghosheh Balagh, Naderkhani, and Makis 2014).

The Toronto Motor Vehicle Crash Data documents vehicle crashes in Toronto involving injuries and fatalities since 2006. I analyzed this data set statistically, controlling for other variables that could affect the results, graphing the factors taken into account and analyzing what was obtained from the graphs, with the goal of offering the possibility of reducing the occurrence of such accidents. In Section @ref(about-motor-vehicle-collisions-involving-killed-or-seriously-injured-persons), I introduce the motor vehicle collision dataset and organized it to visualize changes in area and roadway information about traffic crashes from year to year. I then discuss the possible implications of these visualized data. Through Figure 4 I conclude that Toronto and East York have the most accidents and Scarborough has the highest fatality rate. With Figure 3 I conclude that more collisions occur at night than during the day. Combined with Figure 1 which states that more collisions occur on sunny days, Figure 2 which states that there are more non-fatal collisions than fatal collisions, and that the largest percentage of collisions occur on dry pavement. I conclude that there is no significant correlation between inclement weather and collision risk.More implications and further discussions on this result is presented in Section @ref(discussion).

Data

About Motor Vehicle Collisions involving Killed or Seriously Injured Persons

In order to investigate motor vehicle collisions in Toronto, I downloaded the "motor vehicle collisions involving fatalities or serious injuries" dataset from torontoopendata (Gelfand 2020) The "Motor Vehicle Collisions Involving Fatal or Seriously Injured Persons" dataset was downloaded from it. This dataset was provided by the Toronto Police Service and includes all

vehicle collisions resulting in injury or death in Toronto between 2006 and 2022.The location of crime occurrences have been deliberately offset to the nearest road intersection node to protect the privacy of parties involved in the occurrence.(Services 2021) Therefore this data will not involve personal privacy. And it will not affect the result of my REPORT because this REPORT does not need the intersection data.

The dataset contains 16840 collision observations and 54 variables that contain various information about each collision. Because this report investigates the effects of severe weather, only information for January, February, November, and December of each year has been retained here to avoid bias from not having snow in the summer. This report focuses on six of these variables, year, district, light, visibility, road condition, and fatality. By using R (R Core Team 2020), and R packages "tidyverse" (Wickham et al. 2019), "dplyr" (Wickham et al. 2021), "janitor" (Firke 2021), "kableExtra" (Zhu 2021) and "knitr" (Xie 2021), an extract of the cleaned dataset (Table @ref(tab:dataextract)) is shown below.

Table 1: Random select ten rows from the Motor Vehicle Collisions data

Year	District	Light	Visibility	Road condition	Fatality
2006	Scarborough	Daylight	Clear	Dry	Non-Fatal Injury
2006	North York	Dark	Clear	Dry	Non-Fatal Injury
2006	North York	Dark	Clear	Dry	Non-Fatal Injury
2006	Toronto and East York	Dark	Clear	Dry	Non-Fatal Injury
2006	Toronto and East York	Dark	Clear	Dry	Non-Fatal Injury
2006	North York	Dark	Rain	Wet	Non-Fatal Injury
2006	North York	Dark	Rain	Wet	Non-Fatal Injury
2006	Toronto and East York	Dark	Rain	Wet	Non-Fatal Injury
2006	Toronto and East York	Dark	Rain	Wet	Non-Fatal Injury
2006	Toronto and East York	Dark	Rain	Wet	Non-Fatal Injury
2006	Toronto and East York	Dark	Clear	Dry	Non-Fatal Injury

Table @ref(tab:dataextract) shows ten random rows of Motor Vehicle Collisions data. The variable "Year" represents the year in which the accident occurred. The variable "District" represents the district where the accident occurred. The variable "Light" represents the brightness of the light at the time. The variable "visibility" represents the environment condition, e.g. "Clear", "Rain". The variable "road_condition" represents the road condition, e.g. "dry", "wet", etc. The variable "fatality" involves fatal conditions. "Fatality" involves fatal and non-fatal injury. I was curious about the effects of bad weather on traffic accidents. So for the next report I wanted to look at the trends in the percentage of motor vehicle crashes involving rain, snow, and limited light levels from 2006-2022 winter. It reveals whether or not Toronto has improved road safety management over these years to reduce these percentages and thus reduce the number of traffic accidents that should not be happening.

