

# Data Collection and Curation

We will collect our data from multiple sources:

- (a) [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- (b) [http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data)
- (c) <https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/transportation/#7c8e7c62-7630-8b0f-43ed-a2dfe24aad9>
- (d) Foursquare API

The first one contains the names of the postal codes areas in the Toronto area.

The second one contains the latitude and longitude of these areas.

The third one contains the traffic information of the streets in the central Toronto area.

The Foursquare API will give us all the results for the search “Restaurant” in each Postal code Area in the centre of Toronto within a 250m radius. We will eliminate duplicate restaurants and obtain the number of tips and their rating. For those restaurants without a rating we will assign a value of 5. Finally we will save the dataframes that we have curated for further analysis, visualization and Machine learning predictions.