Traffic Collisions in Toronto: Analysis of Key Variables*

An Analysis of Traffic Collision Data from 2014-2023

Andy Jiang

September 24, 2024

This paper analyzes traffic collision data in Toronto from 2014 to 2023. It explores trends in collisions by month, year, and neighborhood, examining variables such as fatalities, injuries, and property damage. The analysis is done using R, with visualizations produced through ggplot2.

1 Introduction

This paper analyzes traffic collision data in Toronto from 2014 to 2023. We explore trends in monthly, yearly, and neighborhood-level collisions, with a focus on variables such as fatalities, injuries, property damage, and modes of transport involved. The data is sourced from Open Data Toronto and contains detailed information about each collision.

The remainder of this paper is structured as follows. Section Section 2 provides details of the dataset used in this analysis. Section Section 3 presents the results of the analysis, followed by a discussion in Section Section 4. Finally, Section Section 5 summarizes the key findings of the study.

2 Data

The dataset used in this analysis was obtained from Open Data Toronto. It includes traffic collision data collected between 2014 and 2023. The variables include the date and time of the collision, whether there were fatalities, injuries, or property damage, and the type of vehicle(s) involved. Table 1 provides a sample of the cleaned dataset.

*Code and data are available at: [LINK]

Table 1: Sample of the cleaned collision data

Month	Day of Week	Year	Hour	Fatalities	Injury Collisions	FTR Collisions	PD Collisions	Neighbourhood	Automobile	Motorcycle	Passenger	Bicycle	Pedestrian
January	Wednesday	2014	4	NA	YES	NO	NO	Bendale South (157)	YES	NO	YES	NO	NO
January	Wednesday	2014	14	NA	NO	YES	NO	Kensington-Chinatown (78)	YES	NO	NO	NO	NO
January	Wednesday	2014	2	NA	YES	NO	NO	NSA	YES	NO	NO	NO	NO
January	Wednesday	2014	3	NA	NO	NO	YES	NSA	YES	NO	NO	NO	NO
January	Wednesday	2014	5	NA	YES	NO	NO	NSA	YES	NO	NO	NO	NO
January	Wednesday	2014	5	NA	NO	NO	YES	NSA	YES	NO	NO	NO	NO
January	Wednesday	2014	8	NA	NO	NO	YES	NSA	YES	NO	NO	NO	NO
January	Wednesday	2014	8	NA	NO	NO	YES	NSA	YES	NO	NO	NO	NO
January	Wednesday	2014	9	NA	NO	NO	YES	NSA	YES	NO	NO	NO	NO
January	Wednesday	2014	9	NA	YES	NO	NO	NSA	YES	NO	YES	NO	NO

3 Results

Here we analyze the data by looking at the frequency of traffic collisions based on various factors such as month, year, and neighborhood.

Collisions by Month The following figure shows the distribution of collisions by month across all years.

Collisions by Hour of the Day

```
Warning: Removed 8039 rows containing non-finite outside the scale range
(`stat_bin()`).
```

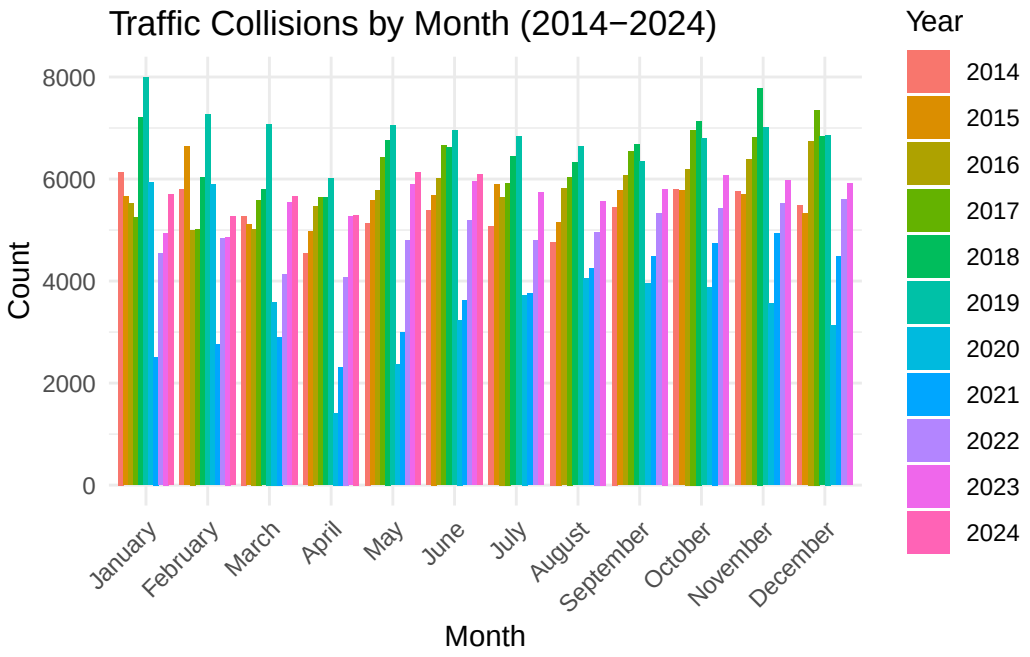

Figure 1: Collisions by Month

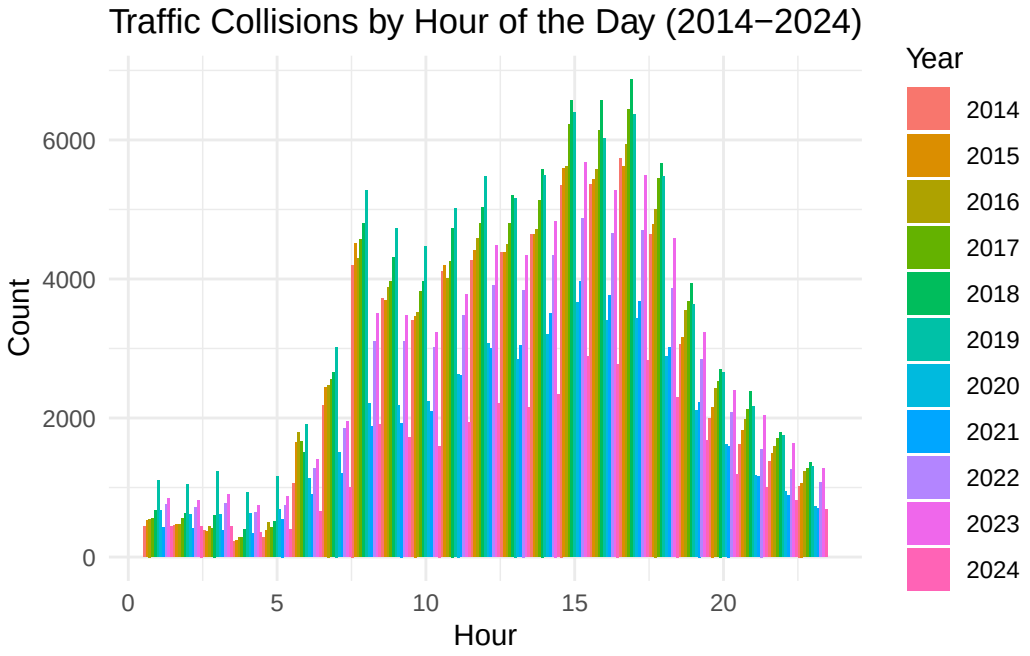

Figure 2: Collisions by Hour of the Day

4 Discussion

5 Conclusion

6 References