Rotman

TPS Case Competition

Team Co-MMA

November 26, 2020
Prepared by Humza Butt, Kevin Zhang, Alan Liu, Devashree Kumar

Rotman School of Management
UNIVERSITY OF TORONTO

Agenda

Toronto Police Service

- **Executive Summary**
- **Background & Business Problem**
- **Problem Statement**
  - Managerial Question
  - Analytical Question
- **Key Data Sources Overview**
- **Exploratory Analysis**
- **Further Analysis**
- **Modeling Approach**
  - Model Trustworthiness
- **Key Findings / Recommendation**
- **Next Steps**

Executive Summary

Rotman

TPS Case

Business Problem

- Reduction of Fatal casualties
- Reduction in # of Severe injuries
- Address safety of most vulnerable users of the transportation system
  - Pedestrians, cyclists, adults, teens and etc.

Data Used

Dataset Name	Description
KSI Data	Traffic Collision: Killed or Seriously Injured (KSI) from 2006 – 2019
Reported Crime	Auto Theft, Break and Enter, Robbery and Theft from 2014 – 2019

Main Findings

- Non-fatal Accidents 

  Fatal Accidents 

- Alcohol and Speeding peaks at **2:00 AM**
- Rush Hours **=** Aggressive Driving
- Correlation: Crime & Fatal Accidents

Recommendation

- Use Dashboard
- Predict the fatal beforehand to allow for officers to act more urgently.
- Allocate resources in specific neighborhoods
- Allocate more police force on specific streets for speeding and Alcohol

Business Problem

Rotman

Current Situation

Business Overview

- Established in 1834
- It was the first local police service created in North America
- Primary agency responsible for providing law enforcement and policing services in Toronto

Purpose / Problems

- 318 fatal & non-fatal injuries in Toronto for 2019.
- Non-fatal injuries - Declining

  Fatal traffic accidents - Consistent
- Due to limited resources, device smarter ways to reduce traffic related injuries and fatalities.

Vision Zero Plan

<u>Vision Zero Plan</u>

Objective : The city of Toronto is committed to reduce traffic-related deaths & injuries to zero.

- Comprehensive 5-year (2017-2021) action plan focused on reducing traffic-related fatalities & serious injuries
- About 10% reduction in pedestrian fatalities from 2013 to 2019

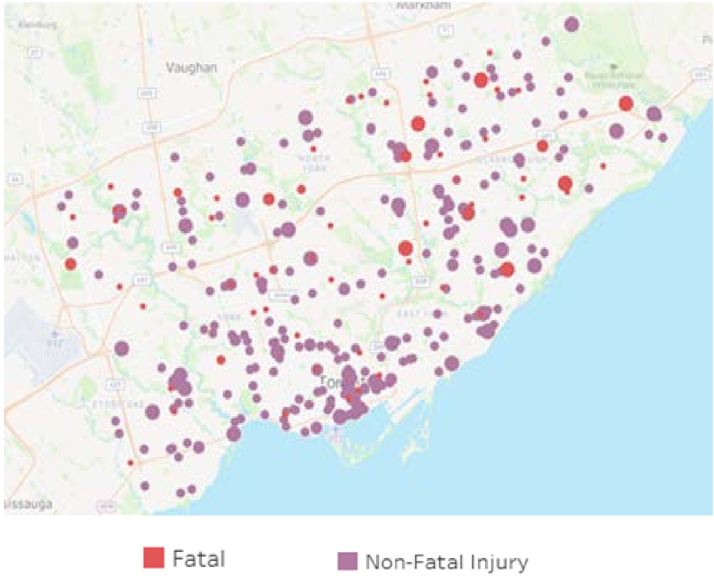
Managerial and Analytical Objectives

Problem Statement

- **Managerial question**
  - How can TPS reduce fatality and serious injuries related to traffic accidents?
  - Who/What should TPS focus their resources? What safety measurement can TPS implement?
  - How can TPS evaluate these implementations?

- **Analytical question**
  - What are the significant variables that impact the likelihood of an accident being Fatal or Non-Fatal?
  - What are the external factors that influence traffic accidents?

2019 Traffic Fatality Map



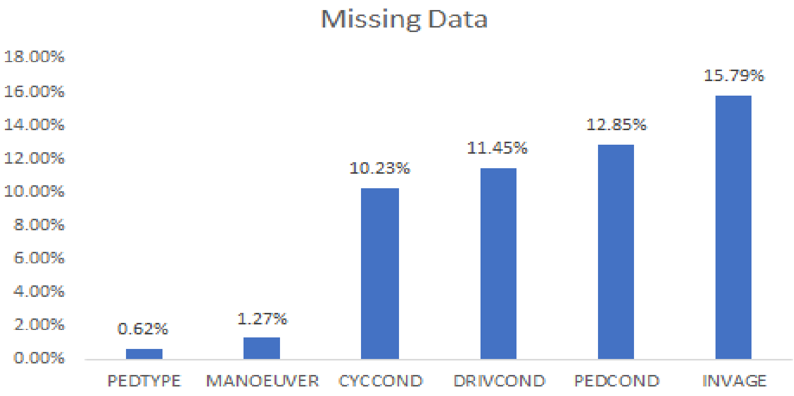
Key Data Sources

Rotman

Handling Data

Problems in Data

- Missing value “unknown” and various empty cells.
- Created indicator variables of variables where only “Yes” is shown and the rest is blank.
- Handle missing data by keeping missing data and categorizing as its own factor.