Figure 1

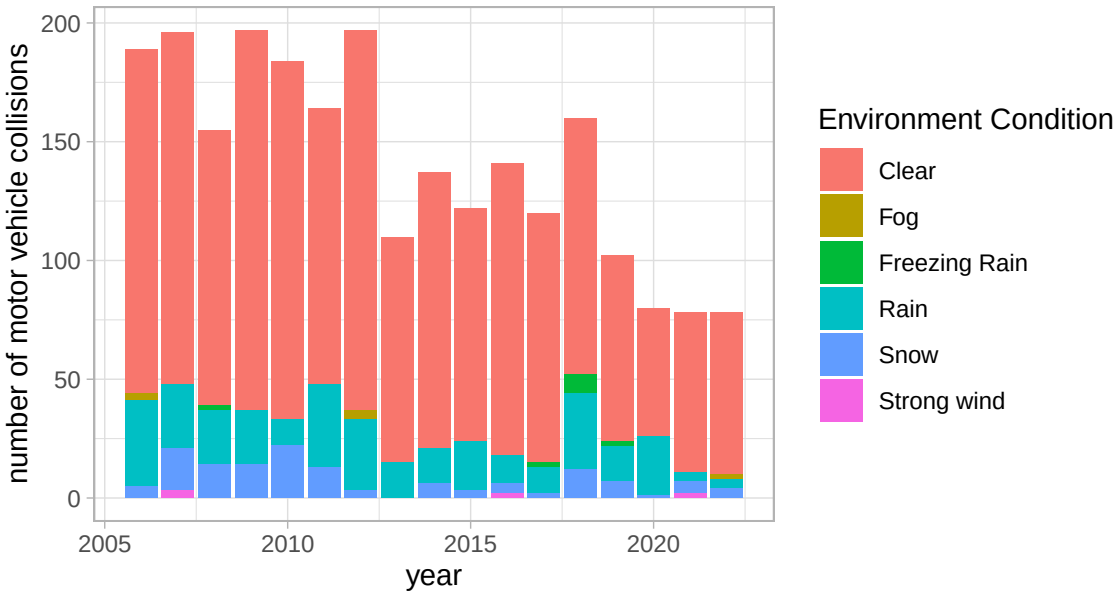

Figure 1: Comparing the Environment Condition of collisions each year from 2006 to 2022

Figure 1 illustrates the number of motor vehicle collisions in the Toronto region for several years between 2006 and 2022, categorized by different environmental conditions such as sunny, foggy, freezing rain, rain, snow, and high winds. As you can see in the graph, the higher percentage of collisions that occurred in 2018 due to rain, snow or freezing rain compared to other years is due to the frequent snowstorms and freezing rain that hit the Toronto area in 2018. When severe weather is frequent, collisions also occur frequently, so collisions and environmental condition are related. The interesting point is that even in the winter when rain and snow are frequent most collisions still happen on sunny days. Rain and snow do cause a large number of crashes, but not as many as on sunny days. In this type of weather drivers usually drive at lower speeds, pay more attention than usual, and there are few cars on the road. In sunny weather, the roads tend to be busy and drivers do not have the same high level of concentration as in rain and snow.

Figure 2

Figure 2 illustrates a comparison of the number of fatal and non-fatal crashes after categorization by road surface condition. This figure supports the point made in Figure 1 that slippery pavement affects the number of crashes, but the highest number of crashes occurs on dry pavement, in sunny weather, and regardless of whether it is a serious crash or not, fatal crashes do not increase and outnumber non-fatal crashes due to the effects of inclement weather.

