Traffic Volumes at Intersections Affects Traffic Collisions in Toronto*

Wentao Sun

25 January 2024

The City of Toronto, due to its busy intersections and high traffic volumes, is often involved in vehicle collisions that not only have a significant impact on the city's transportation system, but also pose a threat to drivers and pedestrians. This report focuses on investigating traffic collisions that occurred between January 2020 and September 2023 in the City of Toronto. By analyzing the frequency, severity, and specific vehicle types and locations of occurrence of traffic crashes during this time period, we aim to reveal the correlation between traffic flow and the occurrence of crashes. Despite the fact that the City of Toronto has taken several measures to improve road safety, the number of traffic crashes remains high on high volume roads during certain peak hours. This phenomenon suggests that better safety measures should be put in place to address the traffic safety risks associated with high traffic volumes. Data on high-traffic roads and traffic collisions are further analyzed and utilized.

Table of contents

*Code and data supporting this analysis is available at: https://github.com/TonySun1107/STA302H1.git

1 Introduction

Toronto is one of Canada's most cosmopolitan and modern cities. However, a large population inevitably brings with it a host of transportation problems. Increased car ownership has made Toronto's roads busy. During the morning and evening rush hours, the influx of people and vehicles slows down vehicle traffic.The result of a case study of Toronto, Canada exhibits that entertainment related activities are more likely to appear during evening peak hours(Huang et al. 2019). Heavy traffic may increase the number of safety hazards and increase the likelihood of vehicle collisions. Vehicle collisions can jeopardize the lives of drivers and pedestrians.

Toronto traffic collision data records crashes involving different types of vehicles in a given neighborhood, where the data covers the specific number of injuries or fatalities and the type of accident. After analyzing the data, it is possible to summarize the content of the data to get conclusions as a way to reduce the number of traffic accidents in Toronto. This paper investigates the specific causes of traffic collisions at different times of the day from 2020 to 2023, focusing on whether or not they are hit-and-run and whether or not they cause material damage.

2 Data

The data used in this paper were obtained from Toronto's Open Data Portal by the library (Gelfand 2022). The dataset was processed using R, a computer language designed for statistical computing and graphics (R Core Team 2023), which included stages of downloading, cleaning, analysis, and visualization. And using some support package from 'dplyer'(Wickham et al. 2023), 'tidyverse'(Wickham et al. 2019), 'knitr'(Xie 2023), 'janitor'(Firke 2023), 'here'(Müller 2020), 'kableExtra'(Zhu 2021), 'ggplot2'(Wickham 2016).

2.1 Police Statistical Report - Traffic Collisions

In order to investigate motor vehicle collisions in Toronto, I obtained the "Police Statistical Report - Traffic Collisions" dataset from the Toronto Open Data Portal(Gelfand 2022). This data includes all traffic collision records from December 2020 to 2023. The dataset consists of 257,303 observations traffic collisions samples.

Table 1 below shows the top fifteen rows of traffic collisions. The variable "Year" indicates the year in which the accident occurred. The variable "Hour" indicates the time of day on which the accident occurred. The variable "Neighbourhood" provides the neighborhood where the accident occurred. The variable "FTR_Collisions" usually stands for "Failure to Remain", which means whether the driver fled the scene of the crash. The variable "PD_Collisions"

Table 1: Extract fifteen rows from the traffic collisions

Year	Month	Hour	Neighbourhood	FTR_Collisions	PD_Collisions
2020	January	9	Woburn North (142)	NO	YES
2020	January	2	West Humber-Clairville (1)	YES	NO
2020	January	11	Kennedy Park (124)	NO	YES
2020	January	3	Glenfield-Jane Heights (25)	YES	NO
2020	January	2	Newtonbrook West (36)	NO	YES
2020	January	1	East End-Danforth (62)	NO	YES
2020	January	11	Mount Dennis (115)	NO	YES
2020	January	1	Wellington Place (164)	YES	NO
2020	January	1	Newtonbrook West (36)	NO	YES
2020	January	12	Highland Creek (134)	NO	NO
2020	January	13	York University Heights (27)	YES	NO
2020	January	11	Englemount-Lawrence (32)	YES	NO
2020	January	9	Palmerston-Little Italy (80)	YES	NO
2020	January	4	High Park North (88)	YES	NO
2020	January	12	Bayview Village (52)	NO	YES