Data Sources Used

Dataset Name	Description
KSI Data	Traffic Collision: Killed or Seriously Injured (KSI) from 2006 – 2019
Reported Crime	Auto Theft, Break and Enter, Robbery and Theft from 2014 – 2019

Description of Data

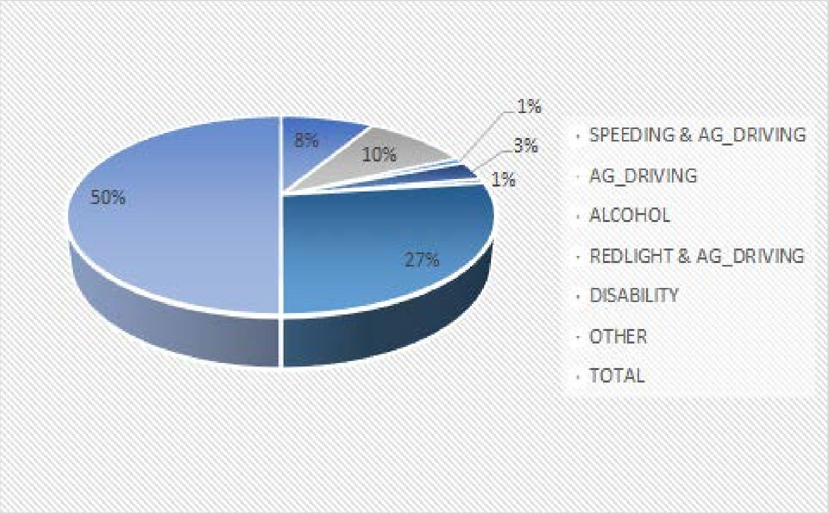
- Single row represents an individual that was involved in a traffic accident at a certain point of time.
- ACCNUM is a unique ID that represents a single traffic accident.
- There can be multiple individuals in an accident, but only one accident per individual.

Exploratory Analysis

Exploratory Stage

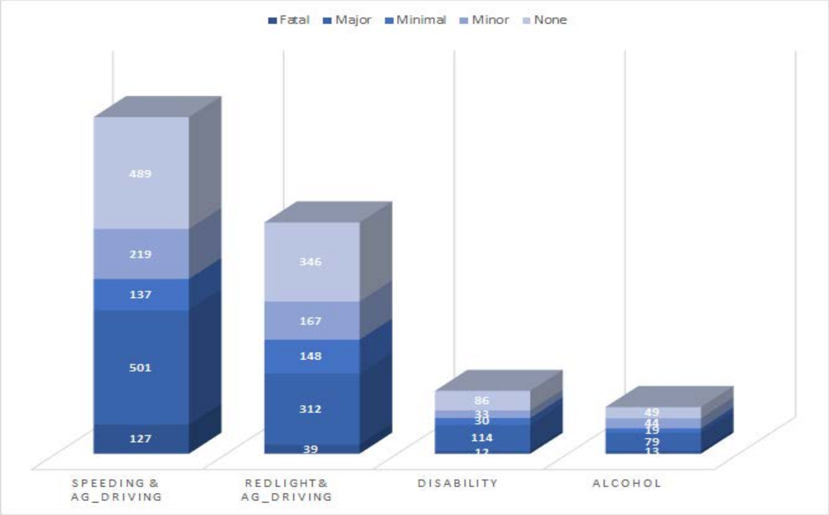
Rotman

Causes of Fatalities



SPEEDING & AG_DRIVING	378	17%
AG_DRIVING	424	19%
ALCOHOL	43	2%
REDLIGHT & AG_DRIVING	136	6%
DISABILITY	32	1%
OTHER	1195	54%
TOTAL	2208	100%

Causes of Injury



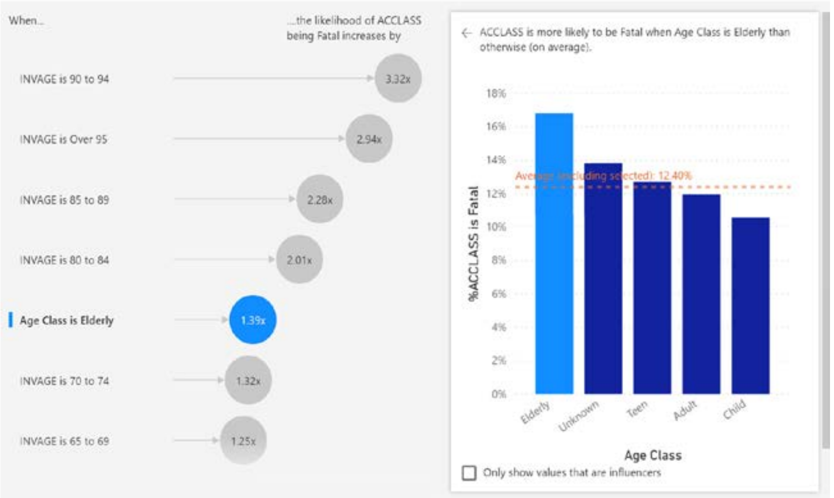
INJURY	SPEEDING & AG_DRIVING	REDLIGHT& AG_DRIVING	DISABILITY	ALCOHOL
Fatal	127	39	12	13
Major	501	312	114	79
Minimal	137	148	30	19
Minor	219	167	33	44
None	489	346	86	49
TOTAL	1650	1116	319	233

