Survey of availability of medical healthcare facilities

Description of the data:

Since, the place is Toronto itself, I shall consider the almost the same approach for collecting the data set as it will serve the purpose of solving the business problem:

1. Build code to scrape the following Wikipedia page, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, and takeout relevant information of the table of postal codes and then to transform the data into a pandas dataframe.

2. The dataframe will consist of three columns: PostalCode, Borough, and Neighborhood as shown:

	PostalCode	Borough	Neighborhood
0	M5G	Downtown Toronto	Central Bay Street
1	M2H	North York	Hillcrest Village
2	M4B	East York	Parkview Hill, Woodbine Gardens
3	M1J	Scarborough	Scarborough Village
4	M4G	East York	Leaside
5	M4M	East Toronto	Studio District
6	M1R	Scarborough	Wexford, Maryvale
7	M9V	Etobicoke	South Steeles, Silverstone, Humbergate, Jamest...
8	M9L	North York	Humber Summit
9	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...
10	M1B	Scarborough	Malvern, Rouge
11	M5A	Downtown Toronto	Regent Park, Harbourfront

3. Geographical coordinates of each postal code can be found from http://cocl.us/Geospatial_data. From this dataset, a data frame of geospatial data can constructed that includes longitude and latitude information.

4. So, finally a merged dataframe would be constructed from the two datasets that will have two additional columns latitude and longitude . This latitude and longitude column information would be essential to for using Foursquare API to achieve the next steps.

Usage of the data constructed above:

Foursquare has a rich set of APIs that help find venues category wise. It is found that each category has subcategories.With different health care related venues against each neighborhood, the neighborhoods can be clustered/grouped.

For each place, latitude and longitude data can be feed to the foursquare api and expected venue can be filtered.