**Course Name:** Data Analytics Tools 1

**Course Code:** BIA-5201-0LA

—

**Course Instructor:** Na Sui
Longo Faculty of Business / Humber College

The purpose of this project is to apply data analysis concepts and tools to explore public safety data provided by the Toronto Police Service, with a specific focus on identifying patterns in traffic-related incidents and generating insights and recommendations to improve road safety in alignment with the City of Toronto's Vision Zero Plan.

## Group Members:

| Student Name | Student ID |
|---|---|
| Thanabodee Wongwichai | N01725320 |
| Tracy Abigail George | N01678319 |
| Valerie Obasi | N01656280 |
| Darren Fernandes | N01712659 |
| Mohamed Waseem Khan Ayub Khan | N01725737 |
| Mohammed Faisal Khan | N01740854 |

# Abstract

Toronto, the most populous city in Canada, is facing increasing challenges in maintaining public safety due to rapid urbanization and rising traffic volumes. In line with its commitment to transparency and making informed decisions based on data, the Toronto Police Service has introduced the Public Safety Data Portal to facilitate open data access and generate valuable insights. This initiative, led by a newly appointed data scientist in the Analytics & Innovation unit, utilizes the portal to examine incidents related to traffic, with a specific focus on road safety.

This study examines high-risk areas for severe traffic collisions through exploratory data analysis and geospatial modelling, particularly around sensitive locations such as schools and hospitals. The analysis reveals temporal trends, weather factors, and spatial patterns linked to increased incident rates. A significant finding indicates a surge in severe collisions during specific time periods near school zones, highlighting the urgent need for focused interventions.

In alignment with the City of Toronto's Vision Zero Plan, this study provides actionable recommendations, including enhanced traffic calming measures in sensitive areas, dynamic signal timing adjustments during peak traffic hours, and the use of predictive analytics tools to identify high-risk locations proactively. The project concludes with recommendations for future data collection, such as gathering pedestrian and cyclist density data to enhance our understanding of road safety risks. This work highlights the significance of data science in developing effective public safety strategies that can adapt to Toronto's evolving urban landscape.

# Introduction

As urbanization progresses, the City of Toronto encounters mounting challenges in safeguarding the safety and well-being of its residents, especially on increasingly congested roads. With a population exceeding 2.9 million, Toronto stands as Canada's largest city and one of the most densely populated areas in North America. Addressing growing concerns about traffic-related injuries and fatalities, the Toronto Police Service (TPS) has introduced the Public Safety Data Portal to enhance transparency, foster public trust, and inform data-driven policy-making decisions.

This project aims to utilize data from the Public Safety Data Portal to investigate traffic collision patterns throughout the city of Toronto. It focuses on key variables, including time of day, location, weather conditions, and proximity to sensitive areas such as schools and hospitals, to reveal actionable insights for enhancing road safety strategies. This analysis aligns with the City's Vision Zero Plan, an initiative to eradicate traffic fatalities and severe injuries. Ultimately, the goal is to deliver recommendations based on evidence that support proactive safety efforts and promote safer streets for all Toronto residents.

This project employs exploratory data analysis and visualization to address a specific issue in road safety, offering a data-driven approach for TPS leadership and policymakers to inform their future decisions.

# Problem Definition

Toronto's rapid expansion and urban concentration have led to a surge in road traffic, which increases the risk of collisions among vehicles, pedestrians, and cyclists. Although the City and law enforcement continue to strive for safer streets, traffic-related injuries and fatalities persist as significant public safety issues. Initiatives such as the Vision Zero Plan aim to address these challenges, but successful execution requires a thorough understanding of where, when, and why these incidents occur.

This project addresses the critical issue of identifying patterns and risk factors linked to severe traffic incidents in Toronto. While existing high-level reports provide a general overview of trends, they often overlook detailed insights that are crucial for guiding targeted interventions. This analysis aims to bridge that gap by leveraging open data from the Toronto Police Service, uncovering significant relationships between collision frequency and contextual variables, including time, location, environmental conditions, and proximity to vulnerable areas such as schools and hospitals.

This project aims to produce actionable recommendations that extend beyond general policy suggestions by focusing on a specific question or theme, such as the impact of traffic incidents near schools. Its goal is to assist the TPS in shifting from reactive measures to proactive prevention strategies, ultimately enhancing urban safety.

# Data Cleaning & Preparation

To ready the dataset for valuable analysis and guarantee the reliability of insights, various data cleaning and transformation procedures were implemented:

1. **Encoding Binary Values:** Columns with "Yes" and "No" answers were converted into numerical values for analysis. "Yes" was represented as 1 and "No" as 0, simplifying statistical evaluation operations.
2. **Season Classification:** A new column has been added to categorize each incident by season—Winter, Spring, Summer, or Fall, according to the date. This allows for the detection of seasonal trends in traffic incidents.
3. **Standardization:** Addresses that were missing or incomplete were marked as "Unknown" to maintain the integrity of the records while indicating the lack of location information. This method kept the dataset's integrity without discarding potentially valid entries.
4. **Time of Day Categorization:** The hour of occurrence was converted into a numerical format to generate a new Time of Day column. This column categorizes incidents into four segments: Early Morning, Morning, Evening, and Night. This categorization provided a clearer insight into the high-risk periods throughout the day.
5. **Date Formatting:** The date fields have been standardized to a consistent format, enabling efficient sorting, filtering, and time-series analysis.