Exploratory Stage

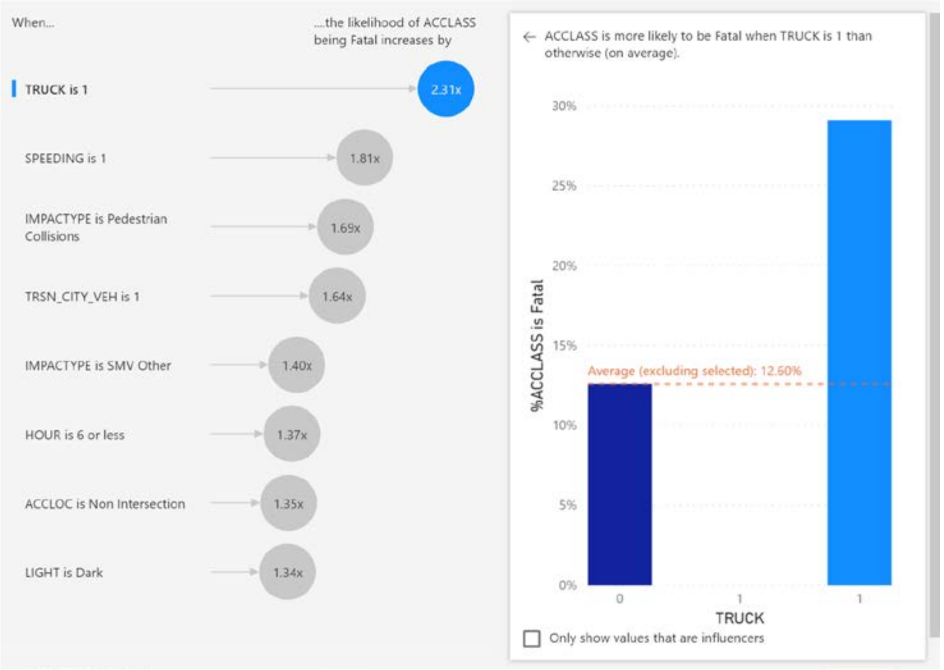
Rotman

What are the factors that are highly related to fatal accidents?

Age



Other Factors: Truck, Speeding, Collision Type...



- **Age Class:**
  - Variable derived from INVAGE.
- **Factor variables (eg. Truck):**
  - 1 means the accident involves truck.

Dashboard

Overview

Graph for Fatal and Non-Fatal Injury Count by Year

Filter for Fatal & Non-Fatal

Total Accident #

Pie chart for the factors involving in accidents (alcohol, speeding, etc.)

Pie chart for the type of vehicle involving in accidents (cyclist, truck, etc.)

Accident Map (color indicate districts)

Date Slicer (select date range)

Bar chart for Accidents by Districts

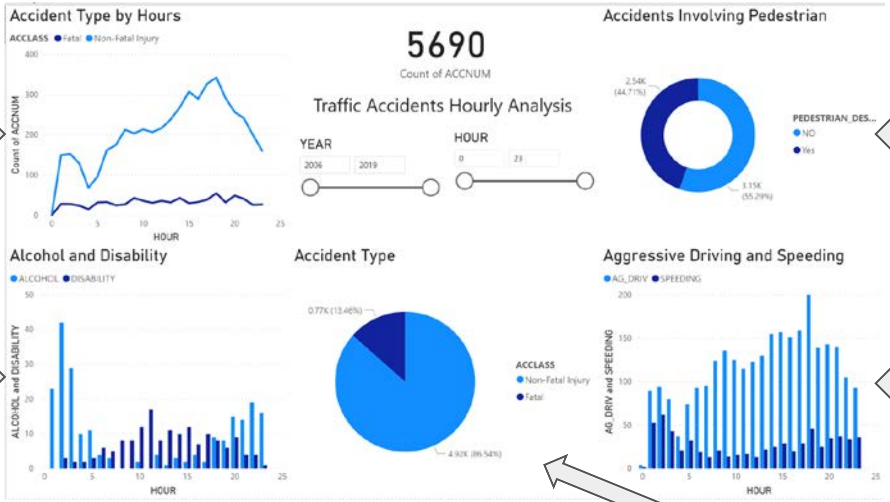
Pie chart for the impact type

Dashboard

Hourly Analysis

Number of Accident type by hours (Fatal vs. Non-Fatal)

Number of Accident involving Alcohol and Disability by hours



Number of Accident that involves Pedestrians

Number of Accident involving Aggressive Driving and Speeding by hours

Number of Accident type (Fatal vs. Non-Fatal)

- Alcohol Consumption and Speeding related accidents peaked at 2 am
- Aggressive Driving related accidents peaked at 9 am and 6 pm (Rush Hours)
- Disability involved in accidents majority during daytime.