usually stands for "Property Damage", which means whether the accident resulted in physical damage. Whether the traffic collision involved damage to vehicles, buildings, roadway infrastructure, etc.

The main focus from the traffic collision data is on the rate of accidents at different times of the day and the extent of accident damage from 2020 to 2023. These data are in a position to prove that there is a relationship between the number of traffic collisions and the time of day. It also responds to the traffic situation in Toronto so that the government can develop different traffic controls to protect the safety of the people on the road. Reduce the harm brought to people by traffic collisions in the future.

2.2 Traffic Volumes at Intersections for All Modes

In order to investigate motor vehicle collisions in Toronto, I obtained the "Traffic Volumes at Intersections for All Modes" dataset from the Toronto Open Data Portal(Gelfand 2022). This data includes traffic and pedestrian volumes in four different general directions at various times of the day from 2020 to 2023. The dataset consists of 108,310 observations traffic and pedestrian volumes samples.

By summarizing and analyzing the traffic Volume in different time periods through this dateset, a 24-hour day can be roughly divided into morning and evening peak traffic hours and other ordinary hours. Since activities related to commuting and recreation are more likely to occur during the evening rush hour, usually more vehicles and pedestrians enter the city traffic at

Table 2: Extract ten rows from the traffic volumes

Time Start	Time End	South car	North car	West car	East car	Pedestrians
2020-01-15 07:30:00	2020-01-15 07:45:00	45	39	6	16	4
2020-01-15 07:45:00	2020-01-15 08:00:00	66	42	5	26	1
2020-01-15 08:00:00	2020-01-15 08:15:00	64	43	7	27	3
2020-01-15 08:15:00	2020-01-15 08:30:00	87	53	3	40	3
2020-01-15 08:30:00	2020-01-15 08:45:00	85	45	6	37	2
2020-01-15 08:45:00	2020-01-15 09:00:00	82	46	6	38	2
2020-01-15 09:00:00	2020-01-15 09:15:00	88	46	12	30	1
2020-01-15 09:15:00	2020-01-15 09:30:00	66	38	11	27	5
2020-01-15 10:00:00	2020-01-15 10:15:00	68	20	9	13	3
2020-01-15 10:15:00	2020-01-15 10:30:00	67	39	4	12	0

this time than usual. After that, it is easy to find the relationship between the rate of vehicle collisions and the morning and evening peak hours. The hypothesis can then be tested.

Table 2 shows the first ten rows of traffic flow. The variable "Time Start" indicates the start time of the recorded traffic. The variable "Time End" records the end time of the traffic situation. Splits the day into many smaller time periods. Each time period displays 15 minutes of the traffic situation. Traffic flow tends to increase or decrease rapidly, so a 15-minute time period provides a better view of changes. The variable "South car" indicates the number of cars traveling southbound. The variable "North car" indicates the number of cars traveling northbound. The variable "West car" indicates the number of cars traveling weatbound. The variable "East car" indicates the number of cars traveling eastbound. The variable "pedestrians" represents the number of pedestrians within the statistical area.

