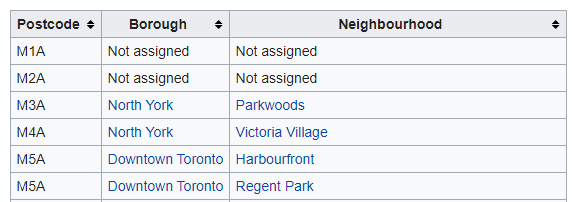
1. Data

Data requirements for the project are as listed below:

* Toronto’s location information: Toronto is chosen as the observation area. In particular, the schools within the boroughs of East Toronto, Central Toronto, Downtown Toronto and West Toronto have been considered for this project. Information on the boroughs, the neighborhoods and the postcodes were scraped from [Wikipedia](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M). A snippet of the data downloaded from Wikipedia is as below:



* Schools information and Test results: List of schools and the OSSLT results were downloaded from Ontario’s data catalogue, which is available as an excel spreadsheet. The file includes details of all schools within Ontario, including:
* board information
* school information
* grade 3 and 6 EQAO student achievements for reading, writing and mathematics
* grade 9 EQAO academic and applied student achievements
* grade 10 OSSLT student achievement
* student demographic percentages on student parents, special education, first language spoken and new students to Canada

School information includes the location information in longitude and latitude, postcodes, various other data pertaining to the school and students demographics. [Link](https://stage.data.ontario.ca/dataset/school-information-and-student-demographics).

* Venues in a particular location: To get the list of venues surrounding a school, **Foursquare** location data was utilized. Foursquare is a technology company that built a massive dataset of location data. They crowd-sourced their data and had people use their app to build their dataset and add venues and complete any missing information they had in their dataset. Currently its location data is the most comprehensive out there, and quite accurate that it powers location data for many popular services like Apple Maps, Uber, Snapchat, Twitter and many others, and is currently being used by over 100,000 developers, and this number is only growing.

The json file downloaded through the API provides the list of all venues present within a specified radius from the school. For each venue we get its name, unique ID, location, and category.