Clustering strategy for finding potential optimal location of a Chinese restaurant in Toronto

Yilun Zhang

2021.05.21

1. Introduction

1.1 Background
Canadians who identify themselves as being of Chinese ethnic origin make up about 4.6% of the Canadian population, or about 1.57 million people according to the 2016 census. In addition, 11.1% of total population in Toronto are Chinese.

1.2 Business Problem
Under this background, opening a Chinese restaurant in Toronto seems to be a great business opportunity for our clients/stakeholders. However, according to my personal experience, not all Chinese restaurant can have a big success or live long in Toronto. Although there are many factors accounting for this situation, I think a cautious decision on restaurant location is the first thing needs to be considered for our clients/stakeholders. In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening a Chinese restaurant in Toronto. By intuition, population, the number and types of existing Chinese restaurants should be taken into consideration when deciding a good location candidate.

2. Data
Based on the population of a neighborhood, the number of Chinese restaurants in a neighborhood and the types of Chinese restaurants in a neighborhood, we can use clustering strategy to segment these neighborhoods in Toronto and decide which neighborhoods our clients/stakeholders should choose when they want to start up a new Chinese restaurant in Toronto.

To be more specific, we will use web-crawling technique to get all neighborhoods in Toronto with the corresponding postal codes and boroughs. See example below:

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

Then we will use the GeoSpatial Dataset given by Coursera to get the latitude and longitude coordinates of each neighborhood. See example below:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494

We will also use the census data to get the population information of each neighborhood downloaded from Canada open data portal. See example below:

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Population
0	M3A	North York	Parkwoods	43.753259	-79.329656	34615
1	M4A	North York	Victoria Village	43.725882	-79.315572	14443
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	41078
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	21048
4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494	10

Lastly, we will utilize Foursquare API to get the venue information of each neighborhood, or in other words, Chinese restaurant related information. See example below:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.753259	-79.329656	The Bean Sprout	43.742229	-79.313577	Chinese Restaurant
1	Parkwoods	43.753259	-79.329656	Omni Palace Noodle House	43.771047	-79.331570	Chinese Restaurant
2	Parkwoods	43.753259	-79.329656	Dragon Pearl Buffet 龍珠	43.753693	-79.349730	Chinese Restaurant
3	Parkwoods	43.753259	-79.329656	China Gourmet	43.755189	-79.348382	Chinese Restaurant
4	Parkwoods	43.753259	-79.329656	Spicy Chicken House	43.760639	-79.325671	Chinese Restaurant
...	...	...	...	...	...	...	...
2044	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999	Ancom Chinese Restaurant	43.624481	-79.509448	Chinese Restaurant
2045	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999	Mandarin Buffet	43.621352	-79.523015	Chinese Restaurant
2046	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999	Li's Oriental Kitchen	43.637787	-79.539134	Chinese Restaurant
2047	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999	Lemongrass	43.645010	-79.522379	Thai Restaurant
2048	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999	Chinese Food Gallery	43.643905	-79.533167	Chinese Restaurant