

Toronto Neighborhood Analysis for Chinese Parent Immigrants

IBM DATA SCIENCE PROFESSIONAL CERTIFICATE CAPSTONE

WANG JINGWEN

Table of Contents

<i>Introduction to Business Problem</i>	<i>2</i>
<i>Data Scours and Data Preparation</i>	<i>2</i>
<i>Exploratory Results.....</i>	<i>2</i>
Toronto Ethics Group.....	2
Toronto Schooling Distribution – North York vs. Scarborough.....	3
Chinese Restaurant by Regions.....	5
<i>Cluster Analysis – Narrow Down the Neighborhood.....</i>	<i>6</i>
<i>Result & Discussion</i>	<i>7</i>
<i>Conclusion & Further Improvement</i>	<i>7</i>

Introduction to Business Problem

Moving to a new country and settling down in a new city has never been an easy choice. Canada, as one of the most immigration-friendly countries, has always welcomed people from all over the world to find new homes in Canada. According to *2016 Census: Housing, Immigration and Ethnocultural Diversity, Aboriginal Peoples*, in 2016, 47% of the residents of Toronto were immigrants. The main challenge faced by new immigrants is to adapt to the new cultural environment, including languages, food cultures and lifestyle. Therefore, by choosing a community where people have the same or similar cultural backgrounds, new immigrants can adapt to the new environment in a more comfortable way. This project further narrows down the target audience to be Chinese parents, who have less flexibility in moving as a family and in greater needs to help their children to adapt to the new place.

The purpose of this project is to help newly landed Chinese parents to better understand the neighborhoods and regions of Toronto. The selection criteria include ethics, schooling and Chinese restaurants.

Data Scours and Data Preparation

Three main sources are used for this project:

- Foursquare API: used to retrieve location data, including nearby venues.
- Wikipedia:
 - “List of Postal Code of Canada: M”: obtained neighborhood and region information of Toronto, including postal code, borough, and neighborhood. (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
 - “Demographics of Toronto”: obtained ethics density data over regions of Toronto. (https://en.m.wikipedia.org/wiki/Demographics_of_Toronto#Ethnic_diversity)
- Open Data by City of Toronto: used to obtain school information of Toronto area.

Major Data Cleaning: Toronto Location Information Table from Wiki - Only process the cells that have an assigned borough. Ignore cells with a borough that is Not assigned. Separated multiple neighborhoods with comma in each cell.

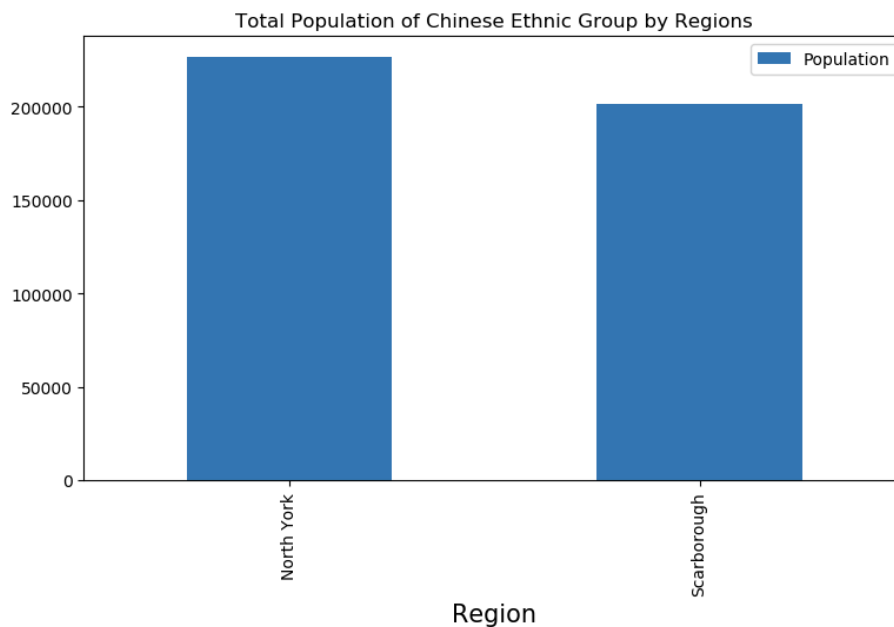
Exploratory Results

Toronto Ethics Group

By reviewing the ethics density in different regions of Toronto. Top 3 Ethics groups are selected with recording the number of populations. North York and Scarborough are two regions containing ridings where Chinese population is the top 1 ethic group. By comparing the total population in

North York and Scarborough regions, North York has slightly higher total population than Scarborough region, but the level is very comparable. Therefore, Chinese immigrants are more likely to find people from Chinese background in these two areas comparing to other regions in Toronto.

