

## 2. Description of the data and how it will be used to solve the problem.

The general idea is to check all neighborhoods in Toronto and to find out if and how much they match one of the four categories, which scientifically we will call clusters here, mentioned above.

First, it is necessary to define what exactly is a neighborhood. Since Canada has a **postal code system** which is similar to the British one, already representing a high degree of localization, it is decided to call an area with one unique ZIP code a neighborhood for this task. This specific data for Toronto is, fortunately, already available on a Wikipedia page<sup>1</sup>.

In a next step, we need to find out the exact position of each of these neighborhoods, defined by **latitude and longitude**. There are several ways to get this data, for example the crowd sourcing project of Openstreetmap offers a so-called 'Overpass API'<sup>2</sup> that could be used for free, Geopy would be another option. For this example, we will simply import a \*.csv file<sup>3</sup> that already includes this data and add it to our neighborhoods list.



The screenshot shows the Wikipedia page titled 'List of postal codes of Canada: M'. It includes a table with the following data:

Postal Code	Borough	
M1A	Not assigned	
M2A	Not assigned	
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights

The data of existing **venues in each neighborhood** can be accessed in our Python code via the Foursquare Places API<sup>4</sup>. We will create a ranking which are the most common venues in these neighborhoods. We will import the Folium library and data provided from that **source to display maps** using Openstreemap of all specified clusters in the end.

Which kind of data will exactly match **our four clusters** mentioned above will be described more detailed in next week's task 2, bullet point (c) 'methodology section'.

<sup>1</sup> See [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

<sup>2</sup> See [https://wiki.openstreetmap.org/wiki/Overpass\\_API](https://wiki.openstreetmap.org/wiki/Overpass_API)

<sup>3</sup> See [https://cocl.us/Geospatial\\_data](https://cocl.us/Geospatial_data)

<sup>4</sup> See <https://developer.foursquare.com/places>