Analysis of the Business Opportunities in Toronto

IBM Capstone Final Submission

Jeff Bowyer

6 June 2020

Table of Contents

Business Problem Introduction

Encouraging investment in a city is essential for the health and growth of the area. How do investors find opportunities to invest in and then have the confidence to move forward with the business opportunity?
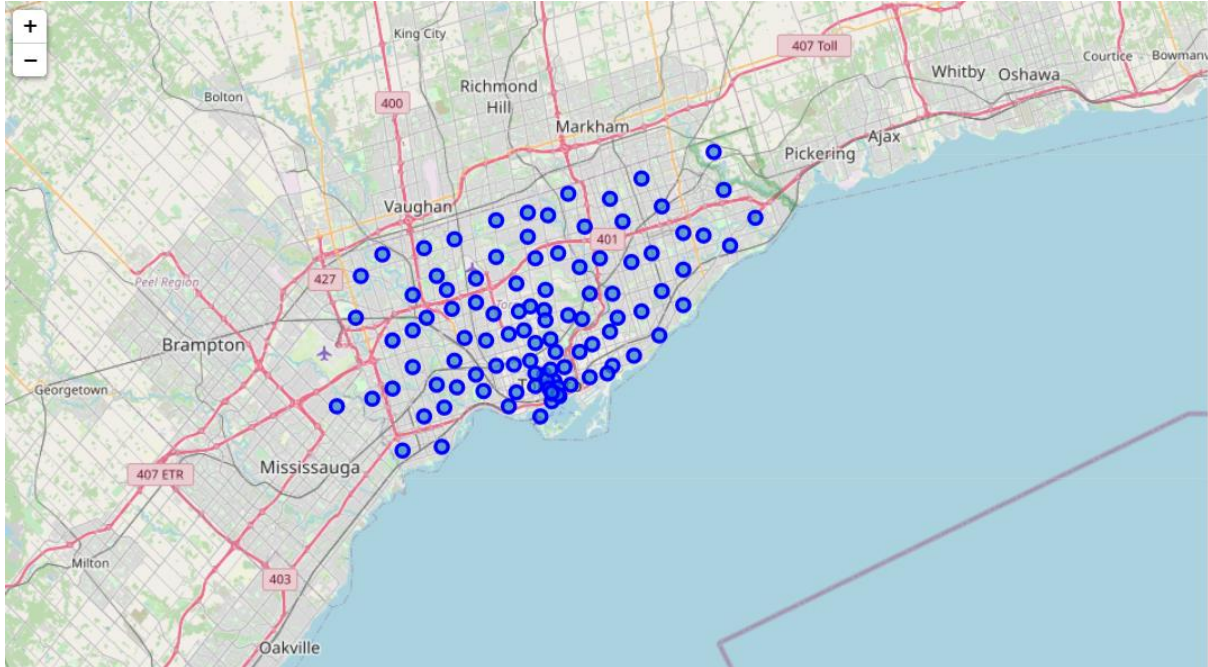

Figure 1 The Neighborhoods in Toronto, Canada

With a large city like Toronto, it would be hard for an investor to know where to start looking. This is especially true for investors that do not know a city well. For example, a restaurateur, would want to know which areas of the city are popular for her product and how many other competitors are active in the area. Without this information it would be difficult for that investor to move forward with the initiative.

Is there a way to provide this information to investors, in a clear and convenient way that will allow them to narrow down their search quickly?

The intention of my final project is to provide valuable data analysis that can be used by entrepreneurs to identify business opportunities in Toronto for investment.

Specific questions to be answered:

- Where are the popular areas for service businesses in Toronto?
- What are the most popular service businesses in Toronto?
- Which popular areas are under served with specific types of businesses?
- What are the top service business opportunities in Toronto right now?

Data

Data Sources

The first data source will be the Wikipedia page that holds the Neighborhoods within Toronto.

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

This data contains 3 columns: Postal Code, Borough and Neighborhood. An example of the data is: M3A, North York, Parkwoods.

This will be combined with a list of coordinates for each of these Neighborhoods. The postal codes of each Neighborhood will be used to connect the two data sources.

https://cocl.us/Geospatial_data

This is a CSV file that has 3 columns: Postal Code, Latitude and Longitude. An example of the data is: M1B, 43.8066863, -79.1943534.

From the resultant table holding Neighborhoods and Coordinates we will use the Foursquare API to pull all the venues within a 500 meter radius of each Neighborhood.

https://foursquare.com/

This data is returned in JSON format with the specific fields requested. The intention will be to pull nearby service businesses and venues. An example of the data would be: Park, Pizza Place etc.

From this dataset I will then do the analysis outlined under the Business Problem Section.

Data Wrangling

Link to Jupyter Notebook in Github
https://github.com/jeffman1000/Coursera_Capstone/blob/master/IBM%20-%20Capstone%20Final%20Submission.ipynb

Overall, the a Jupyter Notebook using Python was used to wrangle the data along with the Pandas library.