These data preparation steps guaranteed that the dataset was clean, consistent, and prepared for thorough exploration, which ultimately facilitated a more precise and targeted analysis of road safety trends in Toronto.

# Time-of-Day Pattern Analysis

We conducted a One-Way ANOVA (Analysis of Variance) test to investigate if traffic collision times significantly vary across the days of the week. The objective was to ascertain whether the average occurrence times of collisions meaningfully differ from Monday to Sunday.

Each incident in the dataset was recorded with an occurrence hour, allowing us to calculate the average accident time for each day. Summary statistics indicated that, although the mean times for all days were fairly similar, ranging from approximately 13.1 to 13.7 hours (1:00–2:00 PM), statistical validation was necessary to determine if these differences were statistically significant.

**Hypotheses:**

- **Null Hypothesis ($H_0$):** There is no difference in the average time of occurrence across the days of the week ($\mu_1 = \mu_2 = \mu_3 = \dots = \mu_7$).
- **Alternative Hypothesis ($H_1$):** At least one day has a significantly different mean accident time.

data preparation

| Monday | Tuesday | Wednesda | Thursday | Friday | Saturday | Sunday |
|---|---|---|---|---|---|---|
| 17 | 20 | 18 | 16 | 20 | 20 | 20 |
| 17 | 14 | 8 | 15 | 12 | 18 | 8 |
| 12 | 9 | 19 | 12 | 11 | 19 | 1 |
| 8 | 19 | 21 | 16 | 18 | 13 | 6 |
| 18 | 17 | 19 | 15 | 11 | 19 | 16 |
| | | | | | | 3 |
| | | | | | | 2 |
| | | | | | | 1 |
| | | | | | | 6 |
| 19 | 9 | 1 | 2 | 15 | 19 | 0 |
| 6 | 12 | 0 | 12 | 20 | 16 | 21 |
| 18 | 17 | 17 | 10 | 19 | 16 | 12 |
| 13 | 4 | 18 | 8 | 17 | 11 | 10 |
| 19 | 5 | 18 | 11 | 0 | 14 | 19 |
| 5 | 7 | 18 | 12 | 21 | 16 | 23 |
| 18 | 8 | 17 | 12 | 7 | 7 | 21 |

for data preparation for this ONE-WAY ANOVA test,
I filtered all occurrence times separated by each day of week

**H0:** μ1 =μ2 = μ3 = μ4 = μ5 (no differences between days of week)
**H1:** At least two means differ

Anova: Single Factor

SUMMARY

| *Groups* | *Count* | *Sum* | *Average* | *Variance* |
|---|---|---|---|---|
| Monday | 37368 | 489932 | 13.111 | 24.99279 |
| Tuesday | 43388 | 576348 | 13.28358 | 23.71643 |
| Wednesday | 43085 | 576366 | 13.37742 | 24.10769 |
| Thursday | 44884 | 604264 | 13.46279 | 24.13214 |
| Friday | 47889 | 657908 | 13.73819 | 25.46453 |
| Saturday | 38491 | 527614 | 13.70746 | 29.21382 |
| Sunday | 29991 | 402604 | 13.42416 | 31.59891 |

since there are 7 days in a week and each day has its own accident occurrence time data, we want to examine whether these time significantly differ from one another by conducting **ONE-WAY ANOVA** test

From **One-Way ANOVA test**,
we can see that on each day of week, **the most frquent time that accidents to occur is around 1-2 PM** which suggest that we should pay extra attention during this hour in order to prevent accidents or improve road safety.

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Between Groups | 12278.13 | 6 | 2046.356 | 79.09128 | 3.1E-99 | 2.098629 |
| Within Groups | 7376206 | 285089 | 25.87334 | | | |
| | | | | | | |
| Total | 7388484 | 285095 | | | | |

p-value < 0.05 reject H0,

therefore, there is a significant difference in the time of occurence across each day

As the p-value is well below 0.05, we reject the null hypothesis. This suggests that there is a statistically significant difference in the average occurrence timing of traffic collisions on different days of the week.

Even with these differences, the peak time for accidents throughout the week was approximately 1:00–2:00 PM, indicating this timeframe is crucial for road safety measures. Higher traffic volume due to lunch breaks or school dismissals may play a role. These results underscore the necessity for focused enforcement or awareness initiatives during this period to minimize collisions and improve public safety.

# Weather Impact Analysis

This analysis examines the impact of weather-related factors, identified through seasonal and time-of-day assessments, on the frequency and severity of traffic collisions in Toronto. By examining collision statistics, fatalities, and collision types across various seasons and times of day, we aim to identify patterns that can inform targeted safety initiatives.

**1. Overall Collision Counts by Season**

Data shows significant seasonal variations in traffic incidents. Winter experienced the highest number of collisions (74,725), followed by summer (51,529), spring (46,466), and fall (37,695). This pattern holds true for all types of collisions, including fatal, injury, property damage, and hit-and-run incidents. The higher collision rates in winter likely result from hazardous driving conditions, such as snow and ice, which increase the risk of accidents.

