A Neighborhood and Temporal Analysis of Toronto's Car Accidents*

My subtitle if needed

Cristina Burca

September 27, 2024

Abstract

This study examines traffic accidents in Toronto from 2006 to 2023, focusing on neighborhood and temporal patterns to identify high-risk areas and peak accident times. The analysis of accident data by hour, day, and month revealed a significant spike in accidents at 3:00 PM during rush hour, particularly in neighborhoods like West Humber-Clairville and Yonge-Bay Corridor. Accidents also peaked in the summer months, from June to September. These findings underscore the need for improved traffic management during peak hours and targeted interventions in high-traffic neighborhoods.

Table of contents

*Code and data are available at: https://github.com/crisburca/Toronto_Accident_Analysis.

1 Introduction

Toronto is a city infamous for its heavy traffic, recently ranked as the city with the third worst traffic in the world (cite). The traffic is due to several factors. Toronto has experienced one of the fastest-growing populations in North America, with an increase of approximately a million residents in the past 15 years. The city has not kept pace in growth in terms of road infrastructure or public transport. Numerous planned road expansion and maintenance projects have been delayed, resulting in prolonged construction periods and bottlenecks across major routes. Furthermore, with years of under-investment in public transit and only two major transit lines leading into the center of the city, public transport has become overcrowded and unreliable. It is difficult for commuters to choose public transport over driving, and thus a heavy car dependency has been established. With roads heavily congested during peak hours, accidents have become a frequent occurrence. On average, there are more than 4 documented collisions each day. These accidents not only cause significant disruptions to daily commutes but become a safety concern for commuters.

As accidents have become a common occurrence, it is crucial to analyze their underlying causes in order to implement solutions. Understanding where and when these accidents happen can provide insight into how to prevent them. This paper aims to identify accident hotspots across Toronto's neighborhoods, as well as examine the frequency of accidents during specific hours of the day. By analyzing these patterns, we hope to give insight to policy makers on the infrastructure, traffic flow, or road design issues that might be contributing to higher accident rates. Section 2 includes an introduction to the dataset and important variables, Section 4 analyzes the data by neighbourhoods and intersections, Section 5 analyzes the data by hours, days of the week, and months, Section 6 discusses the results, limitations, and next steps, and Section 8 includes a simple map of the city of Toronto, for reference.

This study reveals that traffic accidents in Toronto spike during rush hours, particularly around 3:00 PM, with neighborhoods like West Humber-Clairville (119) and Yonge-Bay Corridor (170), St Lawrence-East/ Bayfront The Islands (166) and South Riverdale (70) and Wexford/Maryvale (119) witness the highest accident rates. The analysis also shows a seasonal trend, with more accidents occurring between June and September, likely due to increased travel and pedestrian activity. Despite Toronto's harsh winters, accident counts were lower during these months, possibly due to reduced road activity. These patterns highlight the critical role of urban planning and traffic management, particularly in high-traffic neighborhoods and during peak hours.

2 Data

2.1 Data source and referencing

The data is sourced from the City of Toronto Open Data (Gelfand 2022), a portal containing Licensed official data of Toronto. R(R Core Team 2023) was used to compile this paper as well as packages Tidyverse (Wickham et al. 2019), Dplyr (Wickham et al. 2023), and Lubridate (Grolemund and Wickham 2011) were used for the cleaning, analysis, and graphing of the data, Knitr (Xie 2014) was used for compiling tables, and Sf (Pebesma and Bivand 2023) and (Kahle and Wickham 2013) were used for the location based mapping plots.

2.2 Introduction to the Data

This dataset consists of 18,763 observations of automobile accidents in the Region of Toronto, recorded from January 1st, 2006, to December 29th, 2023. It includes six key variables of interest:

- `Date`: the date of the accident,
- `Time`: the time of the accident,
- `Street 1` and `Street 2`: the nearest intersection where the accident occurred,
- `Hood`: the neighborhood ID (Toronto is divided into 158 neighborhoods),
- `geometry`: containing the latitude and longitude coordinates of the accident location.