The Toronto neighborhoods and postal codes were screen scraped from the web page using the BeautifulSoup library.
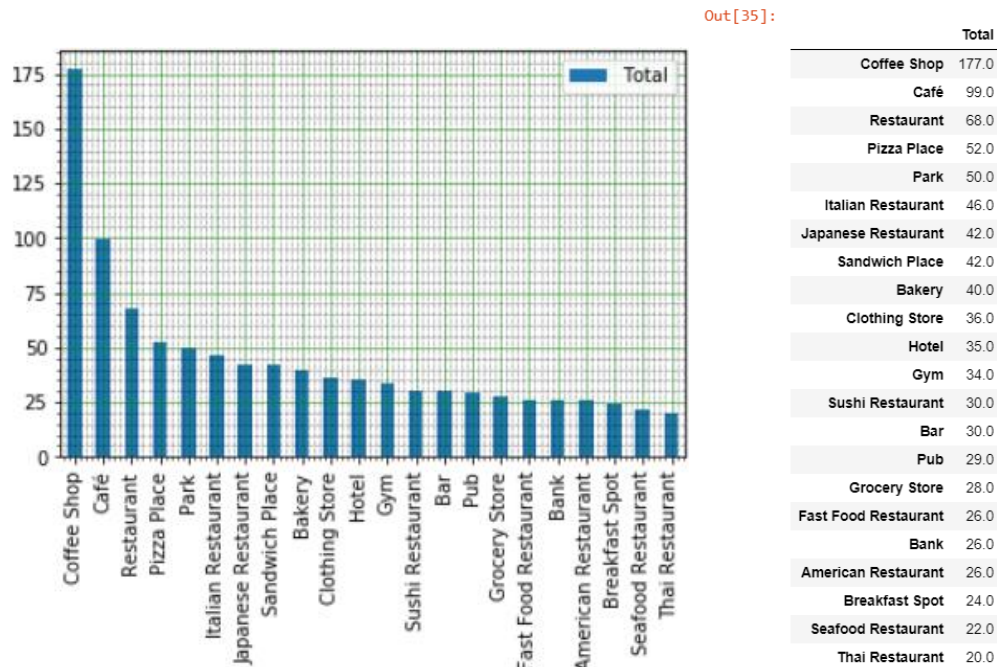
This was linked to the .csv file with the longitude and latitude for each postal code area.

Onehot encoding was used to create a simple table with all the detail of the venues. This was then sliced, as required, to get the data required.

Results

What are the popular areas for Service business in Toronto?

By doing a count of each type of business in Toronto, we can then rank the businesses by popularity.

Out[35]:

	Total
Coffee Shop	177.0
Café	99.0
Restaurant	68.0
Pizza Place	52.0
Park	50.0
Italian Restaurant	46.0
Japanese Restaurant	42.0
Sandwich Place	42.0
Bakery	40.0
Clothing Store	36.0
Hotel	35.0
Gym	34.0
Sushi Restaurant	30.0
Bar	30.0
Pub	29.0
Grocery Store	28.0
Fast Food Restaurant	26.0
Bank	26.0
American Restaurant	26.0
Breakfast Spot	24.0
Seafood Restaurant	22.0
Thai Restaurant	20.0

Coffee Shops are, by far the most popular type of business in Toronto. This would suggest that they are popular venue and locals are accustomed to making use of their services.

Following on are three types of eateries including Café's, Restaurant's, and Pizza places.

Most venues are restaurant type and confirms demand for these types of establishments. 16 of the 22 venue types are restaurants and 80% of total venues are restuarants.

What Toronto Areas are Popular for Business?

By mapping the areas with the most businesses we can see which parts of the city are most popular. We know that the most popular businesses are restaurants, therefore these areas are popular places that people go to eat.
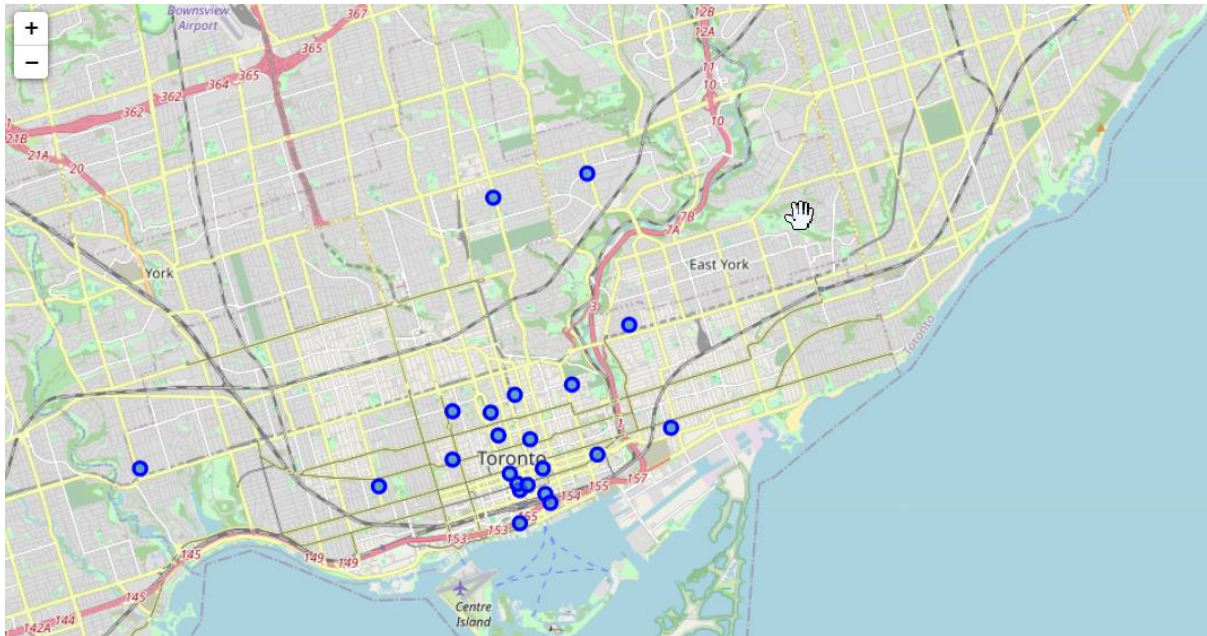
The map shows that "downtown" is a popular area. However, there are also some popular hotspots to the north, east and west of the city. It would make sense to focus a new business venture close to one of these areas.