**2. Collision Counts by Season and Time of Day**

Analyzing fatality data by season and time of day reveals that most fatal collisions take place during the afternoon and evening across all seasons. In winter, fatalities peak during these times, with 23 deaths in the afternoon and 26 in the evening. Similarly, summer exhibits this trend, recording 21 fatalities in the afternoon and 29 in the evening. Although night and early morning periods typically report fewer fatalities, the numbers are still considerable.

Property damage collisions reach their highest numbers in winter afternoons (32,970) and summer afternoons (24,660), indicating a rise in traffic and possibly difficult road conditions during these times.

**3. Collision Counts by Season and Day of the Week**

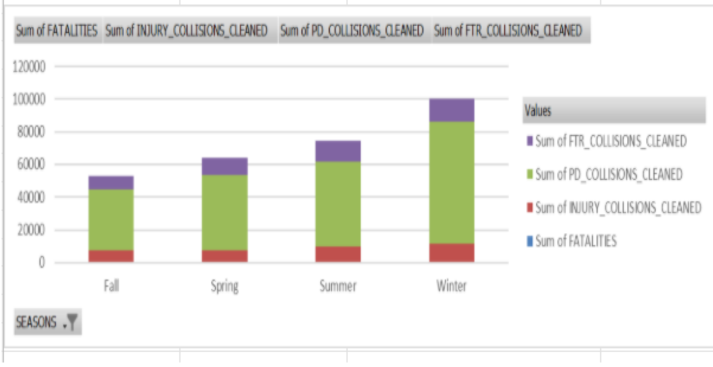
Fatal collisions fluctuate by season and day of the week. Winter consistently shows the highest overall fatalities, but these deaths occur throughout the week. Notably, Fridays and Saturdays typically report higher fatality numbers, possibly linked to greater travel or social events. This seasonal and weekly distribution offers valuable insights for law enforcement resource allocation planning.

**4. Involved Parties by Season**

Analyzing the parties involved in collisions reveals seasonal variations in the participation of automobiles, motorcycles, bicycles, and pedestrians. Collisions involving automobiles are prevalent throughout all seasons, although there is a minor decline in winter, which may be attributed to the decreased activity of motorcycles and bicycles. In contrast, the involvement of bicycles and pedestrians rises during spring and summer, aligning with the seasonal shifts in travel habits and outdoor activities.

## 1. Overall Collision Counts by Season

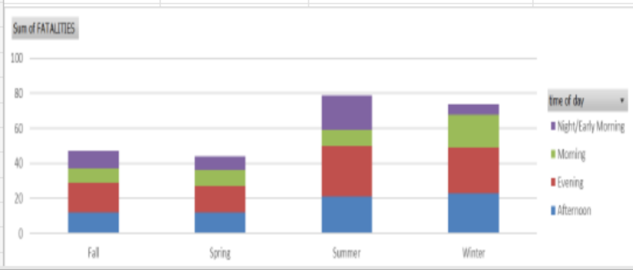
| Row Labels | Sum of FATALITIES | Sum of INJURY_COLLISIONS_ | Sum of PD_COLLISIONS_ | Sum of FTR_COLLISIONS_CLEANED |
|---|---|---|---|---|
| Fall | 47 | 7330 | 37695 | 7779 |
| Spring | 44 | 7257 | 46466 | 10370 |
| Summer | 79 | 10013 | 51529 | 12634 |
| Winter | 74 | 11308 | 74725 | 14275 |
| **Grand Total** | **244** | **35908** | **210415** | **45058** |




## 2. Collision Counts by Season and Time of Day

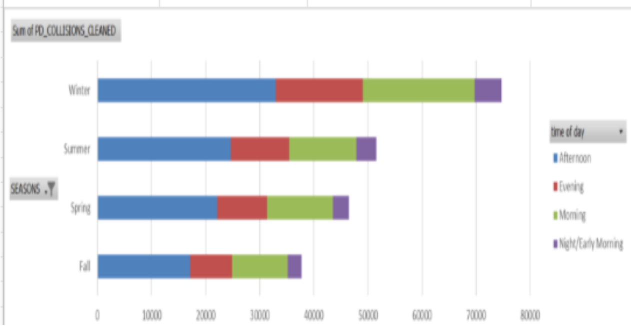
### A)Fatalities by Season and Time of Day

| Sum of FATALITIES | Column Labels | | | | |
|---|---|---|---|---|---|
| **Row Labels** | **Afternoon** | **Evening** | **Morning** | **Night/Early Morning** | **Grand Total** |
| Fall | 12 | 17 | 8 | 10 | 47 |
| Spring | 12 | 15 | 9 | 8 | 44 |
| Summer | 21 | 29 | 9 | 20 | 79 |
| Winter | 23 | 26 | 19 | 6 | 74 |
| **Grand Total** | **68** | **87** | **45** | **44** | **244** |