There are many benefits to start a new life at where new immigrants can find people originally from the same country. First of all, new immigrants can communicate easier with people having no language barriers. Landing into a new environment, new Chinese immigrants may find it easier to ask around from their neighbors. To facilitating the communication and avoid misunderstanding, it can be helpful if they can speak in Chinese with their neighbor when it is necessary. Building new relationships and expanding networks are always at the top challenge for new immigrants, and therefore, choosing to live at where they can easily find other families sharing similar background can help new immigrants to settle down quicker and more smoothly.

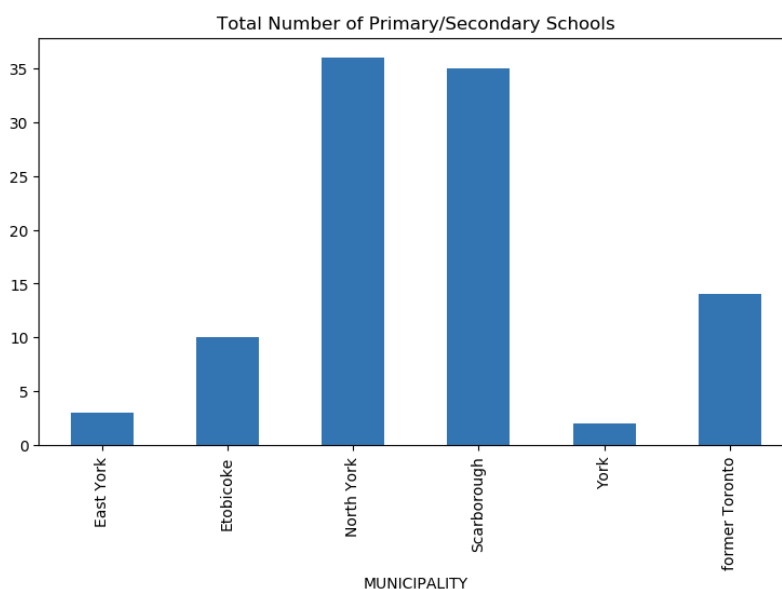


Toronto Schooling Distribution – North York vs. Scarborough

Comparing to single and young immigrants, parents have less flexibility in moving as a family. More importantly, children, especially young children, are more sensitive to the changing environment. Newly immigrant parents pay very much attention to how well their children can adapt to the new environment, and they need to ensure that their movement aligns with the educational goal as well. Unlike post-secondary educational institutions, children attending primary or secondary schools are mainly dependent to the location. Parents are more likely to send their kids to nearby schools in order to take on advantages such as scheduled school bus and school

activities arranged by the community. In addition, it is a safer choice comparing to send their children to somewhere far from home.

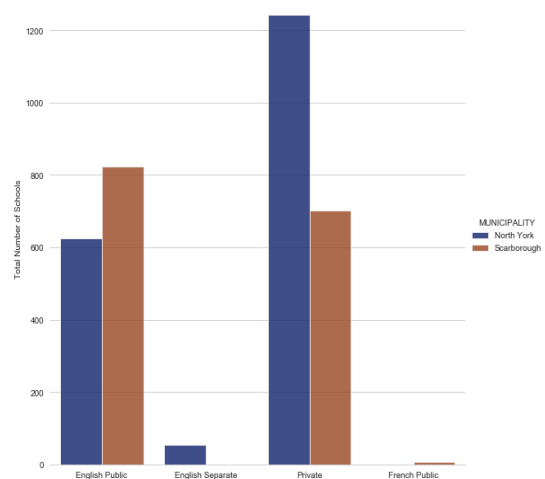
The data source file is directly imported using API from Open Data from City of Toronto website. A geographical file covering the City of Toronto that contains all the point location of schools that we have. It also contains the private schools to the best of our knowledge.



