Choose a Most Suitable Neighborhood When Move to an Unfamiliar City

Lulei Ni

1. Introduction

Moving to another city is a situation many people face in their life. But if you are not familiar with the new city, pick which neighborhood to live would be a big problem. Ask friends? May be someone's best neighborhood is not suitable for you. May be your friend like to Chinese food so a Chinese restaurant nearby is very important to him but not to you. So, here is the question, how to choose the neighborhood that is most suitable for you.

Here is the example, suppose Mary decided to move to Toronto, but she is not familiar with Toronto. At the suitation, she doesn't know the neibothood is suitable for her or not. So she makes a list which including the important nearby venues categories to her to help making the decision. For example, there should be a shopping mall near the neighborhood. Since she has 2 children, a park near by neighborhood would be good for her children to play at the weekend or holiday. She likes drinking coffee, she need to do Yoga 3 times a week and her huaband goes to gym 2 times a week. They like to visit Museum or watching moive, they like to ...

So Mary made a list to give priority to the categories of the venuse, now she needs the geographical data with venues information of the Toronto.

2. Data acquisition

The data has three part:

Part 1: Toronto's neighborhood information, from the Wiki:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

	Postal Code	Borough	Neighborhood
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront
5	M6A	North York	Lawrence Manor, Lawrence Heights
6	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

Part 2: The geographical coordinates of the neigborhood:
http://cocl.us/Geospatial_data

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Part 3: Venues information
I use the **Forsquare API** to get the most common venues of given Borough of Toronto

After getting all the data, merge them together:

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
Regent Park, Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa
Regent Park, Harbourfront	43.65426	-79.360636	Dominion Pub and Kitchen	43.656919	-79.358967	Pub

Now we got the all data needed to help Mary make the decision.

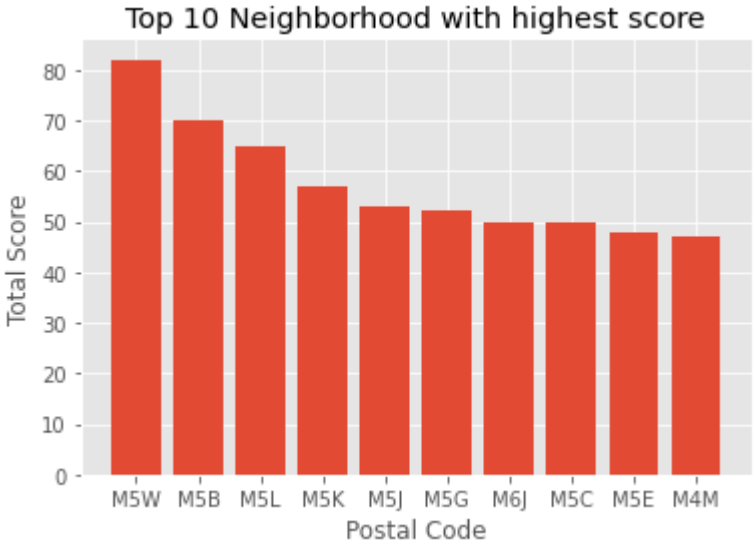
3. Methodology

First we need to make a scorecard to Mary's list to give priority to different categories of venues:

Venue Category	Scores
Shopping Mall	15
Park	13
Coffee	10
Yoga	7
Gym	7
Italian Restaurant	5
American Restaurant	5
Japanese Restaurant	5
Movie	5
Theater	5
Museum	5

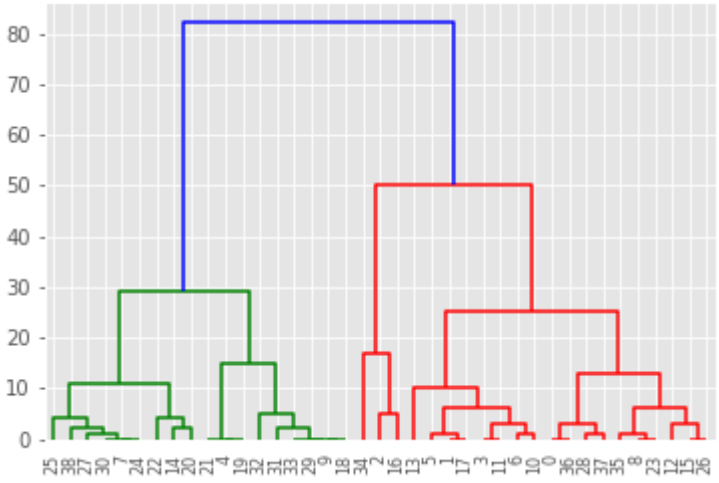
According to the scorecard, when nearby the neighborhood there is a shopping mall, the neighborhood got 15 scores. As well as if there is a park, additional 13 socres, and so on.