The data was cleaned to ensure that the term "automobile accidents" includes any incidents involving cars, trucks, motorcycles, transit vehicles, or emergency vehicles. Accidents involving pedestrians or cyclists are also included. Local or side roads are excluded form this data, and intersections with less than 5 accidents are omitted.

2.3 Measurement

The dataset records motor vehicle collisions in Toronto, capturing the date, time, and location (latitude and longitude) of each incident. The variables used include accident location (nearest intersections), and neighborhood ID (HOOD_158). For the analysis, we focused on major roads and highways to better capture high-traffic areas. The spatial data allows us to identify accident hotspots, while the temporal data helps analyze peak accident times.

While the dataset offers an overview of accidents in the Greater Toronto Area, it is not entirely complete. Some accidents, such as minor collisions or hit-and-run cases, may go unreported and are thus missing from the dataset. Additionally, focusing on areas with high accident frequency means that the analysis may overlook less accident-prone neighborhoods, limiting the context for baseline comparisons.

3 Data Overview

The map in Figure 1 illustrates high-density accident intersections across Toronto. The darker and larger circles represent intersections with a higher number of accidents. These accident hotspots are concentrated along major roadways and intersections, as expected. The densest regions appear in Downtown Toronto, particularly near major highways, such as the intersections along Highway 401, 427, and 409. Additionally, several clusters appear in densely populated neighborhoods, such as North York and Scarborough.

Figure 2 displays the top five neighborhoods with the highest number of accidents. Neighborhood 1 (West Humber-Clairville) is located at the west of the map, neighborhoods Yonge-Bay Corridor (170), St Lawrence-East/ Bayfront The Islands (166) and South Riverdale (70) at the south of the map and Wexford/Maryvale (119) at the east of the map. The number of accidents in each of these neighborhoods points to potential urban planning challenges, such as high traffic density, poorly designed intersections, or insufficient traffic control measures. For example, the West Humber-Clairville area (1), with its proximity to major highway interchanges, experiences a high volume of vehicular movement, which may contribute to its high accident rates. Similarly, the Yonge-Bay Corridor (170), may experience heavy traffic due to commuter flows, particularly during peak hours.

4 Neighbourhood Analysis

The map Figure 3 provides an overview of the top 5 neighborhoods in Toronto with the highest number of accidents from 2006 to 2023. Each point represents a specific location where an accident occurred, with the size and color of the point reflecting the accident count in that area.