### D) Property Damage Collisions by Season and Time of Day

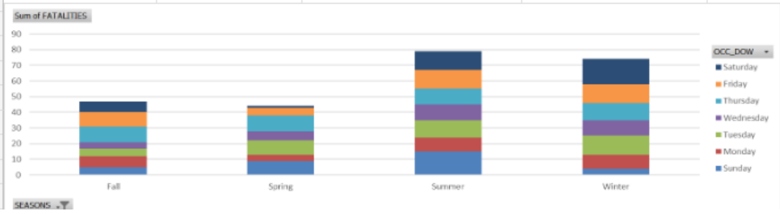
| Sum of PD_COLLISI( | Column Labels | | | | |
|---|---|---|---|---|---|
| **Row Labels** | **Afternoon** | **Evening** | **Morning** | **Night/Early Morning** | **Grand Total** |
| Fall | 17195 | 7715 | 10238 | 2547 | 37695 |
| Spring | 22141 | 9254 | 12143 | 2928 | 46466 |
| Summer | 24660 | 10768 | 12458 | 3643 | 51529 |
| Winter | 32970 | 16116 | 20627 | 5012 | 74725 |
| **Grand Total** | **96966** | **43853** | **55466** | **14130** | **210415** |



**3. Collision Counts by Season and Day of Week**

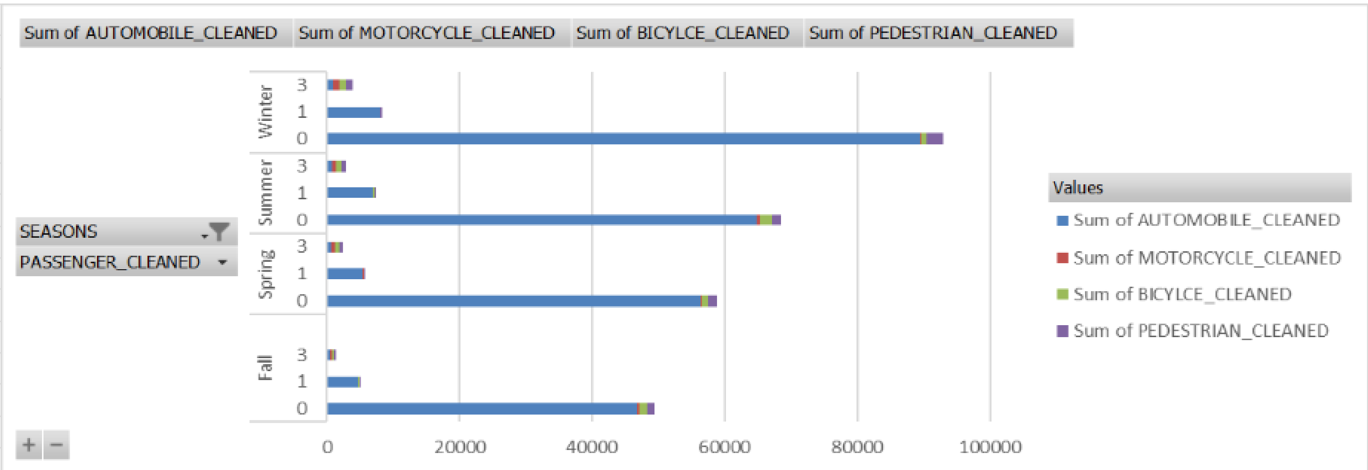
**A)Fatalities by season and day of week**

| Sum of FATALITIES | Column Labels | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Row Labels** | **Sunday** | **Monday** | **Tuesday** | **Wednesday** | **Thursday** | **Friday** | **Saturday** | **Grand Total** |
| Fall | 5 | 7 | 5 | 4 | 10 | 9 | 7 | 47 |
| Spring | 9 | 4 | 9 | 6 | 10 | 5 | 1 | 44 |
| Summer | 15 | 9 | 11 | 10 | 10 | 12 | 12 | 79 |
| Winter | 4 | 9 | 12 | 10 | 11 | 12 | 16 | 74 |
| **Grand Total** | **33** | **29** | **37** | **30** | **41** | **38** | **36** | **244** |




## 4. Involved Parties by Season

| Row Labels | Sum of AUTOMOB | Sum of MOTORCYCLE_CLEA | Sum of BICYLCE_CLEAN | Sum of PEDESTRIAN_CLEANED |
|---|---|---|---|---|
| **Fall** | **51742** | **725** | **1707** | **1670** |
| 0 | 46699 | 329 | 1205 | 1182 |
| 1 | 4674 | 27 | 133 | 119 |
| 3 | 363 | 363 | 363 | 363 |
| | 6 | 6 | 6 | 6 |
| **Spring** | **62428** | **910** | **1710** | **1942** |
| 0 | 56337 | 260 | 985 | 1194 |
| 1 | 5464 | 23 | 98 | 121 |
| 3 | 627 | 627 | 627 | 627 |
| **Summer** | **72392** | **1347** | **2777** | **2099** |
| 0 | 64768 | 541 | 1838 | 1221 |
| 1 | 6889 | 71 | 204 | 143 |
| 3 | 735 | 735 | 735 | 735 |
| **Winter** | **98520** | **1154** | **1850** | **3628** |
| 0 | 89507 | 160 | 792 | 2419 |
| 1 | 8032 | 13 | 77 | 228 |
| 3 | 981 | 981 | 981 | 981 |
| **Grand Total** | **285082** | **4136** | **8044** | **9339** |



# Location-Based & Hotspot Analysis

This section analyses traffic collisions in Toronto, pinpointing major high-risk areas, neighbourhoods, and trends among various road user groups.