Second let's combine the scorecard with the Toronto's neighborhood, making the scores to every neighborhood, at last every neighborhood got its total score. We can rank the neighborhood based on the score so Mary can choose the most suitable neighborhood from the high score ones.



Third, we can use the **cluster** method based on the scores to make different classes of the neighborhood.

Based on the hierachical clustering chart:



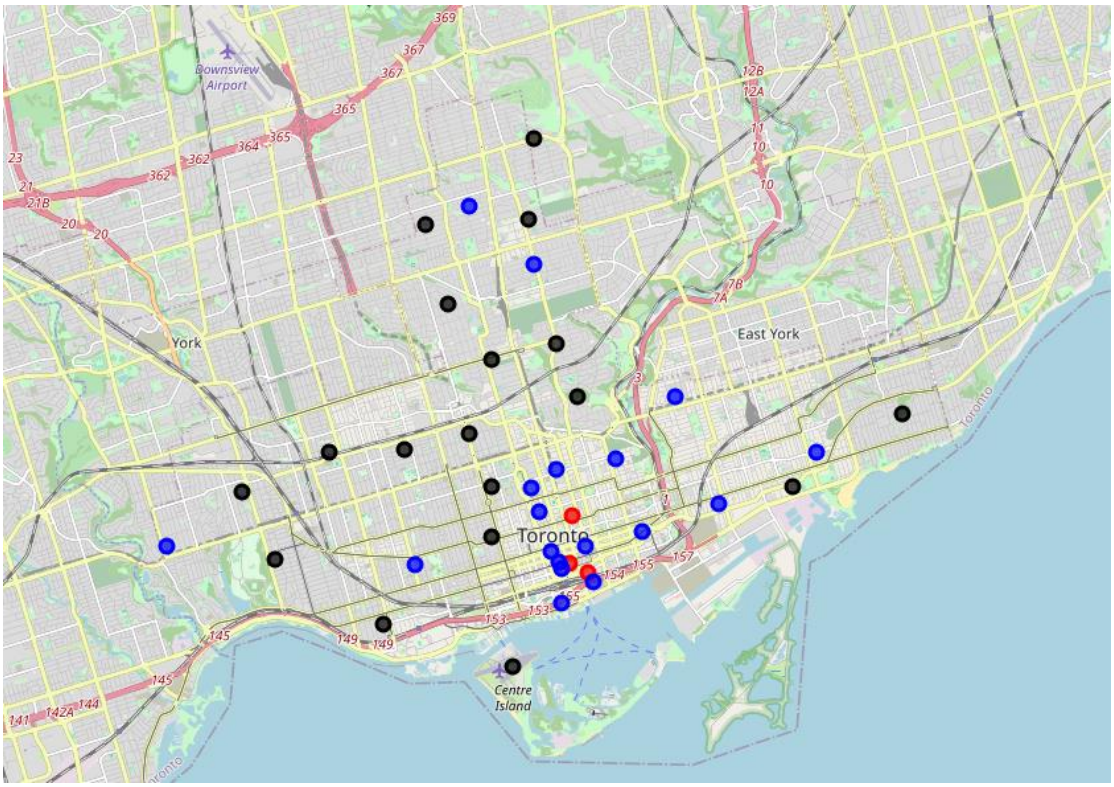
We cluster the neighborhood to 3 classed, and the highest scores class is the Mary's first choice.

Postal Code	Borough	Neighborhood	Latitude	Longitude	Total Score	labels
M5W	Downtown Toronto	Stn A PO Boxes	43.646435	-79.374846	82	2
M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	70	2
M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	65	2

4. Result

So, ***Stn A PO Boxes, Garden District, Ryerson, Commerce Court, Victoria Hotel*** should be Mary's first choices.

We can also visualize all the cluster information on the map to make it intuitive:



The red points are Mary's first choices and the black points are Mary's last choice.

5. Discussion

Now we finish all the steps and help Mary choose the most suitable neighborhood. When you decide to move to another city which you are not family with, you can also obtain the data from the internet, use **Foursquare API** to get the venues information, make the scorecard and take scores to all the neighborhoods in the new city. After doing that, you will got your first choice. Also house price and quality is your consideration too, you can take your field trips among the first choices to make the last decision.

Also this is a very simple model to help some to make decisions, you can add house price as another variable to upgrade the model to make it more reasonable.

Thanks for your reading.