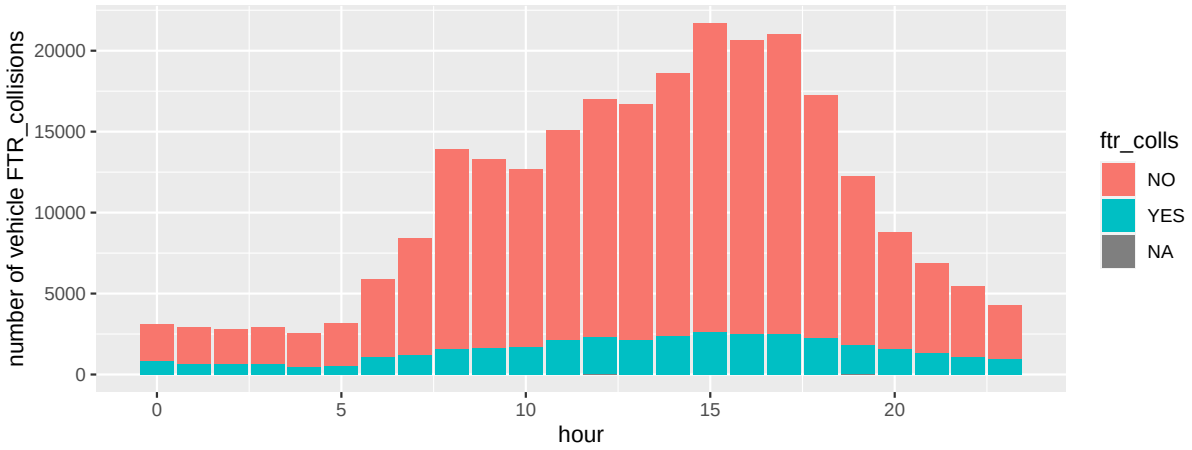

Figure 1: Comparing the number of traffic FTR_collisions each hour from 2020 to 2023

Figure 1 shows the number of vehicle crashes at different times of the day from 2020 through

2023. Also shown are the number of hit-and-run, no hit-and-run, and undetermined collisions. Because there is not a simpler and more efficient way to record and store the data, it is often incorrect or lost. The majority of vehicle collisions occur between the hours of 3:00 PM and 6:00 PM. During times of heavy traffic, drivers may commit more rule violations. Some drivers may disobey traffic rules and arbitrarily change lanes of travel because they are in a hurry. They will rarely even yield the right of way for other pedestrians or vehicles. These behaviors not only increase the probability of vehicle collisions, but also largely increase the danger to pedestrians. With this figure we have no evidence to reject that peak vehicle times increase the risk of vehicle collisions.

3 Discussion

Analysis of the two datasets showed a relationship between vehicle crashes in Toronto and traffic volumes during the morning and evening rush hours. The different time periods bring different data can be seen in Figure 1. The comparison shows that the number of collisions is significantly higher during the evening rush hour off-duty hours than at other times. The implication in section 2.1 is to summarize and analyze the vehicle collision information from 2020 to 2023, which includes a variety of information such as different times of day, different vehicles, and different collision locations. It is not well illustrated by this information that collisions are directly related to time of day, and better evidence is needed to support this idea. The implication in section 2.2 is that it was determined that there is an increase in traffic and pedestrian traffic during the afternoon off-work hours and that too many vehicles can affect the number of crashes. A likely reason to explain this is that the increase in vehicles makes it more difficult for drivers to drive. Collisions and scrapes are more likely to occur on narrow roads.

The Government should take measures to improve the safety of this transportation system and ensure the rights and interests of driving as a driver and the safety of pedestrians. For example, it should increase traffic control and rationalize the distribution of each trunk road. This is because most of the vehicles' routes at the end of the day are from office areas to residential areas. Assigning different roads by license plate number will reduce the number of cars on the main roads to protect drivers and pedestrians, even though some of them are not on the best routes. The government should try different methods of traffic control and observe the results in the following years to determine a safer policy.

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Huang, Wei, Shishuo Xu, Yingwei Yan, and Alexander Zipf. 2019. "An Exploration of the Interaction Between Urban Human Activities and Daily Traffic Conditions: A Case Study of Toronto, Canada." https://www.sciencedirect.com/science/article/abs/pii/S0264275118302786.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://here.r-lib.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

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, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2023. *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/.