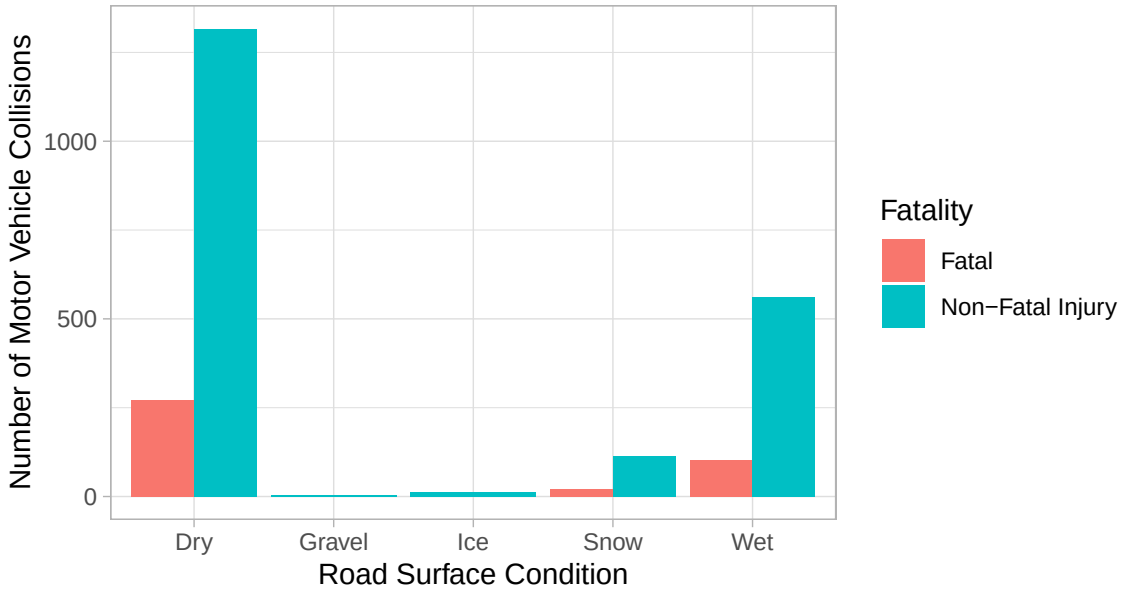

Figure 2: Comparing the Road Surface Condition of collisions by fatality

Figure 3

Figure 3 records the number of motor vehicle crashes from 2006 through 2022, categorized by lighting conditions (dark, dawn, daylight, and dusk) at the time of the crash. As can be seen in the figure, there are many collisions during both daytime and nighttime. During the daytime the roads are busy and densely populated and collisions can be generated by congestion. Collisions occur most often in the dark. Where there is not enough light from the street lamps there are details on the road that drivers don't notice such as dark ice, potholes, and small animals. These can cause a vehicle to lose control and cause an accident. But the overall downward trend in this group of data can be attributed to the gradual improvement of vehicle safety technology breakthroughs, as well as road repair, such as the installation of street lights and fences, repair the road surface. And the data set of dawn and dusk has remained very low, two time periods when the frequency of collisions is not high.

Figure 4

Figure 4 illustrates a visual comparison of motor vehicle collisions by region in Toronto, divided into fatal and non-fatal injury categories. The data shows that Toronto and East York have the highest frequency of collisions, which can be attributed to the dense traffic volumes in these areas as well as the large population base. These areas are characterized by a complex network of intersections and high levels of pedestrian activity, all of which contribute to increased collision rates.