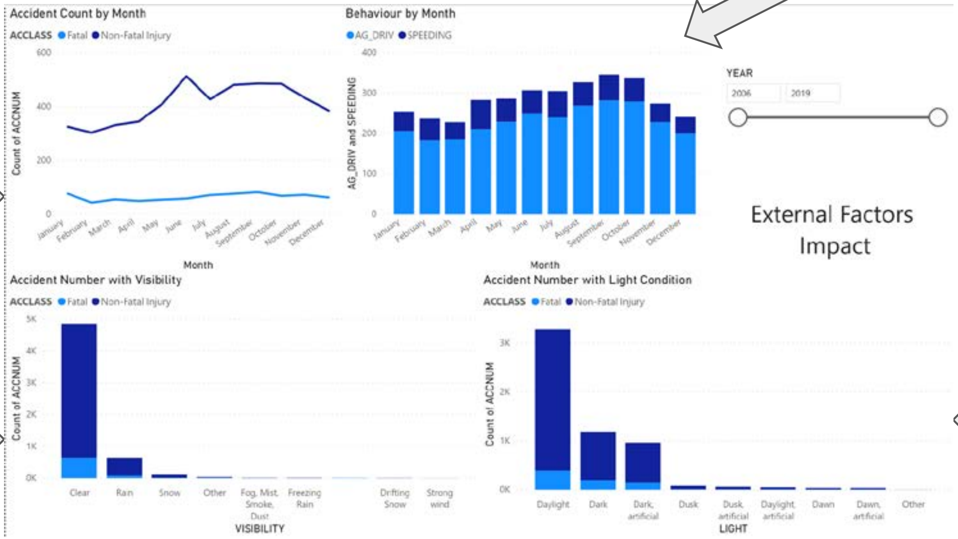
Dashboard

External Factors

- Fatal accidents consistent throughout the seasons
- Road condition does not significantly contribute fatal accidents
- Drivers are less aggressive in Winters (where the road condition is bad)

Number of Accident type (Fatal vs. Non-Fatal)

Number of Accident type by Month (Fatal vs. Non-Fatal)



Number of Accident with respect to Visibility

Number of Accident with respect to Light

Dashboard

Rotman

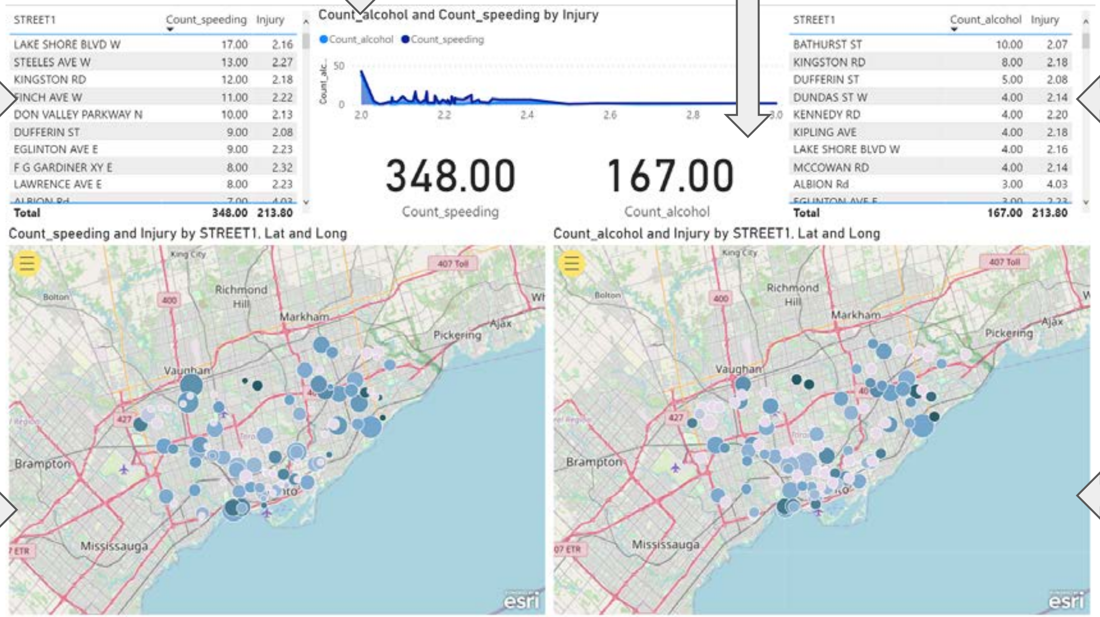
Street Analysis

Graph with the seriousness of injury related to Speeding and Alcohol

Counts related to speeding and alcohol cases

List of Street with most speeding related accidents

List of Street with most alcohol related accidents



Street Map involving Speeding (size indicates the cases of speed related accidents)

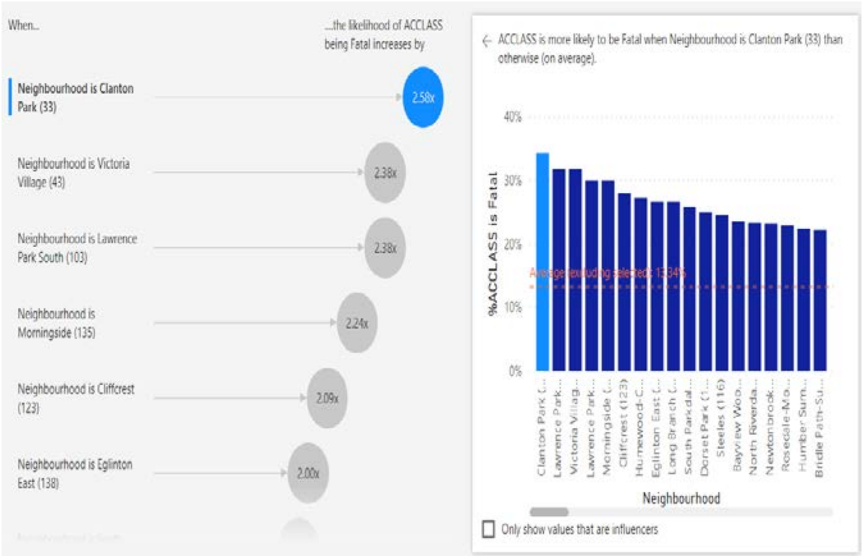
Street Map involving Alcohol (size indicates the cases of speed related accidents)