Knowing the popular areas is important as it shows where large volumes of customers can be found. High footfall is an advantage with any new business.

Which popular areas are under served with specific types of businesses?

A simple algorithm was created to identify underserved areas. This looked at the occurrence of specific businesses in an area in comparison to the average number of that business type in popular areas. Then that number was corrected for by the overall popularity of the area.

Served Level = No. of businesses in Area/ Ave no. of that business in popular areas x No. of businesses in Area/Ave no. of businesses in popular areas

The table below highlights underserved business types in specific areas. A rating of 1 is the perfect level based on the algorithm. Therefore, any value under 1 is a potential area for investment.

	Neighborhood	Coffee Shop	Café	Restaurant	Hotel	Japanese Restaurant	Clothing Store	Italian Restaurant	Seafood Restaurant	Bakery	Gym	American Restaurant	Beer Bar
0	Toronto Dominion Centre, Design Exchange	1.46808	1.69764	1.63807	3.01194	1.50597	0.555775	2.12205	2.33425	1.63807	0.864539	2.59362	1.94521
1	First Canadian Place, Underground city	1.46808	2.26352	1.63807	2.00796	2.00796	0	0.70735	2.33425	0.819037	3.45815	2.59362	1.94521
2	Commerce Court, Victoria Hotel	1.61489	1.98058	2.86663	2.50995	1.50597	0	2.12205	2.33425	1.63807	3.45815	3.45815	1.94521
3	Garden District, Ryerson	1.17447	0.84882	0.409518	1.00398	1.50597	5.00197	2.12205	0.778085	1.63807	0.864539	0	0.972606
4	Harbourfront East, Union Station, Toronto Islands	1.88942	1.12044	1.21627	1.98788	0.49697	0	2.10083	0.770304	0.810846	0.855893	0	0.96288
5	Stn A PO Boxes	1.22879	1.05254	1.14256	0.933702	1.40055	0.516871	1.97351	2.17086	1.52341	1.60804	0.804021	2.71357
6	Richmond, Adelaide, King	1.35064	1.30152	1.50703	1.38549	0.461831	1.53394	0	0.715838	1.50703	2.38613	1.59075	0
7	St. James Town	0.579893	1.11761	0.970558	0.396572	0.396572	0.878124	1.11761	1.22937	0.647039	1.36597	2.04896	1.53672
8	Church and Wellesley	0.695872	0.447045	0.970558	0.793144	1.98286	0.439062	0	0	0	0	0.682985	0.768359
9	Fairview, Henry Farm, Oriole	0.491808	0	0.823132	0	1.009	3.72369	0	0	0.548754	0	0.579241	0
10	Central Bay Street	1.01738	0.713009	0.257997	0.316254	0.632508	0	1.78252	0.490193	0	0	0	0

Fairview Opportunity

Fairview (9th most popular area) is in the north east of the city. It is richly served with clothing stores but underserved in coffee shops and eateries. This suggests that there is an opportunity to invest here. Japanese restaurants are overserved, but all other restaurant types are underserved.

St James Town Opportunity

This is the 7th most popular area and is well served with Cafes, but is underserved with Coffee Shops. Coffee Shops are the most popular business in Toronto.

Church & Wellesley Opportunity

This is a region on the north border of the downtown district. It is well served with Japanese restaurants by has no Italian or Seafood restaurants.

Central Bay Street Opportunity

This region is overserved with Italian restaurants but underserved on all other types.

Conclusion

Ingesting and analysing the data from a city can produce valuable insights. There is no question that the Toronto data provides clear insights into where there are potential business opportunities in the city. By looking at popular areas that are underserved with popular businesses an investor would have a shortlist of areas to investigate.

A similar analysis could be conducted in any city that has the data available.

It would also be possible to do specific analysis around niche businesses. So, a focused analysis could be undertaken on Italian restaurants for example. It may be desirable to find the most popular area for niche restaurants that is underserved in niche suppliers.