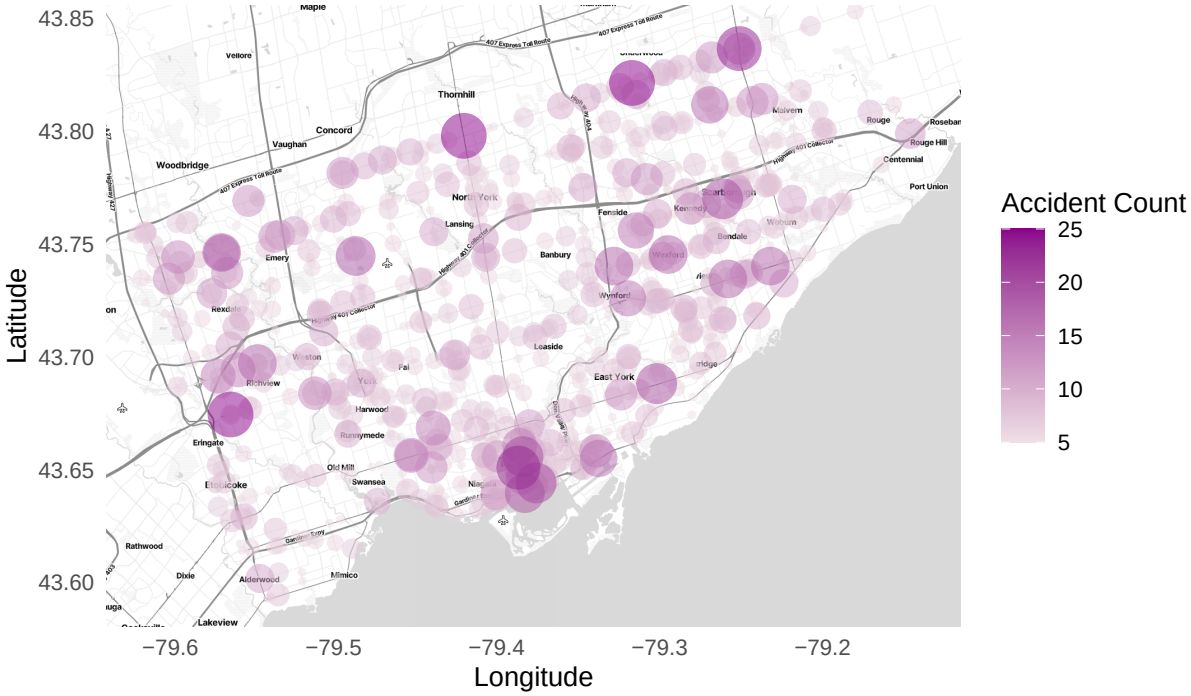


Figure 1: Map of high-density motor vehicle accident points in Toronto from 2006 to to 2023

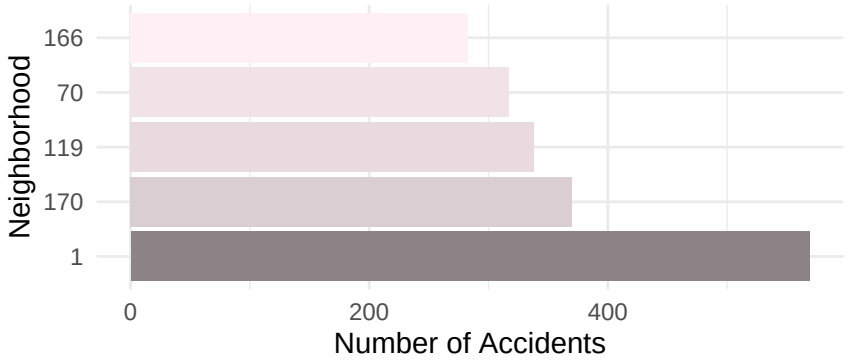


Figure 2: Top 5 neighborhoods in Toronto with the highest number of car accidents

Figure 3: Map of top 5 neighbourhoods in Toronto with highest accident count

4.1 South Toronto

In Figure 4, we see that accidents are heavily concentrated in the Downtown core, particularly along major streets such as Lake Shore Boulevard West, Queen Street, and University Avenue. These areas are likely to experience high traffic volumes due to their proximity to key business, entertainment, tourist districts as well as Union Station and the Gardiner Expressway. Downtown accidents tend to cluster around dense urban blocks.

4.2 West Toronto

Figure 5 provides a closer look at the neighborhoods around West Humber-Clairville. This neighbourhood is located on the West edge of Toronto, consisting of intersections of major highways in Toronto, including the 427, 407, 409, and 401. The accident hotspots are primarily found along major arterial roads and intersections, such as those around the 427 and Highway 401 interchange. These intersections are known for their high traffic volumes, especially during rush hours, and are major connectors for drivers entering and leaving the city.

4.3 East Toronto

In Figure 6, the accidents are concentrated along major streets, including Ellesmere Road, Warden Avenue, and Lawrence Avenue. These streets connect various residential neighborhoods to commercial areas and are likely to experience heavy traffic during commuting hours. Similar to West Toronto, the larger points appear at major intersections, where traffic volumes are highest.