Dashboard

Rotman

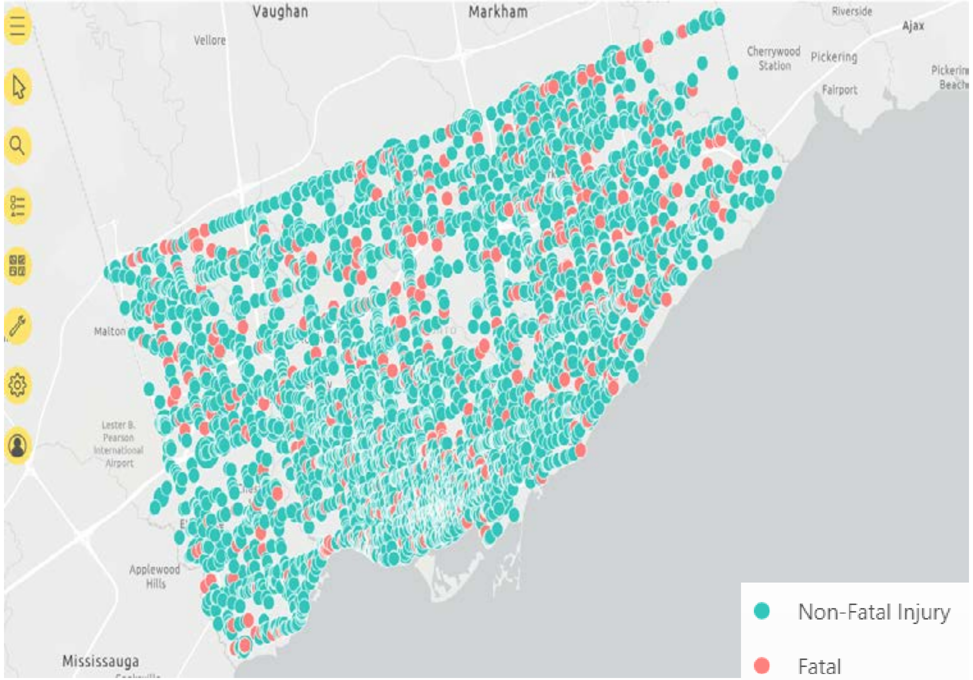
Streetwise Fatality Analysis

Neighbourhood Analysis



- 4 neighbourhoods with highest likelihood of Fatal accidents - Clanton Park, Victoria Village, Lawrence Park(S) & Morningside

