Rotman MMA Datathon 2018: Data Dictionary

Toronto Police Services Killed or Seriously Injured

This dataset includes all traffic collisions events where a person was either Killed or Seriously Injured (KSI) from 2007 – 2017.

Source: http://data.torontopolice.on.ca/datasets/ksi
Size: 12557 rows x 58 columns
Column Descriptions:

Field Name	Description
Index_	Index number
ACCNUM	Accident Number
YEAR	Year Accident Occurred
DATE	Date Accident Occurred
TIME	Time Accident Occurred
HOUR	Hour Accident Occurred
STREET1	Street Accident Occurred
STREET2	Street Accident Occurred
OFFSET	Distance and direction of the accident
ROAD_CLASS	Road Classification
District	City District
LATITUDE	Latitude
LONGITUDE	Longitude
LOCCOORD	Location Coordinate
ACCLOC	Accident Location
TRAFFCTL	Traffic Control Type
VISIBILITY	Environment Condition
LIGHT	Light Condition
RDSFCOND	Road Surface Condition
ACCLASS	Classification of Accident
IMPACTYPE	Initial Impact Type
INVTYPE	Involvement Type
INVAGE	Age of Involved Party
INJURY	Severity of Injury

FATAL_NO	Sequential Number
INITDIR	Initial Direction of Travel
VEHTYPE	Type of Vehicle
MANOEUVER	Vehicle Manouever
DRIVACT	Apparent Driver Action
DRIVCOND	Driver Condition
PEDTYPE	Pedestrian Crash Type - detail
PEDACT	Pedestrian Action
PEDCOND	Condition of Pedestrian
CYCLISTYPE	Cyclist Crash Type - detail
CYCACT	Cyclist Action
CYCCOND	Cyclist Condition
PEDESTRIAN	Pedestrian Involved In Collision
CYCLIST	Cyclists Involved in Collision
AUTOMOBILE	Driver Involved in Collision
MOTORCYCLE	Motorcyclist Involved in Collision
TRUCK	Truck Driver Involved in Collision
TRSN_CITY_VEH	Transit or City Vehicle Involved in Collision
EMERG_VEH	Emergency Vehicle Involved in Collision
PASSENGER	Passenger Involved in Collision
SPEEDING	Speeding Related Collision
AG_DRIV	Aggressive and Distracted Driving Collision
REDLIGHT	Red Light Related Collision
ALCOHOL	Alcohol Related Collision
DISABILITY	Medical or Physical Disability Related Collision
Division	Police Division
Ward_Name	City Ward
Ward_ID	City Ward Identificator
Hood ID	Neighbourhood Identificator
Hood Name	Neighbourhood Name
FID	Alternate Index
Geohash	Geohash at the 7 character level (153m x 153m) that identifies the geohash associated with this hazardous driving area

Geotab Hazardous Driving Areas

This dataset identifies hazardous areas for driving according to harsh braking and accident level events collected within a specific area from Geotab's fleet of vehicles between June 2017 and June 2018 in the Toronto area.

Source: https://data.geotab.com/urban-infrastructure/hazardous-driving
Size: 1537 rows x 20 columns
Column Descriptions:

Field Name	Description
Geohash	Geohash at the 7 character level (153m x 153m) that identifies the geohash associated with this hazardous driving area
Latitude_SW	Latitude of the southwest corner of the geohash
Longitude_SW	Longitude of the southwest corner of the geohash
Latitude_NE	Latitude of the northeast corner of the geohash
Latitude_NE	Longitude of the northeast corner of the geohash
AvgLatitude	Weighted average latitude of points in the cluster where a significant volume of erratic driving events occurred
AvgLongitude	Weighted average longitude of points in the cluster where a significant volume of erratic driving events occurred
City	City (or municipality) within which the geohash resides (U.S., Canada, and Mexico only)
County	County within which the geohash resides (U.S. and Mexico only)
State	State within which the geohash resides (U.S., Canada, and Mexico only)
Country	Country (or territory) within which the geohash resides (English common name)
ISO_3166_2	ISO_3166_2 codes for country and subdivision
SeverityScore	Severity score for each area as the number of harsh braking incidents and accident-level incidents for every 100 units of traffic flow. Traffic flow is defined as total hourly vehicle volume in the geohash.
IncidentsTotal	The total number of harsh braking incidents and accident-level events that have occurred within the geohash
IncidentsCar	Number of incidents involving a car in the specified area
IncidentsMPV	Number of incidents involving a multi-passenger vehicle in the specified area.
IncidentsLDT	Number of incidents involving a light duty truck in the specified area.
IncidentsMDT	Number of incidents involving a medium duty truck in the specified area.
IncidentsHDT	Number of incidents involving a heavy duty truck in the specified area.
IncidentsOther	Number of incidents involving an unknown vehicle type in the specified area.