Figure 4: Map of accident hotspots in neighborhoods 170 (Yonge-Bay Corridor), 166 (St. Lawrence-The Islands), and 70 (South Riverdale), among the top five with the highest accident counts in Toronto

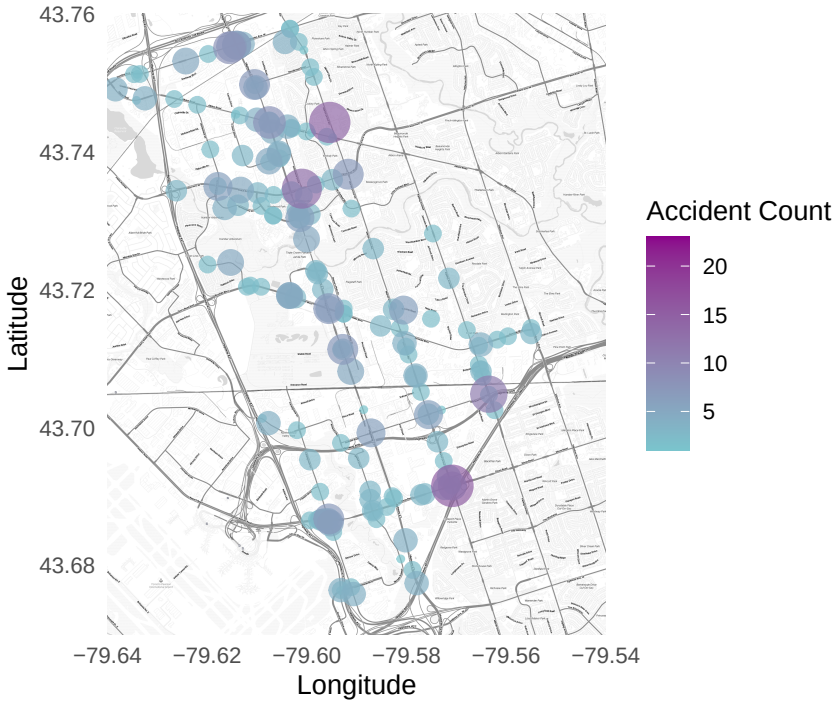


Figure 5: Map of accident hotspots in neighborhood 1 (West Humber-Clairville) in Toronto

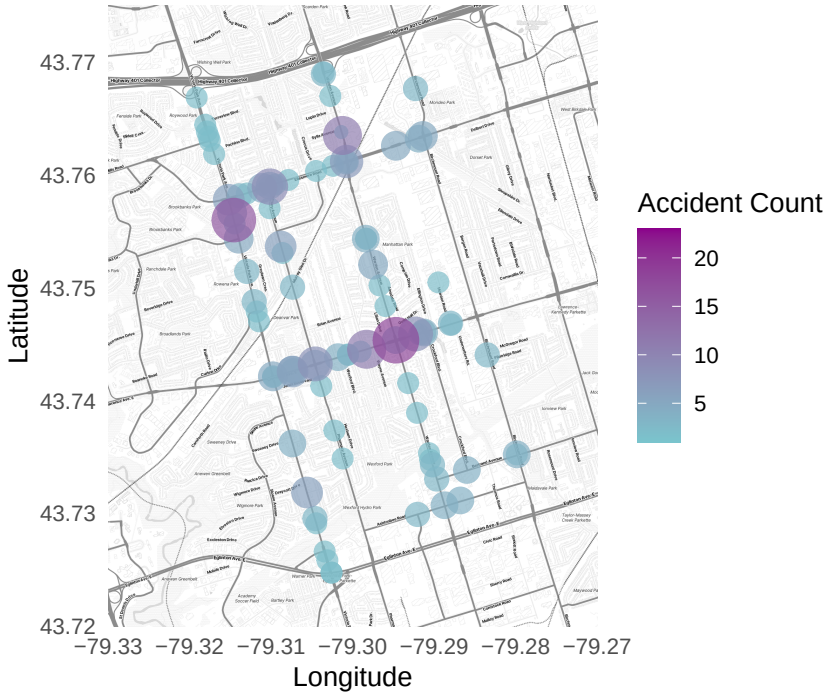


Figure 6: Map of accident hotspots in neighborhood 119 (Wexford/Maryvale) in Toronto