Fatal & Non-Fatal Injury Analysis

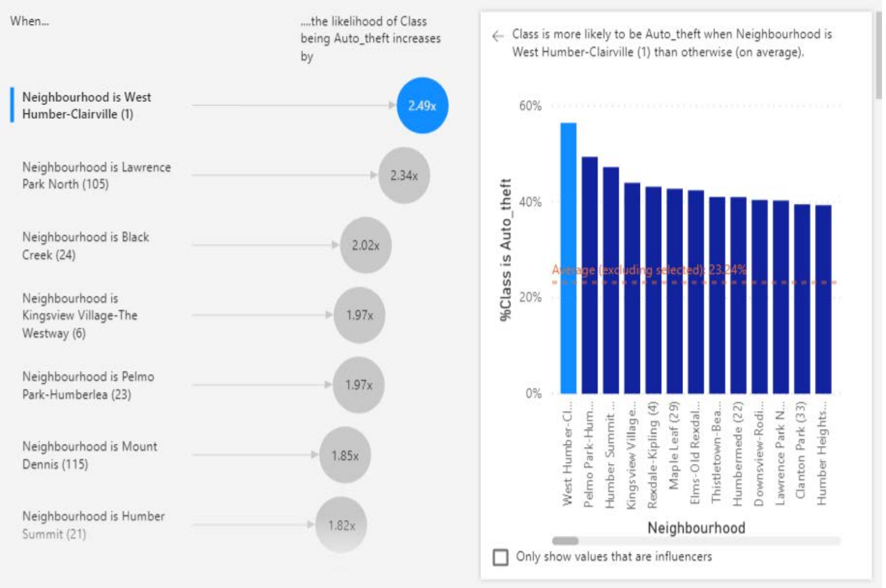


Dashboard

Rotman

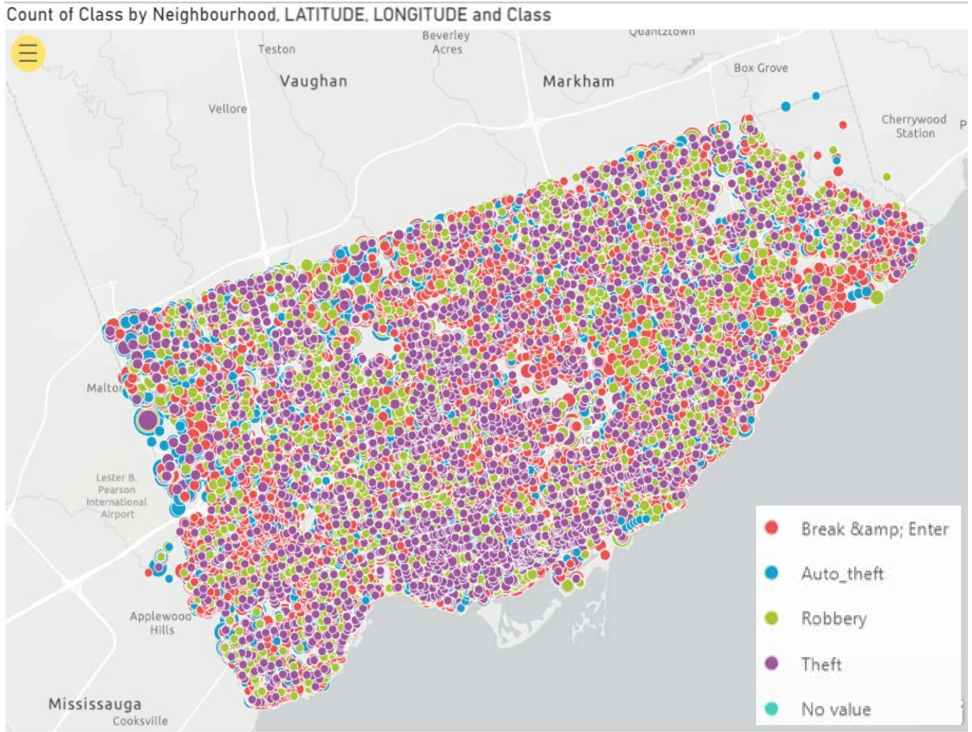
Streetwise Crime Analysis

Likelihood of crime in neighbourhood



- 4 neighbourhoods with highest likelihood of fatal accidents are also the neighbourhoods with higher crime.

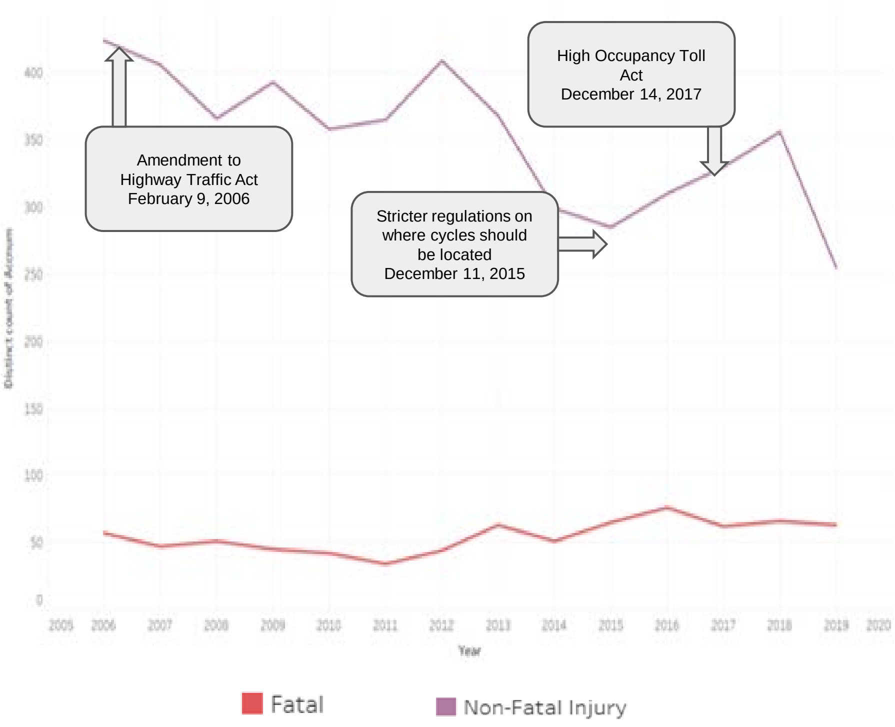
Analysis of various crimes in different neighbourhoods



Fatal vs Non-Fatal Injuries

Trends

Fatal and Non-Fatal Injuries over Time



Insight

- Non-fatal injury traffic accidents have been steadily decreasing over the years.
- Fatal injury traffic accidents have stayed consistent throughout the years.

Recommendations

- Create a model that can be used by officers to determine likelihood an accident may be fatal beforehand.
- Data can be collected from 911 calls and can immediately be used to predict likelihood of fatality.
- Save lives of those that may not have been immediately killed during a crash, but are likely to be.

Modelling Approach

Modeling Approach

Rotman

Logistic Regression

Model Overview

- Logistic Regression to determine likelihood of crash being fatal or non-fatal.
- Only using variables that can be obtained before officer gets to a collision.
- Data was split 70/30 into training and test set to test for overfitting.

Model Accuracy

	Training Set	Test Set
Accuracy	86.85%	86.07%
False Positive Rate	12.82%	13.49%

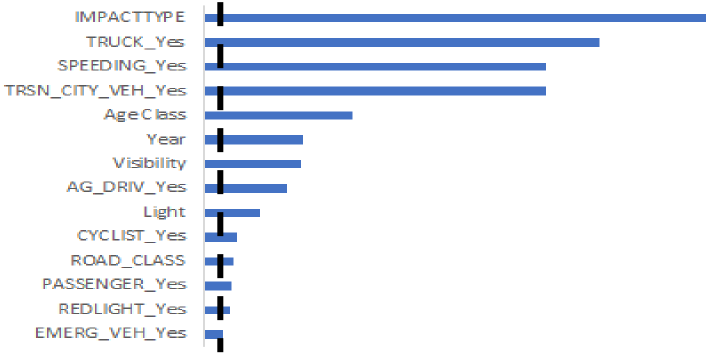
Predictors and Response Variable

Predictors

Road Class	Speeding	Age	Visibility	Redlight
Cyclist	Passenger	Emergency	Vehicle	Light
City Vehicle	Truck	Impact Type		