**1. Collision Counts by District:** The distribution of collisions varies significantly among districts. Districts D32 (20,973 collisions), D42 (21,624 collisions), and D43 (15,765 collisions) report the highest numbers of incidents. These regions are likely aligned with densely populated or high-traffic urban centres, highlighting the necessity for targeted safety measures.

**2. Neighbourhood-Level Collision Hotspots:** At the neighbourhood level, areas such as Wexford/Maryvale (7,316 collisions), West Humber-Clairville (5,761 collisions), and York University Heights (4,639 collisions) stand out as hotspots due to their consistently high collision rates. These neighbourhoods feature a blend of residential, commercial, and institutional areas, leading to intricate traffic dynamics.

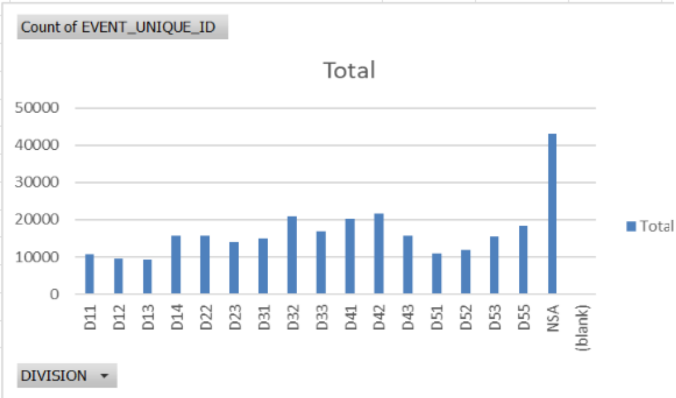
**3. Fatality Distribution by Neighbourhood:** Fatal collisions are spread across different neighbourhoods, including Agincourt North, Banbury-Don Mills, and Avondale, which have recently reported fatalities. While these fatal incidents are less common compared to the total number of collisions, their presence highlights the need for focused interventions in these areas.

**4. Bicycle Collisions by Neighbourhood:** Bicycle collisions display a clear spatial pattern, with neighbourhoods such as Annex (181 collisions), Bay-Cloverhill (74), and Agincourt South-Malvern West (61) exhibiting higher totals. This may indicate increased cycling activity, likely due to their closeness to cycling infrastructure or popular commuter routes. The findings underscore the need for improved cyclist safety initiatives, such as dedicated bike lanes and enhanced signage in these areas.

**5. Year-on-Year Collision Trends:** An analysis of yearly collision data shows varied trends among neighbourhoods. For example, Wexford/Maryvale experienced a rise in collisions from 1,669 in 2023 to 1,943 in 2024, whereas West Humber-Clairville had a modest increase from 1,364 to 1,386 during the same timeframe. Some neighbourhoods exhibit stable or slightly declining collision figures. Keeping an eye on these trends can inform responsive policies that address emerging risks or enhance traffic safety.

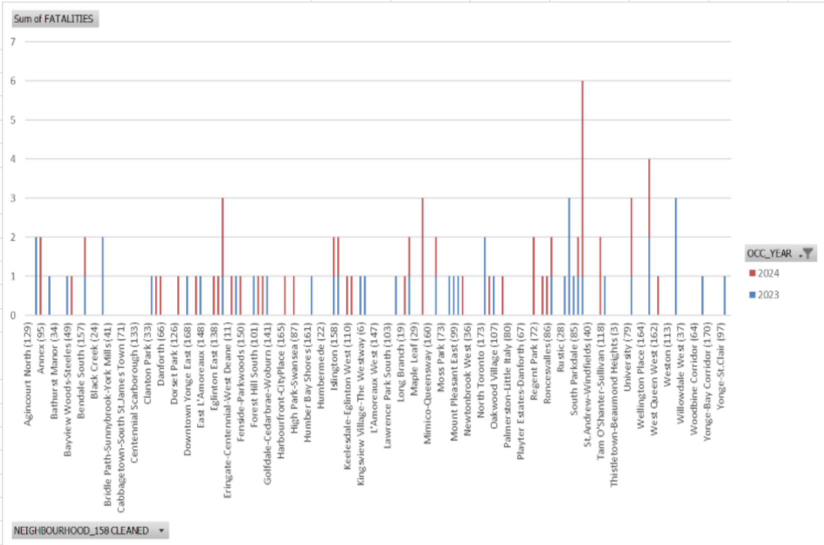
| Row Labels | Count of EVENT_UNIQUE_ID |
|---|---|
| D11 | 10620 |
| D12 | 9635 |
| D13 | 9312 |
| D14 | 15613 |
| D22 | 15689 |
| D23 | 14137 |
| D31 | 15074 |
| D32 | 20973 |
| D33 | 16900 |
| D41 | 20180 |
| D42 | 21624 |
| D43 | 15765 |
| D51 | 10858 |
| D52 | 11964 |
| D53 | 15371 |
| D55 | 18326 |
| NSA | 43055 |
| (blank) | |
| **Grand Total** | **285096** |