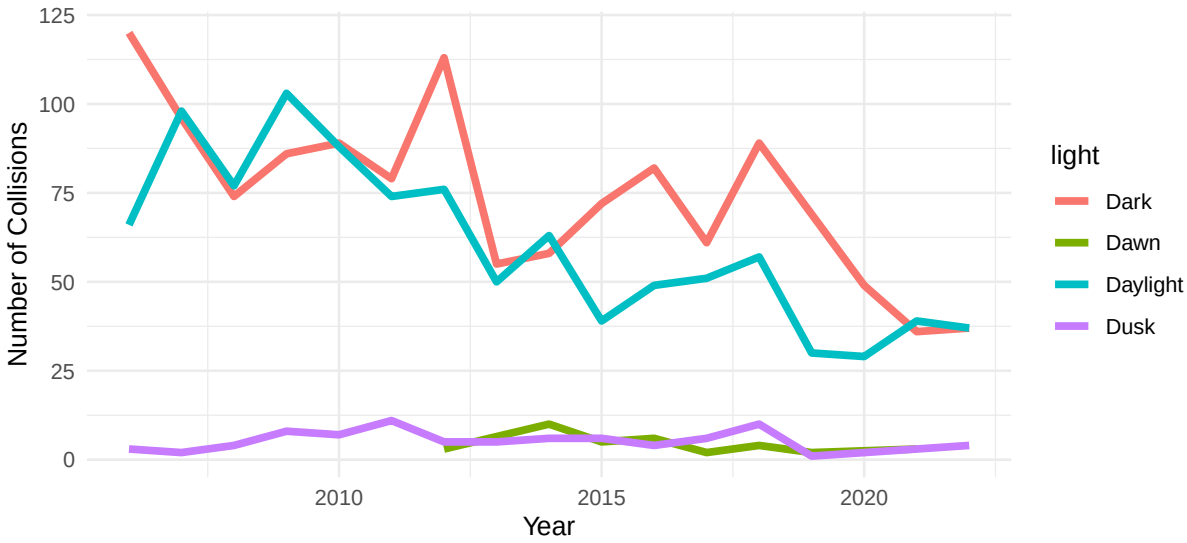

Figure 3: Comparing the Light Condition of collisions each year from 2006 to 2020

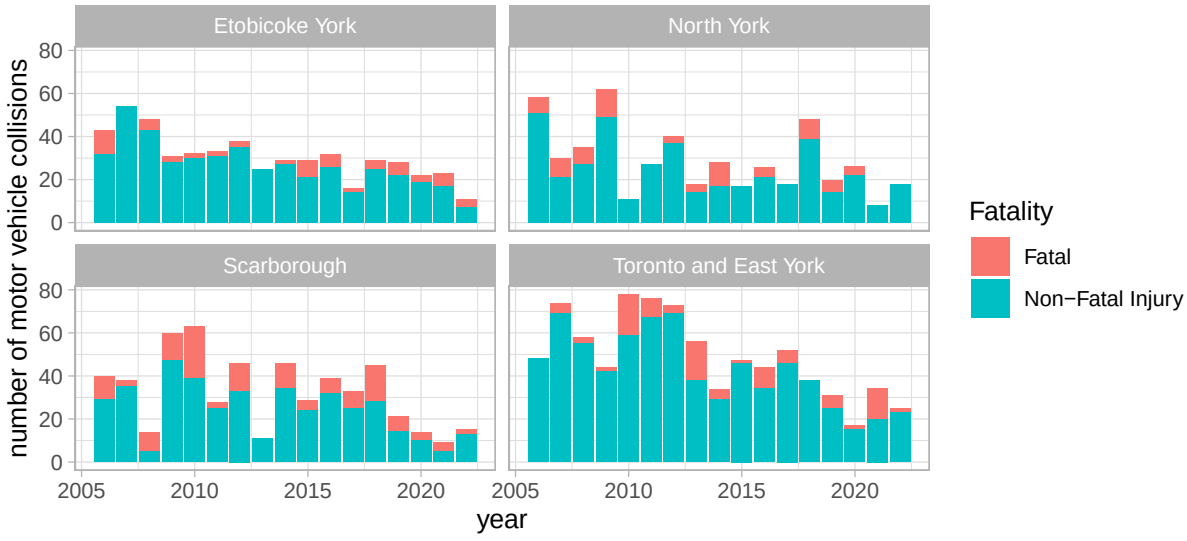

Figure 4: Comparing the number of motor vehicle collisions take place in different districts

North York had no fatalities in some years of data, Etobicoke has had the lowest fatality rate for many years, and that collisions have steadily declined suggests that effective traffic management and road safety measures are being implemented in the area. The steady decline is indicative of ongoing improvements, including better road design, enhanced traffic signal systems, or strict enforcement of traffic laws.