5 Month, Week and Day Analysis

5.1 Hours

The first graph, Figure 7, there is a rise in accidents from 6:00 AM onwards, peaking between 3:00 PM and 6:00 PM. This time period coincides with the evening rush hour, which is when traffic density is at its highest as people commute home from work or school.

Notably, there's also a significant number of accidents occurring late at night, particularly around 11:00 AM-12:00 AM. This could be attributed to factors such as impaired driving, fatigue, or lower visibility at night. The rise throughout the afternoon suggests that as the day progresses and more vehicles are on the road, the likelihood of accidents increases.

5.2 Week

Figure 8 we observe that accidents are relatively consistent from Monday through Thursday, but there is a notable increase on Fridays. The spike in accidents on Fridays could be related to the increased volume of vehicles as people begin their weekend.

Interestingly, accident numbers slightly dip on Saturdays and Sundays compared to weekdays, likely due to lower traffic volumes as fewer people commute to work. However, despite the lower traffic on weekends, the number of accidents remains substantial, suggesting that other factors—such as recreational travel, shopping, and leisure activities—contribute to the accident count.

5.3 Months

Looking at the Figure 9:

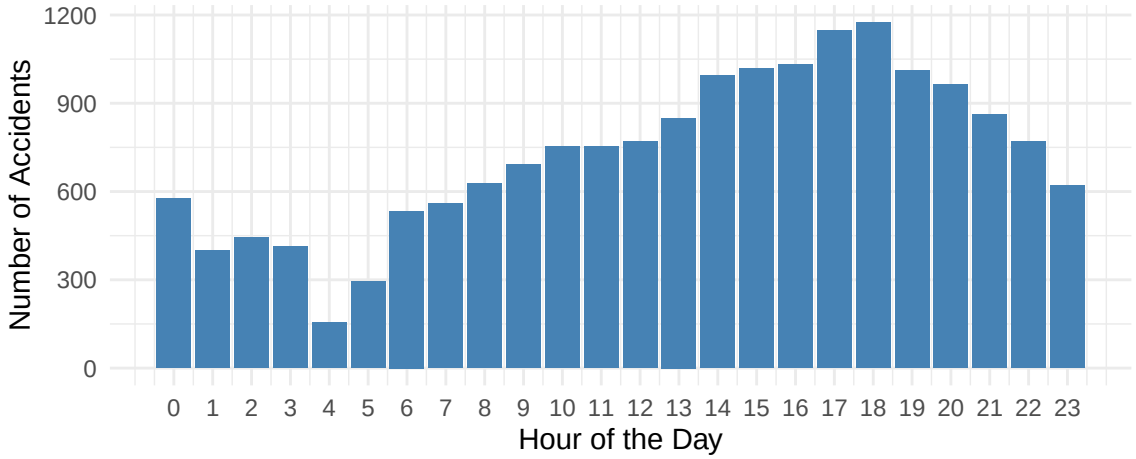


Figure 7: Bar graph of accident count by hour in Toronto

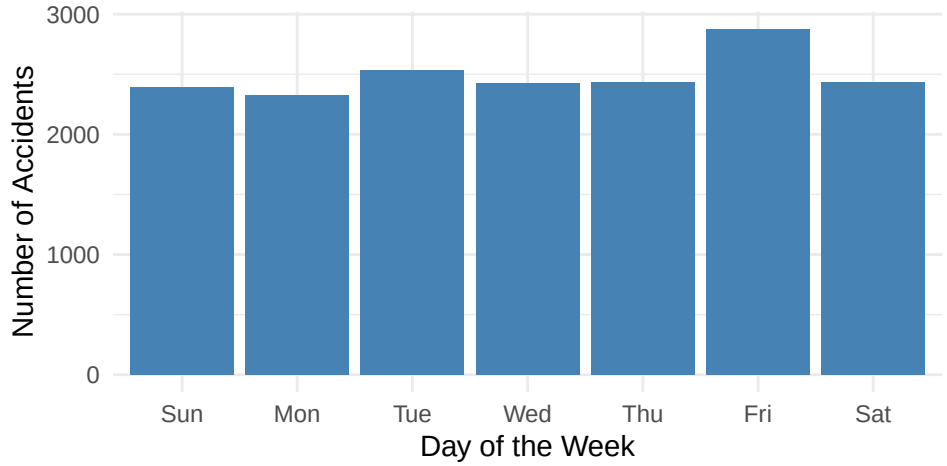


Figure 8: Bar graph of accident count by day of week in Toronto

The data shows a significant increase in accidents during the summer months, with June, July, August, and September seeing the highest number of accidents. Interestingly, while one might expect the winter months (January, February, December) to have the highest accident rates due to icy and snowy conditions, these months have relatively fewer accidents compared to the summer peak.

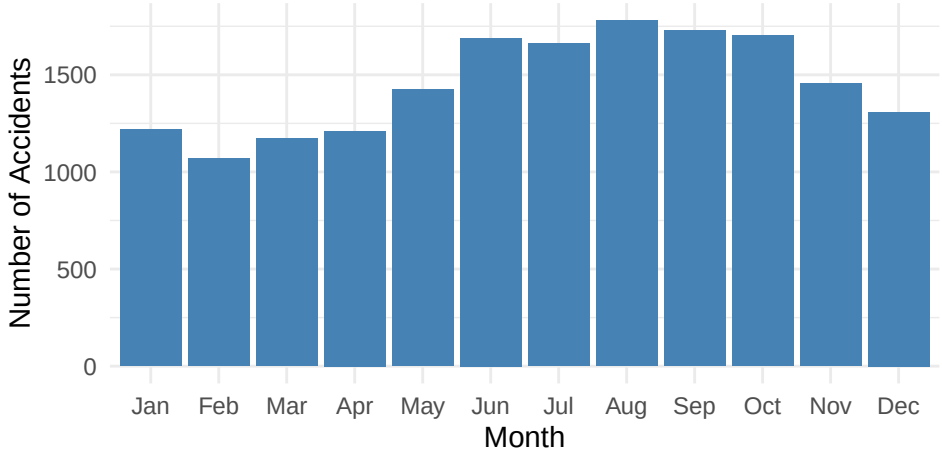


Figure 9: Bar graph of accident count by month in Toronto

Figure 10 illustrates the number of accidents occurring across different neighborhoods during rush hour (3:00 PM - 6:00 PM), broken down by each hour (15, 16, 17, 18).

Overall, 3:00 PM seems to be the most accident-prone hour in most of these neighborhoods, possibly signaling the start of peak traffic congestion.