Response Variable: ACCLASS

Predictor Ranking by Significance

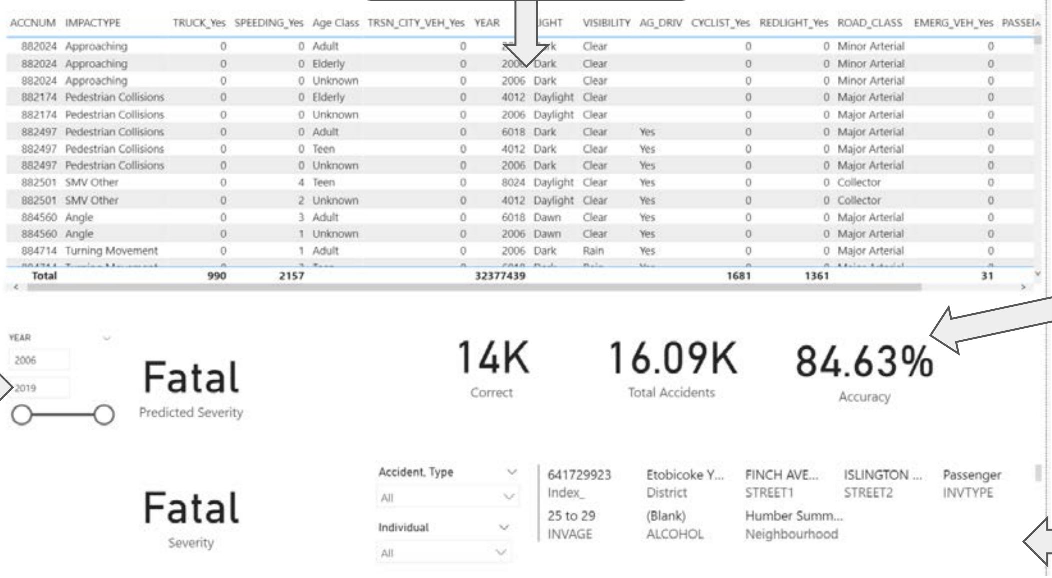


Model Dashboard

Rotman

Implementation of Model

Table that represents the type of data needed for the model to make predictions



ACCNUM	IMPACTYPE	TRUCK_Yes	SPEEDING_Yes	Age Class	TRSN_CITY_VEH_Yes	YEAR	IGHT	VISIBILITY	AG_DRIV	CYCLIST_Yes	REDLIGHT_Yes	ROAD_CLASS	EMERG_VEH_Yes	PASSE
882024	Approaching	0	0	Adult	0	[illegible]	[illegible]	Clear		0	0	Minor Arterial	0	
882024	Approaching	0	0	Elderly	0	[illegible]	Dark	Clear		0	0	Minor Arterial	0	
882024	Approaching	0	0	Unknown	0	2006	Dark	Clear		0	0	Minor Arterial	0	
882174	Pedestrian Collisions	0	0	Elderly	0	4012	Daylight	Clear		0	0	Major Arterial	0	
882174	Pedestrian Collisions	0	0	Unknown	0	2006	Daylight	Clear		0	0	Major Arterial	0	
882497	Pedestrian Collisions	0	0	Adult	0	6018	Dark	Clear	Yes	0	0	Major Arterial	0	
882497	Pedestrian Collisions	0	0	Teen	0	4012	Dark	Clear	Yes	0	0	Major Arterial	0	
882497	Pedestrian Collisions	0	0	Unknown	0	2006	Dark	Clear	Yes	0	0	Major Arterial	0	
882501	SMV Other	0	4	Teen	0	8024	Daylight	Clear	Yes	0	0	Collector	0	
882501	SMV Other	0	2	Unknown	0	4012	Daylight	Clear	Yes	0	0	Collector	0	
884560	Angle	0	3	Adult	0	6018	Dawn	Clear	Yes	0	0	Major Arterial	0	
884560	Angle	0	1	Unknown	0	2006	Dawn	Clear	Yes	0	0	Major Arterial	0	
884714	Turning Movement	0	1	Adult	0	2006	Dark	Rain	Yes	0	0	Major Arterial	0	
Total		990	2157			32377439				1681	1361		31	

Accuracy rate of model and absolute value of number of correct predictions

Predicted and actual results of an accident for an individual. Data is adjustable by year

Summary board to describe selected individual and where they were located

Dashboard Demonstration

Conclusion

Rotman

TPS

Overall Findings

Non-fatal Accidents 

Fatal Accidents 

- Significant factors for fatal accidents

Road Class	Speeding	Age	Visibility	Redlight
Cyclist	Passenger	Emergency	Vehicle	Light
City Vehicle	Truck	Impact Type		

- Aggressive Driving during rush hour, leading to the most accidents.
- Alcohol and speeding related accidents occur mostly around 2AM.
- Neighborhoods with the most likelihood of having a fatal accident coincide with the ones with the highest likelihood of a crime occuring.

Tactical Recommendations

- Create a model to predict the likelihood of a traffic accident becoming fatal beforehand to allow for officers to act more urgently.
- Allocate resources in neighborhoods where there are higher crimes and higher likelihoods of fatal accidents.
- Put more police force on streets where speeding and driving under influence (eg,. Alcohol) are likely to happen at the specific hours (eg,. 2 am).
- Use our dashboard

Next Steps

Rotman

Future Considerations

Data Request

Data Request	Reasoning
Data on when a fatality occurred	To determine if fatalities from traffic accidents occurred at the moment of accident or during ambulance ride or hospital. To better utilize the model to save lives.
Variable: Who is responsible for the accident?	Can be useful predictor to determine who is more injured or likely to be in a fatal accident.

