Investigating motor vehicle collisions in Toronto: Installing more red light cameras do not stop traffic accidents from happening*

Alicia Yang

05 February 2022

Abstract

Vehicle collisions occur constantly in Toronto and sometimes could be severe when resulting in damage in public and personal properties, injuries, or even death. This report aims to investigate motor vehicle collisions in Toronto that result in injuries or deaths from 2006 to 2020. By exploring the trend of these traffic collisions in respect of their frequency, fatality, and location, I find that the total number of accidents has a descending trend. However, the number of vehicle collisions that occurred at intersections has not reduced despite the fact that more red light cameras are installed. This finding suggests that unlike the claimed purpose of this electronic monitor, red light camera doesn't prevent traffic accidents from happening at intersection, and furthur measures should be taken to enhance the safety issue on road.

Contents

*Code and data are available at: https://github.com/Alicia-y/Telling-stories-with-data.

1 Introduction

As one of the busiest and most modern city in Canada, Toronto constantly faces a large number of citizens, and, inevitably, heavy traffic. The road condition in Toronto is busy most of the time. The vast majority of roads and buildings under construction that take place almost everywhere narrow the roadway of vehicles and turn off traffic lights (Webster 2022). Heavy traffic is likely to be one of the main reasons behind car accidents, because it increases the road rage and encourages the vehicle drivers to make dangerous moves to save time. Vehicle collisions have always brought large financial costs to the involving parties, and even more, they even brought injuries and deaths. In just 2019, traffic accidents resulted in more than 1700 deaths and more than 140000 injuries (Canada 2019).

The Motor Vehicle Collisions data of Toronto provides a record of the past severe vehicle collisions in Toronto which result in injuries or deaths, and with the help of statistical data analysis, it can possibly offer insights on the ways to reduce this type of accident. This report investigates the motor vehicle collisions in Toronto from 2006 to 2020, with a focus on the ones result in injuries or deaths. In Section 2, the Motor Vehicle Collisions dataset is introduced and organized to visualize the change in the annual number of traffic accidents with respect to their fatality and their location. Then, I discussed about the possible implications of these visualizations. By Figure 2, I concluded that motor vehicle collisions are always more likely to occur at intersections. Thus, Continuing

You can and should cross-reference sections and sub-sections. For instance, Section 2. R Markdown automatically makes the sections lower case and adds a dash to spaces to generate labels, for instance, Section **??**.

2 Data

2.1 Motor Vehicle Collisions involving Killed or Seriously Injured Persons

To investigate the motor vehicle collisions in Toronto, I obtained the dataset "Motor Vehicle Collisions involving Killed or Seriously Injured Persons" from the Toronto Open Data Portal (Gelfand 2020). Provided by the Toronto Police Service, this dataset includes all vehicle collisions occurred in Toronto between 2006 and 2020 that result in injuries or deaths. One bias aroused by this dataset is the offset from the real accident location. For ethical reason and the protection of the parties involved in the accidents that are considered to be criminal cases, the location of those criminal accidents are offset to the closest intersection node (Services 2021). As a result, the location of those accidents specific to neighborhoods might not be accurate. Therefore, I will use the variable district instead of neighbourhood to categorize the accidents later in the report to avoid this bias.

This dataset contains 16840 observations of collisions and 54 variables with all kinds of information of each collision. This report focuses on four of these variables which are: year, accident location, district, and fatality. By using R (R Core Team 2020), and R packages "tidyverse" (Wickham et al. 2019), "janitor" (Firke 2021), "kableExtra" (Zhu 2021) and "knitr" (Xie 2021), an extract of the cleaned dataset (Table 1) is shown below.

Table 1: Extracting the first ten rows from the Motor Vehicle Collisions data

Year	Collision location	District	Fatality
2006	At Intersection	Toronto and East York	Fatal
2006	At Intersection	Toronto and East York	Fatal
2006	At Intersection	Scarborough	Fatal
2006	At Intersection	Scarborough	Fatal
2006	NA	Scarborough	Fatal
2006	NA	Scarborough	Fatal
2006	NA	Scarborough	Fatal
2006	At Intersection	Toronto and East York	Non-Fatal Injury
2006	At Intersection	Toronto and East York	Non-Fatal Injury
2006	At Intersection	Etobicoke York	Fatal

Table 1 shows the first ten rows of the Motor Vehicle Collisions data. Variable "Year" indicates the year in which the accident occurs. Variable "District" provides the district of the accident. Variable "Collision location" includes 8 types: at intersection, intersection related, at/near private drive, laneway, non intersection, overpass or bridge, trail, and underpass or tunnel. Variable "Fatality" involves fatal and non-fatal injury.