Statistics on total number of primary/secondary schools of Toronto suggests North York and Scarborough regions having more than twice number of schools comparing to the rest of regions in Toronto. This observation is a very good news to Chinese parents because according to the ethics group analysis, these two regions are the top choices on the list.

Out[38]:

	MUNICIPALITY	School Type	Total Number of School
0	North York	English Public	625
1	North York	English Separate	55
2	North York	Private	1242
3	Scarborough	English Public	823
4	Scarborough	French Public	8
5	Scarborough	Private	702

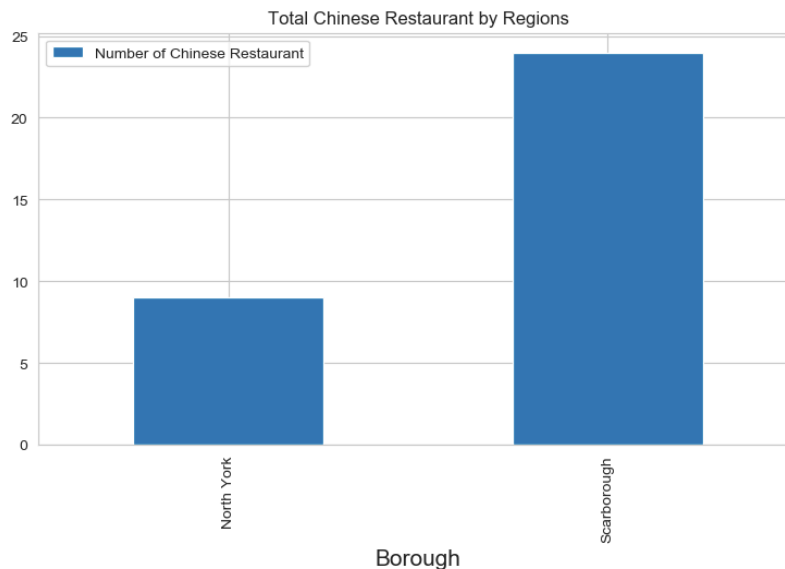


(-- North York --Scarborough)

Noted that English Public and Private schools are the major types in North York and Scarborough regions, which also fits the need of Chinese parents. Scarborough region has slightly higher number in public English school, while North York region has almost twice more private schools than Scarborough region.

Chinese Restaurant by Regions

Chinese cuisine is a very important part of Chinese culture, and it is very differentiable from other cuisines. Moreover, eating habit is one of the most resistant habit that cannot be easily changed over time. Given that North York and Scarborough regions having greater percentage in Chinese ethnic group, it is likely that new immigrants can find more Chinese restaurants there.



Scarborough region has almost twice more Chinese restaurants than North York region. As the final recommendation needs to be as specific as possible, number of nearby Chinese restaurants are also studied at neighborhood level.

Out[50]:

	Borough	Neighborhood	Number of Chinese Restaurant
0	North York	Bathurst Manor, Wilson Heights, Downsview North	1
1	North York	Bayview Village	1
2	North York	Don Mills	2
3	North York	Hillcrest Village	1
4	North York	North Park, Maple Leaf Park, Upwood Park	1
5	North York	Northwood Park, York University	1
6	North York	Parkwoods	1
7	North York	York Mills West	1
8	Scarborough	Agincourt	4
9	Scarborough	Cedarbrae	1
10	Scarborough	Clarks Corners, Tam O'Shanter, Sullivan	1
11	Scarborough	Cliffside, Cliffcrest, Scarborough Village West	1
12	Scarborough	Dorset Park, Wexford Heights, Scarborough Town...	2
13	Scarborough	Kennedy Park, Ionview, East Birchmount Park	4
14	Scarborough	Malvern, Rouge	1
15	Scarborough	Milliken, Agincourt North, Steeles East, L'Amo...	5
16	Scarborough	Steeles West, L'Amoreaux West	4
17	Scarborough	Woburn	1