Next Steps

- Use different modeling algorithms to increase accuracy of predictions and reduce false positive rates.
- Discuss with stakeholders about what information should be collected by police for their reports.
- Evaluate effectiveness of dashboard and model by using fatality count as a KPI.
- Look at legislation and determine the impact of that on accident count in Toronto.

Thank You

Q&A

Appendix

Nominal Logistic Fit for ACCLASS

Effect Summary

Source	LogWorth	PValue
IMPACTYPE	49.124	0.00000
TRUCK_Yes	38.799	0.00000
SPEEDING_Yes	33.496	0.00000
TRSN_CITY_VEH_Yes	14.539	0.00000
Age Class	13.314	0.00000
YEAR	9.598	0.00000
VISIBILITY	9.437	0.00000
AG_DRIV_Yes	8.087	0.00000
LIGHT	5.449	0.00000
CYCLIST_Yes	3.171	0.00067
ROAD_CLASS	2.819	0.00152
PASSENGER_Yes	2.579	0.00263
REDLIGHT_Yes	2.430	0.00371
EMERG_VEH_Yes	1.772	0.01689

Nominal Logistic Fit for ACCLASS

Parameter Estimates

Term		Estimate	Std Error	ChiSquare	Prob>ChiSq
YEAR		0.05322732	0.0084136	40.02	<.0001*
CYCLIST_Yes		-2.5708077	1.1012515	5.45	0.0196*
ROAD_CLASS[Collector]	Biased	22.7177567	35306.149	0.00	0.9995
ROAD_CLASS[Expressway]	Biased	21.7678229	35306.149	0.00	0.9995
ROAD_CLASS[Laneway]	Biased	24.2131959	35306.149	0.00	0.9995
ROAD_CLASS[Local]	Biased	22.5544624	35306.149	0.00	0.9995
ROAD_CLASS[Major Arterial]	Biased	22.6301519	35306.149	0.00	0.9995
ROAD_CLASS[Major Arterial Ramp]	Biased	-136.03603	253887.45	0.00	0.9996
ROAD_CLASS[Minor Arterial]	Biased	22.4809029	35306.149	0.00	0.9995
ROAD_CLASS[Other]	Zeroed	0	0	.	.
LIGHT[Dark]		0.10797348	0.1709183	0.40	0.5276
LIGHT[Dark, artificial]		-0.4696495	0.173654	7.31	0.0068*
LIGHT[Dawn]		0.20952274	0.346259	0.37	0.5451
LIGHT[Dawn, artificial]		-0.3842372	0.3650334	1.11	0.2925
LIGHT[Daylight]		-0.2295281	0.1637789	1.96	0.1611
LIGHT[Daylight, artificial]		-0.4008061	0.3371658	1.41	0.2345
LIGHT[Dusk]		0.27821295	0.2548112	1.19	0.2749
LIGHT[Dusk, artificial]		-0.6192087	0.3417973	3.28	0.0700
TRUCK_Yes		1.38280345	0.0991019	194.70	<.0001*
AG_DRIV_Yes		-0.4265416	0.0751079	32.25	<.0001*
REDLIGHT_Yes		0.4024247	0.1366639	8.67	0.0032*
TRSN_CITY_VEH_Yes		0.89548618	0.1076115	69.25	<.0001*
EMERG_VEH_Yes	Unstable	-22.197082	37234.426	0.00	0.9995
PASSENGER_Yes		0.21680959	0.0718934	9.09	0.0026*
SPEEDING_Yes		1.19810198	0.0970292	152.47	<.0001*
Age Class[Adult]		-0.1091446	0.0665432	2.69	0.1010
Age Class[Child]		-0.157394	0.1804632	0.76	0.3831
Age Class[Elderly]		0.42699396	0.0685872	38.76	<.0001*
Age Class[Teen]		-0.2446483	0.1161012	4.44	0.0351*
VISIBILITY[Clear]	Unstable	8.37911569	11943.754	0.00	0.9994
VISIBILITY[Drifting Snow]	Unstable	-14.275368	47058.874	0.00	0.9998
VISIBILITY[Fog, Mist, Smoke, Dust]	Unstable	8.56397093	11943.754	0.00	0.9994
VISIBILITY[Freezing Rain]	Unstable	-13.791986	30650.166	0.00	0.9996
VISIBILITY[Other]	Unstable	9.93158585	11943.754	0.00	0.9993
VISIBILITY[Rain]	Unstable	8.30932675	11943.754	0.00	0.9994
VISIBILITY[Snow]	Unstable	7.41630348	11943.754	0.00	0.9995

Appendix

Rotman

Nominal Logistic Fit for ACCLASS

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
TRUCK_Yes	1	1	173.058111	<.0001*
AG_DRIV_Yes	1	1	33.2300413	<.0001*
REDLIGHT_Yes	1	1	8.41977831	0.0037*
TRSN_CITY_VEH_Yes	1	1	62.3375482	<.0001*
EMERG_VEH_Yes	1	1	5.70787925	0.0169*
PASSENGER_Yes	1	1	9.04455354	0.0026*
SPEEDING_Yes	1	1	148.788187	<.0001*
Age Class	4	4	68.4359342	<.0001*
VISIBILITY	7	7	58.0714271	<.0001*