I'm interested in finding the trend of motor vehicle collisions from 2006 to 2020 and the types of injuries they result in. It reveals the traffic control in Toronto, and possibly the quality and knowledge of vehicle owners throughout these years.

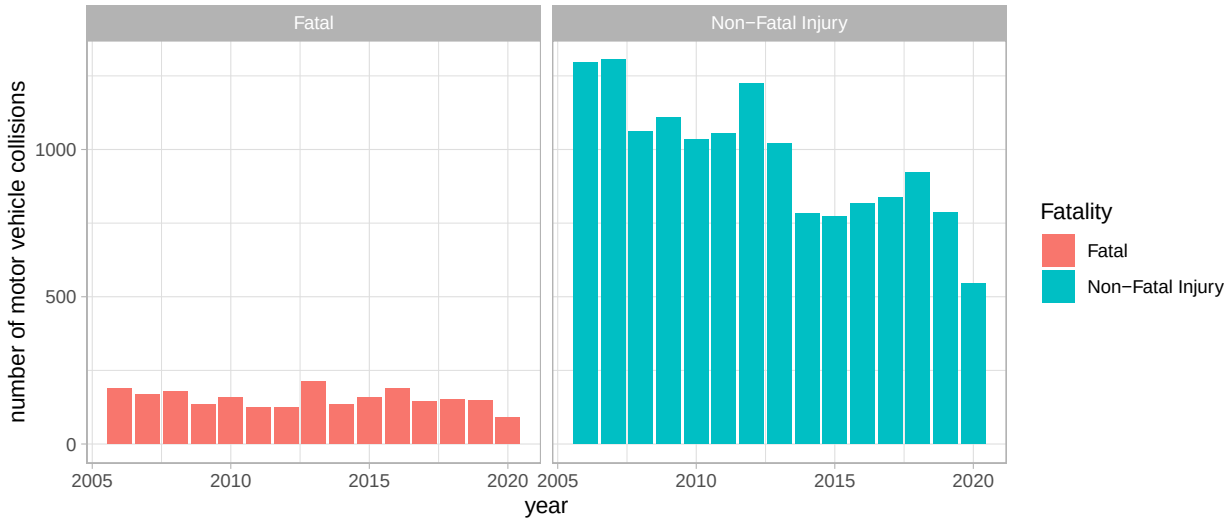

Figure 1: Comparing the number of traffic collisions each year from 2006 to 2020

Figure 1 shows a descending pattern for the number of car accidents involving non-fatal injuries from year 2006 to 2020. The number of fatal accidents does not increase nor decrease much. In general, the number vehicle collisions in Toronto has largely decrease, indicating that when combining this two factors(number and fatality), Toronto has done a better job on traffic control in general. I'm also interested in the trend of motor vehicle collisions if other factors are taken into considerations.