In contrast, Scarborough has a high fatality rate. Historically, the region's infrastructure has not been as robust, with roads in need of repair and safety features that do not meet the standards of the center region. However, recent initiatives aimed at strengthening the infrastructure appear to be bearing fruit, as indicated by a significant reduction in the number of collisions. These efforts may include road resurfacing, modernized traffic control devices, and improved street lighting, all of which are critical to ensuring the safety of both drivers and pedestrians.

Discussion

Some limitations should be noted when interpreting the results of this study. First of all, it cannot be assumed that the results of this study are applicable to all situations because we did not make a statistical survey of the drivers' driving age, and those who have less driving experience will be more unfamiliar with the control of the road surface, and they don't have a lot of experience in dealing with complex situations when the road conditions are poor. These results are based on cross-sectional surveys and therefore cannot be used to draw causal conclusions. Nonetheless, the current analysis suggests that there is no significant correlation between inclement weather and crash risk.

Analysis of this data set suggests that bad weather is not a major cause of collisions. Although bad weather and poor road conditions increase the likelihood of collisions, collisions are most likely to occur in sunny weather conditions and on dry roads. As you can see from figure 1 and figure 2, there is no year where the number of collisions had a higher percentage of bad weather factors than normal weather, and even in 2018, which was hit by frequent snowstorms, the number of collisions under sunny skies was still a bit higher. Then we need to consider other causes of accidents, such as the inexperience of young drivers and the quality level of drivers. Of course there are also external factors including poorly maintained roads, substandard street lighting, numerous potholes and certain speed limits which may be impractical in specific situations.

In light of these findings, measures aimed at reducing the likelihood of accidents must be taken. Initiatives may include enhanced driver education programs, especially for new drivers, focusing on defensive driving techniques for different weather and road conditions. Improvements in roadway infrastructure - such as better pavement, more effective street lighting, timely pothole repair, and review of speed limit policies - may also be critical. In addition, fostering a safe driving culture through public awareness campaigns could further reduce crash risk.

In conclusion, while this study depicts some patterns of collision occurrence in different weather and road conditions, it also emphasizes the multifaceted nature of road safety and the need for a comprehensive approach to accident prevention. Future research should aim to incorporate a wider range of variables, including driver demographics and behavioral factors, to build a more conclusive understanding of the determinants of road safety.

References

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://sharlagelfand.github.io/opendatatoronto/, https://github.com/sharlagelfand/opendatatoronto/.

Khaleghei Ghosheh Balagh, Akram, Farnoosh Naderkhani, and Viliam Makis. 2014. "Highway Accident Modeling and Forecasting in Winter." *Transportation Research Part A: Policy and Practice* 59: 384–96. https://doi.org/https://doi.org/10.1016/j.tra.2013.10.008.

Mills, Brian, Jean Andrey, Brent Doberstein, Sean Doherty, and Jennifer Yessis. 2019. "Changing Patterns of Motor Vehicle Collision Risk During Winter Storms: A New Look at a Pervasive Problem." *Accident Analysis & Prevention* 127: 186–97. https://doi.org/https://doi.org/10.1016/j.aap.2019.02.027.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Services, Toronto Police. 2021. *Motor Vehicle Collisions Involving Killed or Seriously Injured Persons*. https://open.toronto.ca/dataset/motor-vehicle-collisions-involving-killed-or-seriously-injured-persons/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. http://haozhu233.github.io/kableExtra/, https://github.com/haozhu233/kableExtra.