

Case Assignment: 2022 Safe Roads Competition

Case

Toronto and Mexico City are some of the most populous municipalities in North America and significant efforts are being made across all levels of government to manage the challenges posed by increased urbanization. This case will aim to provide insights and recommendations to address some of the road safety challenges posed to Toronto and Mexico City.

Geotab

To address these challenges, more and more cities have been turning towards better use of data, analytics and most recently nascent Internet of Things (IoT) technologies. Geotab Inc. is a leader in GPS fleet management solutions collecting over 40 billion data points per day from sensors installed in road vehicles across North America, including Toronto and Mexico City.

Geotab collects a wide variety of IoT data including GPS telematics on hyper-local weather data, detailed accelerometer data, and even cell phone coverage. Geotab's advanced analytics (<https://data.geotab.com/>) provides higher-level datasets from their detailed sensor data. For this case, they have provided:

- **Hazardous Driving Area Dataset:** Identifies areas for which harsh braking and accidents have occurred based on aggregated onboard sensor data for vehicles that have their solution installed.
- **Road Impediments Dataset:** Identifies areas with potential road impediments by aggregating their z-axis (up/down) accelerometer data.
- **Intersection Metrics:** Provides metrics related to traffic flow and congestion at traffic-signal controlled intersections. Such metrics include average wait times, number of stops before which a vehicle is able to pass through the intersection, and the proportion of vehicles which stop at the light before passing through.

Note that all datasets are aggregations over the vehicles (and fleets) using Geotab GPS devices in a given location.

One of Geotab's initiatives involves a Smart City program where their datasets can be utilized to increase safety, mobility, efficiency, and citizen engagement.

Toronto Police Service

The Toronto Police Service (TPS) is the first municipal police service created in North America and among one of the oldest in the English-speaking world. It is also one of the leaders in data analytics with their Analytics and Innovation Unit which supports operational and strategic initiatives at TPS. They also maintain an open data portal (<http://data.torontopolice.on.ca/>) which is intended to improve transparency and confidence as well as public safety in Toronto.

One key strategic area for TPS is traffic safety as it relates to public awareness and enforcement. Recent data on traffic collisions resulting in serious injuries or death have been released publicly by the Toronto Police Service in partnership with the City of Toronto Transportation Safety Unit in order to collaborate broadly on approaches to traffic safety and to use data and analytics to drive decisions.

- **Toronto Police Services All Motor Vehicle Collisions:** includes a dataset with all motor vehicle collisions reported to TPS from 2014-2020.
- **Toronto Police Services Killed or Seriously Injured Dataset:** includes traffic collision events in the City of Toronto where a person was either Killed or Seriously Injured (KSI) from 2006 – 2020.

This dataset includes specific information about each KSI incident over the specified period. While this makes the KSI dataset more specific with respect to the most serious types of incidents compared to Geotab's Hazardous Driving Areas Dataset (which relates to all incidents of harsh braking – many of which may not have resulted in accidents at all), it is also, by necessity much sparser geographically, since the number of KSI incidents is quite limited. It is hoped that by combining the datasets provided, new insights on the areas with high potential for traffic collisions may emerge.

City of Toronto Data

Traffic safety around schools, hospitals, places of worship, parks and other buildings with heavy public usage are of particular concern. To help you identify these areas, two data sets maintained by the City of Toronto are available:

- **City of Toronto One Address Repository:** the authoritative address database, which includes exact geographical location of each city address, as well as the name, where it exists, of the location (e.g., church, school, park, etc.)
- **City of Toronto Centreline Intersection Dataset:** identifies all intersections in the city of Toronto

Mexico City Data

- **Accidentes por clase:** This dataset includes the number of accidents per class.
- **Accidentes por causa:** This dataset includes number of accidents per reason for the accident.
- **Número de víctimas:** This dataset includes number of victims per type.
- **Accidentes por tipo de accidente:** This dataset includes number of accidents per type.
- **Geotab (ignition-results):** This is the hazardous driving areas in Mexico City based on severity score and total incidents over the last month. The higher the score, the more hazardous the location.

Problem Statement

The main objective of this case is to understand the factors that affect traffic safety utilizing the datasets provided by Geotab, the Toronto Police Service, and other open data sources. The aim of the project is to extract insights from the data and provide specific recommendations to improve traffic safety in your municipality.