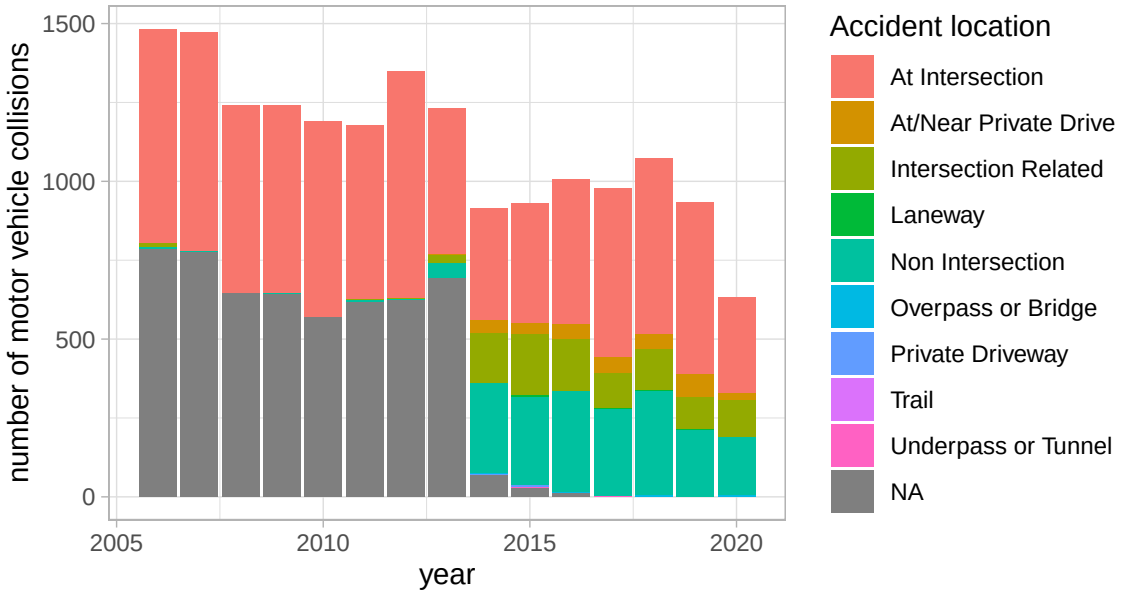

Figure 2: Comparing the location of traffic collisions each year from 2006 to 2020

Figure 2 illustrates the where these vehicle collisions usually take place from 2006 to 2020. Before 2013, half of the data doesn't come with accident location, which is understandable because back in the days there wasn't an easy approach to record and save the data, so they often got lost. A large proportion of the accidents take place either at intersections or is intersection related. The other seven accident locations types add up to less than half of all vehicle collisions. At intersections, a lot more rule breaking behaviors can be committed by drivers. For instance, they could sometimes ignore stop signs, speed up when approaching yellow traffic lights; some even claimed that they rarely yielded at intersections. These behaviors largely increase the possibility of them colliding with other vehicles and the pedestrians at intersections.

Motor vehicle collisions are most likely to take place at intersections. Thus, if measures are taken to restrict the driver's rule breaking behaviors at intersections, the numbers of car accidents would certainly drop a lot in general. Red light camera is claimed to be a measure taken by the government to monitor the vehicles at intersections. I will look into whether it's effective or not in respect of preventing the traffic accidents from happening.

2.2 Red Light Cameras

The number of red light camera instilled in Toronto from 2007 to 2020 is obtained from the dataset "Red Light Cameras" found in the Toronto Open Data Portal (Gelfand 2020). This dataset is shared by the Transportation Services and it covers the locations of red light cameras in Toronto. In its introduction, this dataset claims that red light cameras can improve the safety of the drivers and pedestrians by monitoring the behavior of the drivers (Services 2022). It contains 199 observations and 32 variables. This dataset's downside is that it is missing values in lots of its variables. To reduce errors generated from this downside, I will only utilize the variables with no missing values. R (R Core Team 2020), R packages "tidyverse" (Wickham et al. 2019), "janitor" (Firke 2021), "kableExtra" (Zhu 2021) and "knitr" (Xie 2021) are used to produce an extract of the cleaned dataset (Table 2).

Table 2: Extracting the first ten rows from the Red Light Camera data

District	Year	Activation date
Toronto and East York	2007	2007-11-09 05:00:00
Toronto and East York	2007	2007-11-09 05:00:00
North York	2007	2007-11-09 05:00:00
North York	2007	2007-11-09 05:00:00
Etobicoke York	2007	2007-11-09 05:00:00
Etobicoke York	2008	2008-03-18 04:00:00
Etobicoke York	2008	2008-10-28 04:00:00
North York	2009	2009-03-17 04:00:00
North York	2007	2007-11-09 05:00:00
North York	2008	2008-03-19 04:00:00

Table 2 shows the first ten rows of the Red Light Camera data in Toronto. "Year" is a new variable added by me into this dataset. It is pulled from the original variable "activation date", and it provides the year in which the red light camera is activated. Variable "District" includes the four main districts of Toronto: Toronto and East York, North York, Scarborough, and Etobicoke York.

I'm interested in observing the cumulative number of red light cameras installed in each district from 2006 to 2020 and compare this data with the trend of number of motor vehicle collisions to see if they actually have connections.