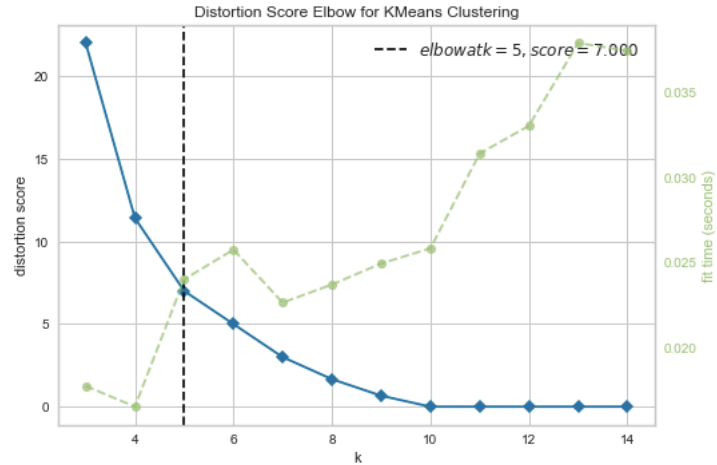
According to the statistics on Chinese restaurant on neighborhood level, it can be seen that all neighborhood has at least one Chinese restaurant nearby. The neighborhood (or neighborhood region) has the highest number of Chinese restaurants surrounding are those near Milliken.

Cluster Analysis – Narrow Down the Neighborhood

The purpose of this project is to help new immigrant Chinese parents to understand the neighborhood and regions based on their particular needs and provide potential recommendation at neighborhood level.

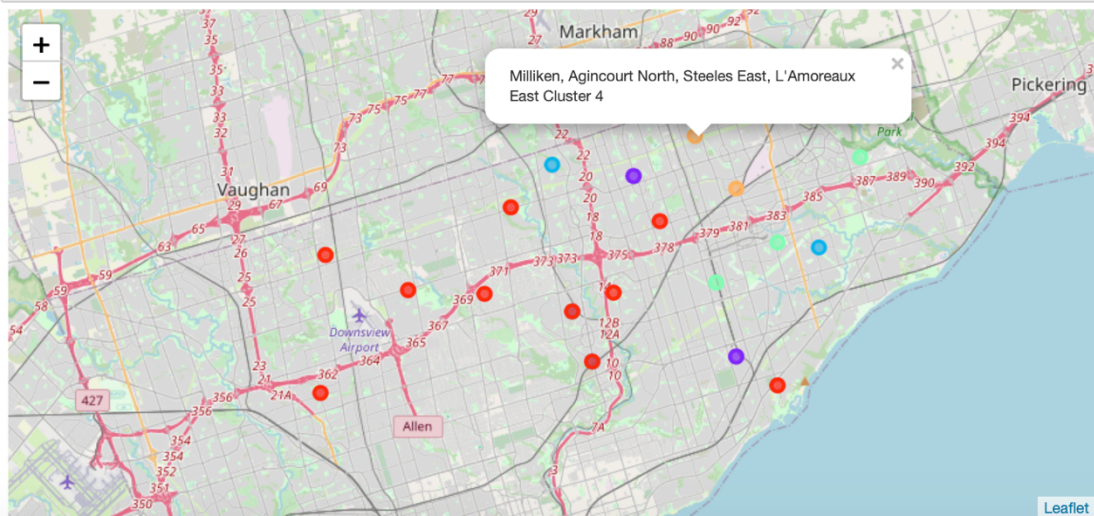
K-means is an unsupervised clustering algorithm designed to partition unlabeled data into a certain number (the “K”) of distinct groupings. In other words, k-means finds observations that share important characteristics and classifies them together into clusters (Alan, Jeffares). By utilizing K-mean cluster analysis, neighborhood with the greatest similarity according to schooling and Chinese restaurants can be grouped together. Therefore, audience can easily find similar neighborhoods fits their need.

To determine the number of groups (the value of K), the elbow method is used, and group number is 5. Package *KElowVisualizer* from *yellowbrick* is used to determine the best value of K.



Result & Discussion

Out[72]:



The cluster analysis grouped neighborhoods in North York and Scarborough regions into 5 groups. The cluster analysis groups neighborhood into less restaurants & less school, more schools & less restaurants, less schools & more restaurants, and more school and more restaurants. Agincourt and Milliken (neighborhood regions) have relatively more choices of Chinese restaurants and more selection of schools, which can be recommended to Chinese parent to settle their first new home.

Conclusion & Further Improvement

As discussed in the results, North York and Scarborough are two recommended regions for new Chinese parent immigrants. Further studies can be done by including more selection criteria such as grocery, shopping mall, medical center and transportations.