Collision rate based on division



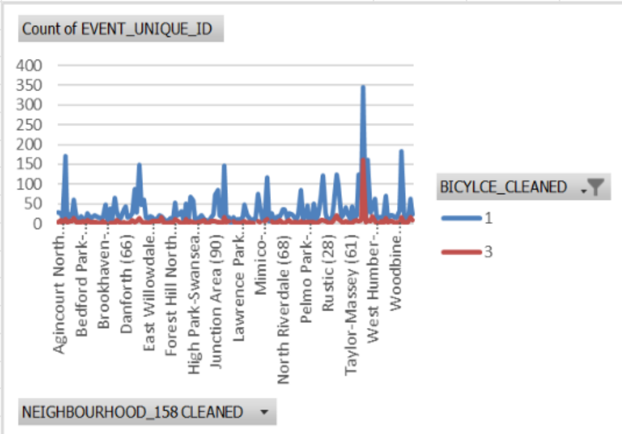
Fatalities in every neighbourhood for the following years

| Sum of FATALITIES | Column Labels | | |
|---|---|---|---|
| Row Labels | 2023 | 2024 | Grand Total |
| Agincourt North (129) | 0 | 0 | 0 |
| Agincourt South-Malvern West (128) | 0 | 0 | 0 |
| Alderwood (20) | 2 | 0 | 2 |
| Annex (95) | 0 | 2 | 2 |
| Avondale (153) | 0 | 0 | 0 |
| Banbury-Don Mills (42) | 1 | 0 | 1 |
| Bathurst Manor (34) | 0 | 0 | 0 |
| Bay-Cloverhill (169) | 0 | 0 | 0 |
| Bayview Village (52) | 0 | 0 | 0 |
| Bayview Woods-Steeles (49) | 1 | 0 | 1 |
| Bedford Park-Nortown (39) | 0 | 1 | 1 |
| Beechborough-Greenbrook (112) | 0 | 0 | 0 |
| Bendale South (157) | 0 | 0 | 0 |
| Bendale-Glen Andrew (156) | 1 | 1 | 2 |
| Birchcliffe-Cliffside (122) | 0 | 0 | 0 |
| Black Creek (24) | 0 | 0 | 0 |
| Blake-Jones (69) | 0 | 0 | 0 |



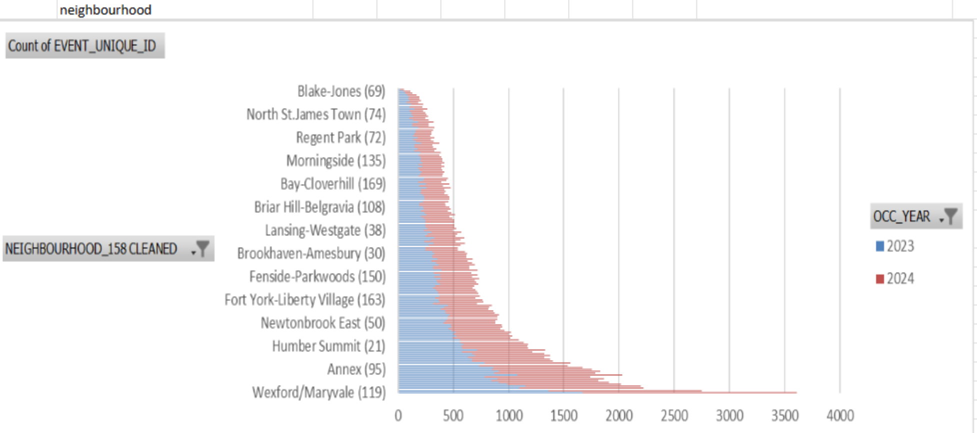

| Count of EVENT_UNIQUE_ID | Column Labels | | |
|---|---|---|---|
| Row Labels | 1 | 3 | Grand Total |
| Agincourt North (129) | 28 | 4 | 32 |
| Agincourt South-Malvern West (128) | 25 | 9 | 34 |
| Alderwood (20) | 13 | 1 | 14 |
| Annex (95) | 170 | 11 | 181 |
| Avondale (153) | 3 | 3 | 6 |
| Banbury-Don Mills (42) | 23 | 7 | 30 |
| Bathurst Manor (34) | 21 | 7 | 28 |
| Bay-Cloverhill (169) | 61 | 13 | 74 |
| Bayview Village (52) | 15 | 3 | 18 |
| Bayview Woods-Steeles (49) | 2 | 3 | 5 |
| Bedford Park-Nortown (39) | 18 | 4 | 22 |
| Beechborough-Greenbrook (112) | 10 | 6 | 16 |
| Bendale South (157) | 11 | 1 | 12 |
| Bendale-Glen Andrew (156) | 26 | 9 | 35 |
| Birchcliffe-Cliffside (122) | 15 | 3 | 18 |
| Black Creek (24) | 15 | 4 | 19 |
| Blake-Jones (69) | 21 | 4 | 25 |
| Briar Hill-Belgravia (108) | 19 | 3 | 22 |