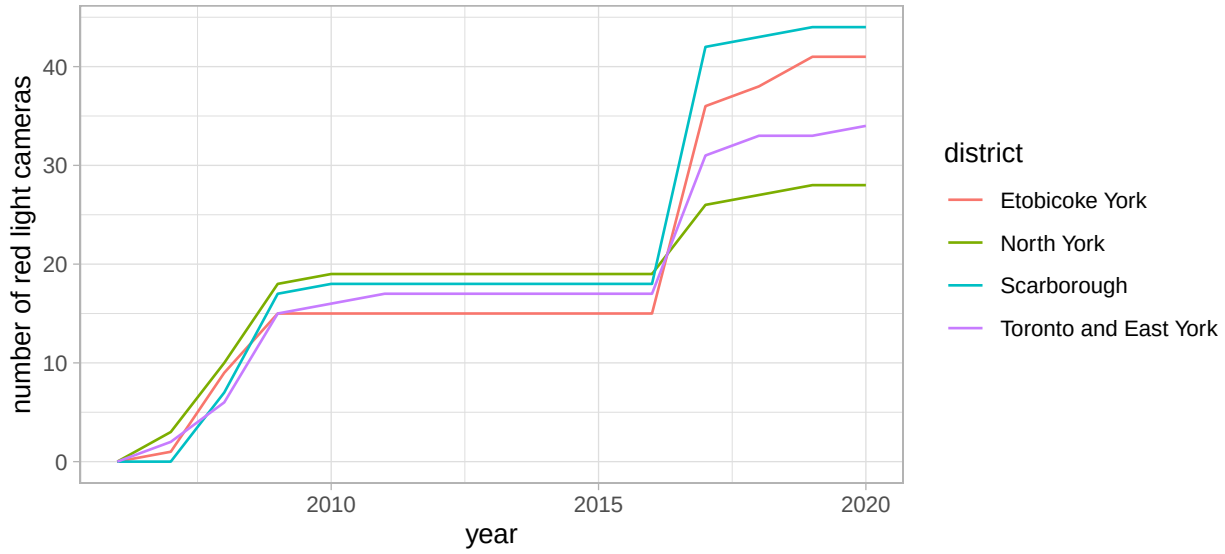

Figure 3: Number of red light cameras in the four districts each year

Figure 3 demonstrates the total number of red light cameras in each district from 2006 to 2020. Scarborough installed the most red light cameras in this period of time, especially after 2016, a large number of red light cameras were installed there. The district of North York, on the contrary, started with the most red light cameras installed among all districts, but has the slowest growth and ended up with the least number of red light cameras in 2020 among all districts. All the four districts installed a large number of red light cameras from 2016 to 2017.

I wanted to see how the increase in number of red light cameras compare with the traffic collision's frequency

in each district. So, I selected the accidents that occurred at intersections or somewhere near intersections from the Motor Vehicle Collision dataset and created a ggplot to see their trend by each district in Toronto.

