

Capstone Project Data

The Location: Scarborough is a popular destination for new immigrants in Canada to reside. As a result, it is one of the most diverse and multicultural areas in the Greater Toronto Area, being home to various religious groups and places of worship.

Longitude and Latitude Data: We will need geo-locational information about that specific borough and the neighborhoods in that borough. It is "Scarborough" in Toronto. This project will require knowledge of the different neighborhoods in Toronto, school ratings and median house prices. As such the neighborhood data required will be:

1. Neighborhood location in terms of latitude and longitude
2. School Ratings
3. Median House Prices

Dataset comprising latitude and longitude, zip codes is already available through the previous notebook. The location of Scarborough would be filtered using the same:

https://github.com/Caiyulu/SSH_Test/blob/master/W3%20Segmenting%20and%20Clustering%20Neighborhoods%20in%20Toronto%20PART%203.ipynb

	Postalcode	Borough	Neighborhood
0	M1B	Scarborough	Rouge, Malvern
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

Foursquare API Data:

We will need data about different venues in different neighborhoods of that specific borough. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API. After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meter.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category