Geotab Road Impediments

The Road Impediments dataset uses the on-board accelerometer located within Geotab's telematics GO devices to help identify potential road impediments within the City of Toronto. Geotab refines this data to remove anomalies and outliers from the published dataset. Z-axis (up/down) accelerometer activity from the second quarter of 2018 is then aggregated to the 40 m x 20 m level.

Note: The geohash has been hashed at the 153mx153m level to match the other datasets. This means that there will be multiple entries for the same 7-character geohash.

Source: https://data.geotab.com/urban-infrastructure/road-impediments
Size: 461481 rows x 17 columns
Column Descriptions:

Field Name	**Descriptions**
Geohash	Geohash at the 7 character level (153m x 153m) that identifies where the road impediment is located
Latitude_SW	Latitude of the southwest corner of the geohash
Longitude_SW	Longitude of the southwest corner of the geohash
Latitude_NE	Latitude of the northeast corner of the geohash
Longitude_NE	Longitude of the northeast corner of the geohash
Latitude	Average Latitude of the road impediment
Longitude	Average Longitude of the road impediment
City	City within which the impediment resides (U.S. and Canada only)
County	County within which the impediment resides (U.S. only)
State	State within which the impediment resides (U.S. and Canada only)
Country	Country within which the impediment resides (English common name)
ISO_3166_2	ISO-3166-2 codes for country and subdivision
AvgAcceleration	The average magnitude of the upwards or downwards acceleration, in m/s2, experienced by vehicles travelling over the impediment at this location.
PercentOfVehicles	The percentage of vehicles travelling past the impediment that show a change in upwards or downwards acceleration.
AvgMonthlyVolume	The number of vehicles that travel over or past this impediment in the span of a month - averaged over the past 3 months.

PercentCar	Percentage of vehicles that were cars.
PercentMPV	Percentage of vehicles that were multi-purpose passenger vehicles.
PercentLDT	Percentage of vehicles that were light duty trucks.
PercentMDT	Percentage of vehicles that were medium duty trucks.
PercentHDT	Percentage of vehicles that were heavy duty trucks.
PercentOther	Percentage of vehicles that were unclassified.

City of Toronto One Address Repository

The One Address Repository (OAR) is the City of Toronto's authoritative addresses. Addresses in OAR are geographic objects with relationship to the City of Toronto's streets, parcels and buildings, comprised of information on street number, civic numbers, street type, direction and geographic location.

All address points have a position on the correct side of the street, and are correctly positioned relative to neighbouring addresses. They are positioned within the parcel or the building footprint.

Source: https://portal0.cf.opendata.inter.sandbox-toronto.ca/dataset/one-address-repository/
Size: 524666 rows x 12 columns
Column Descriptions:

Field Name	Descriptions
Geohash	Geohash at the 7 character level (153m x 153m) that identifies where the road impediment is located
MAINT_STAG	Maintenance Stage (e.g. reserve, regular)
ADDRESS	Address number with suffix
LFNAME	Street Name
DISTANCE	The distance from the start of the street segment along the street segment
FCODE_DES	Feature code description of the location
CLASS	Address classification (primary, structure, entrance address etc.)
NAME	Name of place (if exists e.g. churches, schools, parks etc.)
MUN_NAME	Municipality name (i.e. former municipality name)
WARD_NAME	Ward name
Latitude	Latitude of address
Longitude	Longitude of address

City of Toronto Centerline Intersection

The Toronto Centreline is a data set of linear features representing streets, walkways, rivers, railways, highways and administrative boundaries within the City of Toronto. Each line segment is described with a series of attributes including a unique identifier, name, feature code, and address ranges where applicable.

In addition to retaining historical archives, threaded archives are also retained that record splits and merges of address points, of linear features and of area features. All features are linked and integrated.

Source: https://portal0.cf.opendata.inter.sandbox-toronto.ca/dataset/centreline-intersection/
Size: 53494 rows x 9 columns
Column Descriptions:

Field Name	Descriptions
Geohash	Geohash at the 7 character level (153m x 153m) that identifies where the road impediment is located
INTERSEC5	Unique ID of intersection
CLASSIFI7	Intersection Classification
NUM_ELEV	Number of different elevation in this intersection
ELEVATIO10	Feature code description of the elevation
ELEV_LEVEL	Elevation Level Number
ELEVATION	Elevation height
Latitude	Latitude of segment
Longitude	Longitude of segment