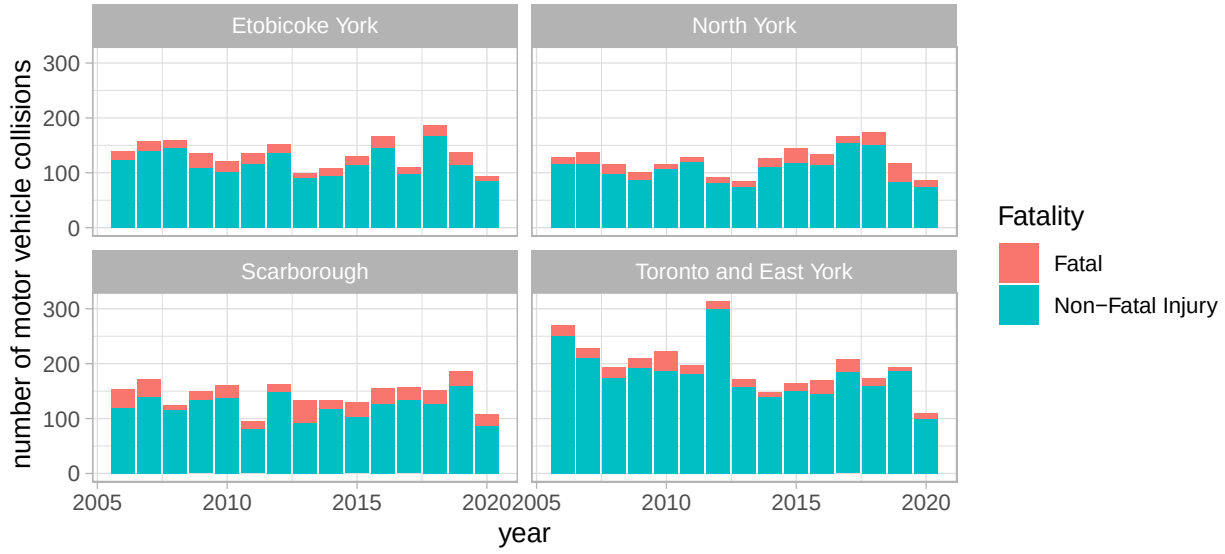

Figure 4: Comparing the number of motor vehicle collisions take place at or near intersections by district

Figure 4 shows the number of annual motor vehicle accidents occurred at or near intersections in each district. For all the districts, the numbers of fatal vehicle collisions stay approximately the same throughout these years, with Scarborough having the highest mortality rate in general. With regards to the total number of accidents, data in Etobicoke York, North York, and Scarborough all has its fluctuations, but the number remains about the same. The number of vehicle collisions in Toronto and East York was the highest in 2006, with more than to 260 cases, but it is the only district where the number has dropped.

By comparing Figure 3 and Figure 4, I found that they haven't show any associations between the number of traffic accidents and the number of red light cameras installed. For instance, Scarborough has installed the most number of red light cameras, especially after the year of 2016, but its number of traffic accidents didn't reduced at all, it even reach its peak in 2018, nor did the fatality of those accidents. We see a large decrease in Toronto and East York's motor vehicle collisions, but it didn't installed at much red cameras as Etobicoke York or Scarborough.

3 Discussion

The analysis of the two datasets show no relationship between the cases of motor vehicle collisions in Toronto and the number of red light cameras. From the comparision between Figure 3 and Figure 4, we see that in none of the districts did the number of vehicle collisions followed the trend of the installations of the red light cameras. For stance, between 2016 and 2017, all four districts made the installations of lots of new red light cameras, but none of the data in Figure 4 seemed to reflect that action. The implication of Section 2 is that: installing more red cameras does not change the number of traffic accidents in Toronto, nor does it influence the fatality of the accidents. A very likely reason to explain this situation is that red light cameras only concerned the drivers who already followed the rule most of the time, but not the rule-breaking drivers. Another possible reason would be that at intersections with red light cameras, more drivers would intend to stop rather than to proceed when they see the light is about to turn red, resulting in a higher chance of rear end accidents (Gallagher 2018).

In recent years, complaints have been raised about whether red light cameras are used mainly to prevent injuries and deaths from traffic accidents or earn revenues for the government (Gabor 2017). Measures should

be taken to improve the effectiveness of this system and to ensure the safety of drivers and pedestrians who are the victims of the vehicle accidents. For instance, drivers should be more notified on the existence of red light cameras, instead of knowing them after they are fined, because red light cameras is supposed to prevent accidents but not to punish people after they cause accidents. Government could also seek for other solutions to reduce these accidents. Toronto and East York has done a great work reducing the number of motor vehicle collisions, which means other district should try to adopt similar traffic controls.

References

Canada, Transport. 2019. *Canadian Motor Vehicle Traffic Collision Statistics: 2019*. https://tc.canada.ca/en/road-transportation/statistics-data/canadian-motor-vehicle-traffic-collision-statistics-2019.

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

Gabor, Robert. 2017. *Red Light Cameras: Are They Effective?* https://diamondlaw.ca/red-light-cameras-are-they-effective/.

Gallagher, Justin. 2018. *Red Light Cameras May Not Make Streets Safer*. https://www.scientificamerican.com/article/red-light-cameras-may-not-make-streets-safer/.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://sharlagelfand.github.io/opendatatoronto/,%0Ahttps://github.com/sharlagelfand/opendatatoronto/.

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/.

Services, Transportation. 2022. *Red Light Cameras*. https://open.toronto.ca/dataset/red-light-cameras/.

Webster, Bell R. 2022. *Why Does Toronto Have so Much Traffic?* https://newcanadianlife.com/why-does-toronto-have-so-much-traffic/.

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.

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/,%0Ahttps://github.com/haozhu233/kableExtra.