Some of the managerial questions you may wish to examine include:

- Are there relationships between locations with high number of harsh braking incidents, road impediments, and KSI incidents? To what extent are these relationships predictive? How can they be used in pro-active accident prevention initiatives?
- Is there evidence that the relationship between KSI incidents and the factors measured in Geotab data sets has been changing over time?
- Are there “sensitive” areas (e.g. in close proximity to schools, hospitals or other public buildings) that are particularly risky with respect to more serious accidents?

Outcomes from this competition can potentially help municipalities with initiatives related to safety and traffic enforcement:

- Examine public safety data through new lenses
- Speeding Reduction Campaigns
- Red Light Camera Locations
- Vision Zero Safety Measures Enforcement & Collaboration
- Service Gaps
- Leveraging IoT datasets to gain further insights

Submission

Your submission should be in the form of a video presentation to the Municipality by addressing some or all of the questions above. The presentation should be limited to 10 minutes and non-technical in nature; it can be followed by an Appendix containing more technical descriptions of your work. While you can also submit your code in a separate file, the presentation should be self-contained (i.e., it should not be necessary for the reader to open your code).

The presentation will be judged on:

- Managerial relevance of your work
- Soundness of your analysis
- Effectiveness of your communication, particularly data visualization
- Actionability of your conclusions and recommendations

Datasets

SAS has preloaded datasets to the [SAS OnDemand for Academics](https://data.sas.com/academic) platform for your use in this competition. We will review how to access these datasets at the kick-off event and during the workshops. You are welcome to upload alternative open data sources for your research and recommendations. Here are some datasets you can start working with:

- Geotab Hazardous Driving Area Dataset (<https://data.geotab.com/urban-infrastructure/hazardous-driving>)
- Geotab Road Impediments Dataset (<https://data.geotab.com/urban-infrastructure/road-impediments>)

- Geotab Intersection Metrics Dataset (<https://data.geotab.com/urban-infrastructure/intersection-metrics>)
- Toronto Police Services Killed or Seriously Injured Dataset (<http://data.torontopolice.on.ca/datasets/ksi>)
- Toronto Police Services All Motor Vehicle Collisions Reported (<https://data.torontopolice.on.ca/datasets/traffic-collisions-asr-t-tbl-001/explore?location=43.865743%2C-79.269698%2C9.72>)
- City of Toronto Address Points (<https://open.toronto.ca/dataset/address-points-municipal-toronto-one-address-repository/>)
- City of Toronto Centreline Intersection Dataset (<https://open.toronto.ca/dataset/toronto-centreline-tcl/>)
- [Accidentes de tránsito \(inegi.org.mx\)](https://inegi.org.mx)
- [Datos Abiertos de México - Infraestructura Carretera - Datos Viales anuales-2021](#)
- [Datos Abiertos de México - Tarifas Históricas y Vigentes](#)

Helpful Tips

Here are some sample questions that you may want to address:

- Which are the most accident/hazardous-prone locations?
- What factors affect accidents/hazards-prone locations (E.g.: proximity to certain buildings, roadway features, vehicle types, road impediments, intersection conditions, elevation)?
- Which accident types are most important to focus on? (E.g.: vehicle, speeding, visibility, time of day, injury age etc.)
- How do certain segments compare to the average for the ward, city, or surrounding areas?

General Tips:

- Perform exploratory analysis to find a particular problem to focus on. Most likely, no actionable insight can be extracted at the entire city-level.
- It is better to focus on one specific question and investigate it in some depth than to obtain limited insights and results on a wide range of topics.
- Use a combination of exploration, segmentation and modelling to come to a relevant conclusion.
- Make use of the geographic data to derive insights by visualizing it on a map.
- Ensure your presentation is non-technical but backed with solid data analysis.
- Keep in mind that your primary goal is not to build many analytical models. Rather it is to provide actionable model-backed insights and recommendations.
- There is no requirement to use all data available to you. Only the data relevant to your chosen analytical focus should be used.
- Keep data limitations in mind. Each dataset you are provided with is limited in some way. While these limitations do not make the data useless (all real-life data has some limitations), they are important to keep in mind when arriving at conclusions and recommendations.