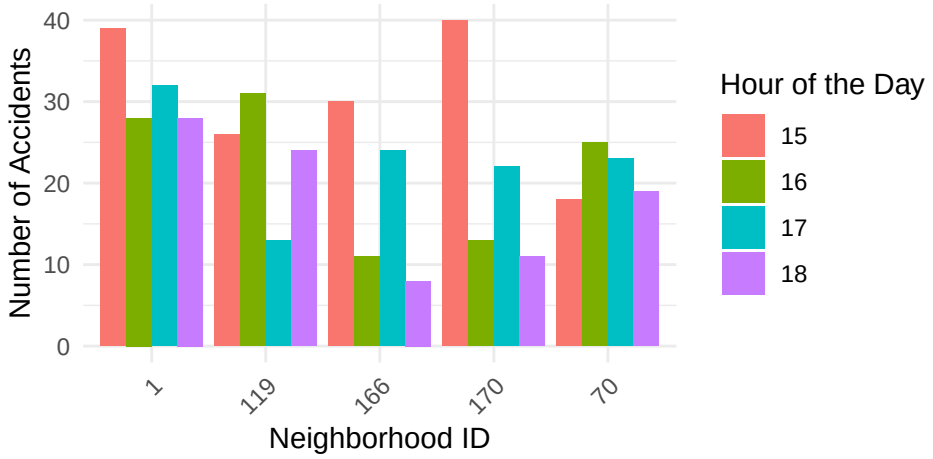


Figure 10: Bar graph of accident count by top 5 neighbourhoods by peak hour

Figure 11 summarizes the average number of accidents for each hour of the rush period (15:00 to 18:00). Key insights include:

3:00 PM (15) has the highest average number of accidents. As rush hour progresses, there is a steady decline in the number of accidents. Both figures highlight how 3:00 PM (15) marks the most dangerous time during rush hour in many neighborhoods. Areas like West Humber-Clairville and the Yonge-Bay Corridor experience a spike in accidents early in the afternoon, likely due to the onset of heavy traffic as people leave work or school.

These insights suggest that targeted traffic management could be most effective if implemented earlier in the afternoon, around 3:00 PM. For example, optimizing traffic signals or increasing law enforcement presence during this peak hour might help reduce accident rates in these high-risk areas.

Hour of Day	Average Accidents
15	113
16	95
17	92
18	79

Figure 11: Average number of accidents in Toronto by peak hour

6 Discussion

The results of this study provide a view of traffic accidents in Toronto, focusing on the spatial and temporal dimensions of the problem. One finding is the spike in accidents during rush hours, specifically at 3:00 PM. This pattern is seen across several neighborhoods, with West Humber-Clairville (119) and the Yonge-Bay Corridor (170) experiencing the most accidents at this time. The high accident rate at the beginning of the rush hour suggests that the sudden increase of vehicles on the road contributes to higher traffic congestion and subsequently more accidents.

Another important observation is the seasonal variation in accident counts, with June to September emerging as the months with the highest number of accidents. This can likely be attributed to the increase in overall traffic during the summer months from increased travel, cyclists and pedestrians, and recreational driving. Interestingly, despite the harsh winter conditions in Toronto, the winter months did not show as many accidents as expected. This may be due to reduced road activity during extreme weather conditions or more cautious driving.

The neighborhood-level analysis reveals important spatial patterns, with certain intersections and streets being particularly accident-prone. Downtown areas, as well as neighborhoods close to major highways, consistently show higher accident counts. This suggests that urban planning and traffic infrastructure play a critical role in accident frequency. From a temporal perspective, the decline in accidents after 3:00 PM, as seen in the hourly accident data, indicates that while rush hour remains a risk factor, the period of highest risk is confined to the start of the rush, possibly due to impatience, aggressive driving, or distractions during the initial wave of commuters.

7 Conclusion

The results of this study provide a comprehensive view of traffic accidents in Toronto, particularly focusing on the spatial and temporal dimensions of the problem. One key finding is the clear spike in accidents during rush hours, specifically at 3:00 PM. This pattern is seen across several neighborhoods, with West Humber-Clairville and the Yonge-Bay Corridor experiencing the most accidents at this time. The high accident rate at the beginning of the rush hour suggests that the sudden influx of vehicles on the road contributes to higher traffic congestion and subsequently more accidents.

Another important observation is the seasonal variation in accident counts, with June to September emerging as the months with the highest number of accidents. This can likely be attributed to the increase in overall traffic during the summer months, as well as the presence of more cyclists and pedestrians. Interestingly, despite the harsh winter conditions in Toronto, the winter months did not show as many accidents as one might expect. This may be due to reduced road activity during extreme weather conditions or more cautious

driving. From a temporal perspective, the decline in accidents after 3:00 PM, as seen in the hourly accident data, indicates that while rush hour remains a risk factor, the period of highest risk is confined to the start of the rush at 3:00 PM.

