Battle of Neighbourhoods

Prateek Shamprasad Pawar

Email: prateekspawar@gmail.com

1)Introduction:

1.1.Background:

Tom is a Resident of Parkwoods, Toronto. His family includes his wife, 10-year-old son and himself. His firm is offering him a promotion with great future prospects but he has to relocate to New York for this opportunity. Tom is inclined to accept this offer and his only problem is finding a similar great neighbourhood in New York as Parkwoods.

1.2. Quantification:

Tom decides to find out the most important factors that make Parkwoods a great neighbourhood for his family. If he can quantify the compatibility of a neighbourhood, it will make the job of finding a similar neighbourhood easy. After giving it considerable thought, he came up with the following factors making a good neighbourhood for them:

1) Plenty of Shops and Services making daily life easy.
2) A great number of food venues available.
3) Good transportation services.
4) Less Nightlife spots in the neighbourhood.

1.3. Problem:

This makes the problem clearer. The steps involved in problem-solving are:

1) Find all the venues in the neighbourhood of Parkwoods.
2) Categorize the venues in different categories. i.e. food, shops, transport etc.
3) Assign proper weight to each category according to the factors specified above.
4) Find the total score of the neighbourhood using the weights and number of venues in each category.
5) Find the total score of each neighbourhood in New York using same weighs.
6) Separate the neighbourhoods having total score more than Parkwoods, same as Parkwoods, less than Parkwoods.
7) Provide the separated list of neighbourhoods to Tom to help him decide which one to select for relocation.

Data:

For finding out all the venues in neighbourhood, we will use the Foursquare API. To call in the Foursquare database, we need Latitude and Longitude values of each of the neighbourhood.

a) Neighbourhoods in New York:

1)The geospatial data of all the neighbourhoods in New York is available in .json file at https://geo.nyu.edu/catalog/nyu_2451_34572 .
2)We can extract data about Neighbourhood name, Latitude, Longitude from the file using json.load() function from json library.
3) We can append this data in a Pandas DataFrame having columns Neighborhood, Latitude and Longitude.

```
In [7]: ny.head()
```

Out[7]:

	Neighborhood	Latitude	Longitude
0	Wakefield	40.894705	-73.847201
1	Co-op City	40.874294	-73.829939
2	Eastchester	40.887556	-73.827806
3	Fieldston	40.895437	-73.905643
4	Riverdale	40.890834	-73.912585

4) Then using the Foursquare API request of venues, we can find all the venues in each neighbourhood of New York.

b) Neighbourhoods in Toronto:

1) The table having name of each neighbourhood in Toronto and its postal code is at https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M .
2) The csv file having Latitude and Longitudes for each postal code in Toronto is at https://cocl.us/Geospatial_data/Geospatial_Coordinates.csv .
3) We can combine these two datasets to make one DataFrame having names of Neighbourhoods and their Latitudes and Longitudes.

```
In [10]: tor.head()
```

Out[10]:

	Neighborhood	Latitude	Longitude
0	Parkwoods	43.753259	-79.329656
1	Victoria Village	43.725882	-79.315572
2	Harbourfront	43.654260	-79.360636
3	Lawrence Heights,Lawrence Manor	43.718518	-79.464763
4	Queen's Park	43.662301	-79.389494

4) Then using the Foursquare API request of venues, we can find all the venues in each neighbourhood of Toronto for further analysis.