Description of business problem and background

There is always an opening of new business. Before opening any new business, the market opportunity analysis for that kind of business is necessary. This is the first step to ensure business success. A restaurant business plans to open a new restaurant in Toronto. Which kind of restaurant and where to open the restaurant are very important factors for the business's success. I use the neighborhood listings and the FourSquare venue listings from prior projects to obtain a picture of which and how many restaurants of different types can be found in the different neighborhoods of Toronto. We then compare the frequencies of different types of restaurants across neighborhoods to find out where there may be fewer restaurants than the market can support, which we can recommend as opportunities to open new restaurants of different types.

Data Description

2.1 Data Sources

Toronto neighborhood data and geospatial data (latitude and longitude) can be found in the website https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M and http://cocl.us/Geospatial_data . Venue listings are queried from FourSquare.

2.2 Data Preprocessing

First, the data from the above sources are merged on postal code to obtain the latitude and longitude for each neighborhood and then the Toronto neighborhood data was extracted.

Here is the snapshot of Toronto neighborhood data

	Postal Code	Borough	Neighborhood	Latitude	Longitude
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
15	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
19	M4E	East Toronto	The Beaches	43.676357	-79.293031
20	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
24	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
25	M6G	Downtown Toronto	Christie	43.669542	-79.422564
30	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568
31	M6H	West Toronto	Dufferin, Dovercourt Village	43.669005	-79.442259
36	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752

Second, the neighborhood venue listings were queried from FourSquare and the Toronto neighborhood venue listings were extracted. Here is some example data queried from foursquare API and merged with Toronto neighborhood data.

	Postal Code	Borough	Neighborhood	Latitude	Longitude	headerLocation	venue.location.lat	venue.location.lng	venue.id	venue.name	category	ne.lat	
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	Corktown	43.653447	-79.362017	54ea41ad498e9a11e9e13308	Roselle Desserts	Bakery	43.658760	-79.
1	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	Corktown	43.653559	-79.361809	53b8466a498e83df908c3f21	Tandem Coffee	Coffee Shop	43.658760	-79.
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	Corktown	43.653947	-79.361149	4ae5b91ff964a520a6a121e3	Morning Glory Cafe	Breakfast	43.658760	-79.
3	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	Corktown	43.653249	-79.358008	574c229e498ebb5c6b257902	Cooper Koo Family YMCA	Distributor	43.658760	-79.
4	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	Corktown	43.654735	-79.359874	50760559e4b0e8c7babe2497	Body Blitz Spa East	Spa	43.658760	-79.
5	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	Corktown	43.656369	-79.356980	5612b1cc498e3dd742af0dc8	Impact Kitchen	Restaurant	43.658760	-79.

After checking FourSquare's neighborhood and postal code neighborhood, I found that FourSquare's neighborhoods are not the same as the postal code neighborhood boundaries, so some venues will show up more than once. I assign each venue to the postal code neighborhood it is closest to.

Third, I checked how many different kinds of venues we have in each neighborhood. Some of these are not restaurants, so I drop them. In addition, since this list is long, I drop any types that have 5 or fewer venues; it may not be feasible to open another one of these restaurants because they are not very popular. Now the data is ready to do analysis.

	Postal Code	Borough	Neighborhood	Latitude	Longitude	headerLocation	venue.location.lat	venue.location.lng	venue.id	venue.name	category	ne.lat	n
0	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Bay Street Corridor	43.660391	-79.387664	4a8355bff964a520d3fa1fe3	Mercatto	Italian	43.666802	-79.38
1	M5W	Downtown Toronto	Stn A PO Boxes	43.646435	-79.374846	St. Lawrence	43.646964	-79.374403	4ada4b1ef964a520022121e3	The Old Spaghetti Factory	Italian	43.650935	-79.36
2	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280	Financial District	43.650243	-79.380820	4adc5c6af964a520da2b21e3	Mercatto	Italian	43.652929	-79.37
3	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188	Greektown	43.677062	-79.353934	4af4e0d0f964a5202ff721e3	7 Numbers	Italian	43.684057	-79.34