The neighborhood-level analysis reveals important spatial patterns, with certain intersections and streets being particularly accident-prone. Downtown areas, as well as neighborhoods close to major highways, consistently show higher accident counts.

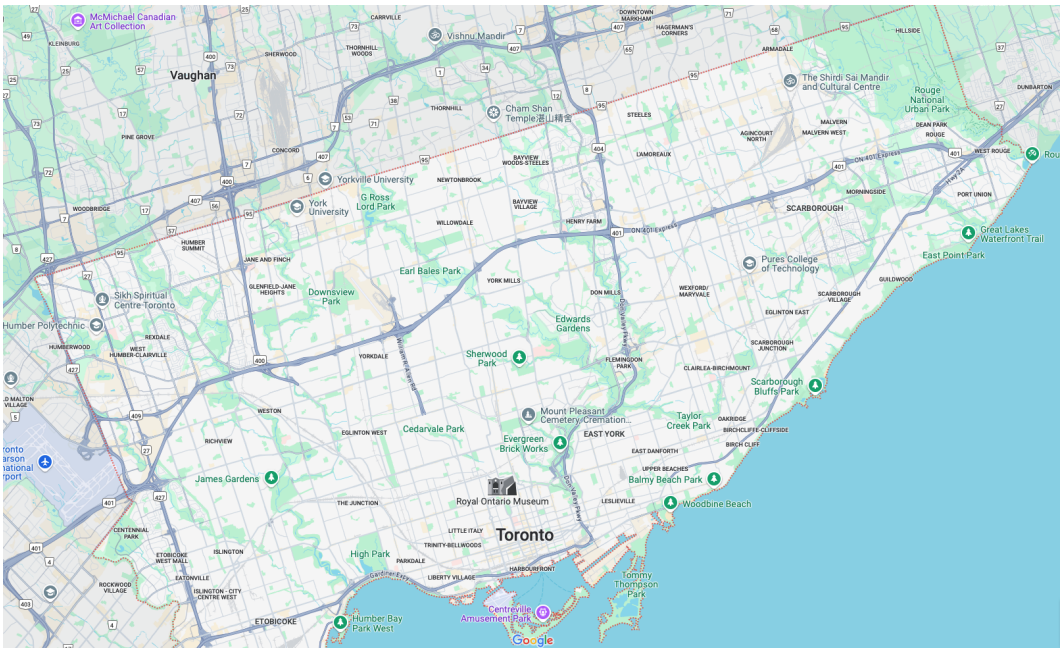
7.1 Limitations

There are a few limitations to this study. It is not mentioned how the accidents happen, whether it is due to human fault, poor infrastructure, or misleading signs. Cases that are not due to human fault, car failure, should not be observed, but misleadings signs, poor road and surface infrastructure, weather conditions resulting in dangerous driving and other reasons should be considered. Data should be studied over the years, to recognize increasing hotspot intersections, and other problems in road construction. Important institutions and buildings that are highly populated should be noted and their areas around them.

7.2 Next steps

Future studies could incorporate weather data and traffic flow information. It would also be useful to explore intervention measures, such as changes in road infrastructure or traffic management, to assess their impact on accident reduction. Furthermore, looking into the role of speed limits, signal timing, and pedestrian crossings could offer additional insights into how urban planning can mitigate accident risks. Finally, expanding the analysis to include detailed information on driver behavior, such as speeding or distracted driving, could further enhance the understanding of accident causes and prevention strategies.

8 Appendix



References

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

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

Kahle, David, and Hadley Wickham. 2013. "Ggmap: Spatial Visualization with Ggplot2." *The R Journal* 5 (1): 144–61. https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf.

Pebesma, Edzer, and Roger Bivand. 2023. *Spatial Data Science: With applications in R*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.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. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.