Bicycle collision based on every neighbourhood




| Count of EVENT_UNIQUE_ID | Column Labels | | |
|---|---|---|---|
| Row Labels | 2023 | 2024 | Grand Total |
| Wexford/Maryvale (119) | 1669 | 1943 | 3612 |
| West Humber-Clairville (1) | 1364 | 1386 | 2750 |
| Dorset Park (126) | 1098 | 1125 | 2223 |
| York University Heights (27) | 1150 | 1047 | 2197 |
| Banbury-Don Mills (42) | 989 | 1028 | 2017 |
| Milliken (130) | 914 | 995 | 1909 |
| Bendale-Glen Andrew (156) | 845 | 970 | 1815 |
| Woburn North (142) | 902 | 964 | 1866 |
| South Parkdale (85) | 782 | 960 | 1742 |
| St Lawrence-East Bayfront-The Islands (166) | 1080 | 953 | 2033 |
| Yorkdale-Glen Park (31) | 870 | 917 | 1787 |
| Etobicoke City Centre (159) | 916 | 916 | 1832 |
| Annex (95) | 846 | 908 | 1754 |
| Yonge-Bay Corridor (170) | 863 | 806 | 1669 |
| Morningside Heights (144) | 735 | 801 | 1536 |
| South Riverdale (70) | 782 | 781 | 1563 |
| Harbourfront-CityPlace (165) | 667 | 736 | 1403 |
| Agincourt South-Malvern West (128) | 668 | 708 | 1376 |
| St.Andrew-Windfields (40) | 633 | 690 | 1323 |
| Wellington Place (164) | 693 | 683 | 1376 |

Year on year collisions in every neighbourhood



# Historical Trends Over Time

This analysis investigates whether the occurrence of traffic collisions in Toronto is significantly associated with seasonal variations over time. Specifically, it examines if collisions (occurrence: Yes or No) differ meaningfully across the four seasons: Fall, Winter, Spring, and Summer.

Using a chi-square test of independence, the observed counts of collision occurrences were compared against expected counts under the assumption that season and collision occurrence are independent. The test results revealed a chi-square statistic of approximately **0.35** with a p-value greater than **0.05**.

Since the p-value exceeds the standard significance threshold, we **fail to reject the null hypothesis** that collision occurrence is independent of season. This indicates that the likelihood of a collision happening does not significantly depend on the season. In other words, collisions occur roughly in the expected proportions across all seasons, suggesting no strong seasonal effect on whether collisions happen at all.

These findings suggest that while seasonal factors may affect the severity or timing of collisions, the overall probability of collision occurrence remains relatively stable throughout the year. Future research could explore more granular time-based trends or interactions with other variables, such as weather conditions or traffic volume, to uncover subtle temporal effects on collision dynamics.

**Problem Q:** Is there a significant association between season and the occurrence of collisions (Yes vs. No)?

**$H_0$:** Collision occurrence is independent of season (no association between season and collision outcome).

**$H_1$:** Collision occurrence depends on season (there is an association).

Yes = 1
N0 = 0

**Actual Range**

| Collision | Fall | Winter | Spring | Summer | Total |
|---|---|---|---|---|---|
| Yes | 253938 | 291363 | 237197 | 247275 | 1029773 |
| No | 492222 | 563925 | 457018 | 477717 | 1990882 |
| Total | 746160 | 855288 | 694215 | 724992 | 3020655 |

