

DESCRIPTION OF DATA

Data from four different data sources will be retrieved, cleaned, and merged using postal district as the common field among all the data.

Accidents Data will be retrieved from the [Toronto open data source portal](#). The data includes accidents data from 2006 till the end of 2019. The accident data is before the start of covid pandemic, so no impact of mandatory closing of businesses due to covid. The data includes the exact time when the accident happened, the location of accident, etc. Accident details are available for more than 16000 accidents over a period of 14 years.

Postal Districts Data with the longitude and latitude is available on the internet. The postal districts will be used to define regions in the Toronto city. The analysis will collect and summarize the data at the postal district level.

Toronto's Neighborhood and Borough's Data is retrieved from the [WIKIPEDIA page](#). This data will allow us to see results at the Borough's level.

FOURSQUARE has the data on all the night life venues organized under categories and subcategories. All the night life venues can be found under one main category using the API. The FOURSQUARE API will be used to search the businesses under the category 'Nightlife Spot' in each of Toronto's postal district.

Data from all the four data sources will be combined and summarized at the postal district level.