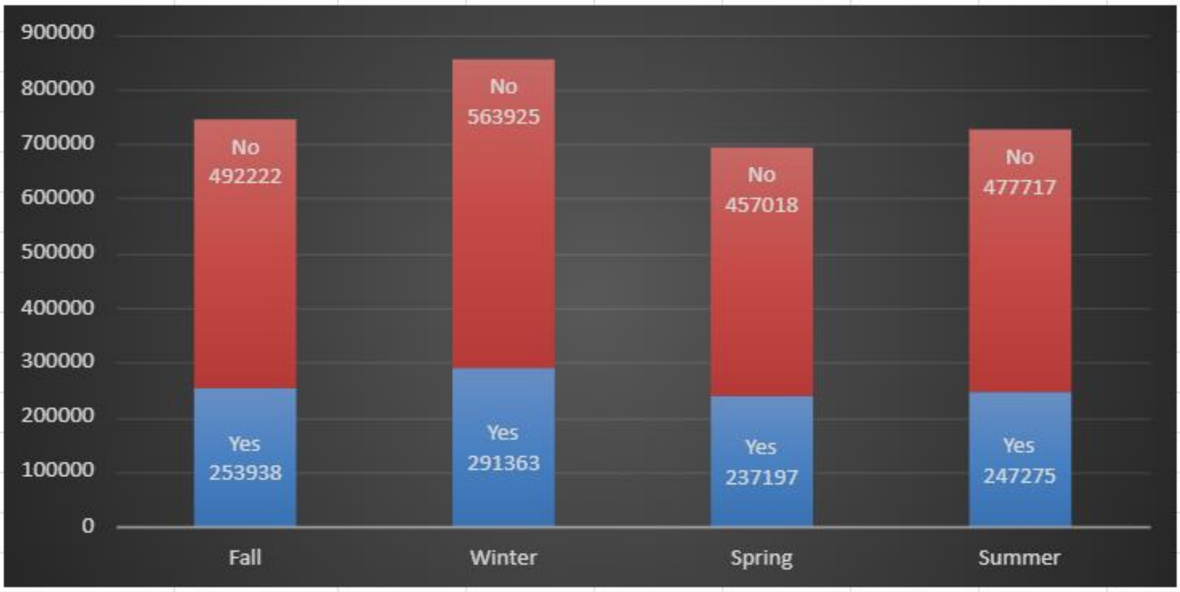
**Expected Range**

| Collision | Fall | Winter | Spring | Summer |
|---|---|---|---|---|
| Yes | 254373.8 | 291576.7 | 236665.2 | 247157.4 |
| No | 491786.2 | 563711.3 | 457549.8 | 477834.6 |

The actual and expected counts are very similar in each cell

**Chi Square Test Value: 0.35206958**

**Since p-value > 0.05, you fail to reject the null hypothesis ($H_0$)**
This means there is **no statistically significant association** between season and collision occurrence.
In other words, collisions (yes/no) seem to occur in about the proportions we'd expect in each season, purely by chance.
Season does not appear to affect the likelihood of a collision happening versus not happening.



# Strategies and Recommendations

Based on the comprehensive analysis of Toronto's traffic collision data, including temporal patterns, weather impact, location-based hotspots, and historical trends, several targeted strategies are recommended to improve road safety and reduce traffic-related injuries and fatalities in alignment with the City of Toronto's

**Vision Zero Plan:**

- **1. Targeted Time-of-Day Interventions**
- **• Focus Enforcement During Peak Hours**: By recognizing that the busiest collision times usually occur between 1–2 PM each day, we can boost police presence and traffic enforcement during this key hour to help discourage risky behaviours like speeding and distracted driving.
- **• Adjust Traffic Signal Timings:** Collaborate with city traffic management to fine-tune signal timings during those high-risk afternoons and evening periods, aiming to make our roads safer by reducing congestion and collision risks.

**2. Seasonal and Weather-Responsive Measures**

- **Improve Winter Road Safety:** Winter months often see the most accidents and serious collisions, so it's essential to focus on road maintenance, which includes prompt snow clearance, salting, and ice management. Share winter driving warnings and encourage careful driving during dangerous weather conditions.
- **Seasonal Public Awareness Initiatives:** Launch safety campaigns before seasonal changes, such as winter and spring, to raise awareness about evolving road conditions and the presence of vulnerable road users.
- **3. Infrastructure and Enforcement by Location**
- **Prioritize High-Collision Areas:** Implement targeted strategies in high-risk hotspots, such as D32, D42, D43, and neighbourhoods with significant bicycle collisions, including the Annex and Agincourt South-Malvern West. These strategies aim to enhance traffic safety through measures such as traffic calming, better street lighting, improved pedestrian crosswalks, and the addition of protected bike lanes.
- **Safeguard Vulnerable Road Users:** Intensify the focus on infrastructure and policies that protect cyclists and pedestrians in areas with high collision rates, ensuring the availability of safe routes near schools, hospitals, and commercial zones.

**4. Data-Driven Proactive Prevention**

- **Utilize Predictive Analytics Tools**: Leverage historical and spatial collision data to create predictive models that pinpoint emerging hotspots, enabling timely enforcement and infrastructure enhancements.
- **Ongoing Monitoring and Assessment:** Set up a real-time collision monitoring dashboard to observe trends, assess the impact of interventions, and adjust strategies in real-time**.**

**5. Community Engagement and Collaboration**

- **Local Stakeholder Partnerships:** Collaborate with community organizations, educational institutions, and local businesses to jointly develop safety initiatives that address the unique risks and needs of each neighbourhood.
- **• Education and Outreach Programs:** Enhance awareness through driver, cyclist, and pedestrian education initiatives that highlight risk factors revealed by data analysis, focusing on the significance of caution during peak collision periods and seasonal dangers.

# Conclusion

Road safety in a vast and bustling city like Toronto presents a multifaceted challenge, yet this analysis highlights distinct patterns that can guide us toward meaningful enhancements. We observe that specific times, particularly early afternoons and the winter months, pose greater risks, while specific neighbourhoods experience more accidents than others. By recognizing these trends, we can better direct our resources to areas that need them most.

Using data as our guide, the city and police services have the chance to develop targeted strategies that surpass one-size-fits-all solutions. By integrating thoughtful infrastructure changes, focused enforcement, and community involvement, we can create safer streets for everyone, whether they're driving, cycling, or walking.

Making Toronto's roads safer is an ongoing effort, but with these insights, we are taking significant steps toward a future where fewer lives are lost and every trip feels safer.

# Appendix

Susser, D. (2020). 12. Predictive Policing and the Ethics of Pre-emption. Social Science Research Network. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3875917

Analyzing True Yield Variation by Maturity Date in FGN Bonds Dataset | Vizly. https://vizly.fyi/share/0876c91c-009d-46a3-9328-04e711e1a87c/4ee62c05-d739-400a-b54d-7f2686fc613a