

Selecting Neighbourhoods in Toronto for Establishing a New Coffee Shop

Table of contents

Introduction: Business Problem

Data

Methodology

Analysis

Results and Discussion

Conclusion

Introduction: Business Problem

A business organization is considering to setup a new coffee shop in Toronto. This business organization is the audience of the data

In this regard, it is desired to provide answers to the following questions:

1. Which neighbourhoods in Toronto have high business activity in terms of existing outlets and venues?
2. Is coffee shop a viable business category in Toronto? Do enough coffee shops exist to prove existence of a coffee consumer market?
3. What are popular business venues in Toronto? Are some coffee shops popular venues?
4. Which neighbourhoods in Toronto have highest and lowest number of coffee shops?
5. Which neighbourhoods in Toronto have high business activity and low competition from existing coffee shops? These neighbourhoods will be potential places or areas to be selected for establishing a new coffee shop.

Data

The following data sources have been used in preparation of this report:

1. List of postal codes of Canada: This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with M are located within the city of Toronto in the province of Ontario.
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
2. Geospatial data: This comma separated value (CSV) file contains geographical coordinates (latitude and longitude values) of the postal codes in Toronto.
https://cocl.us/Geospatial_data
3. Information of venues, venue categories, location latitude and location longitude have been collected through API calls to Foursquare website.
<https://api.foursquare.com/>

Methodology

Python programming language has been used to conduct the analysis. BeautifulSoup library has been used for webscrapping, Pandas for data wrangling, Matplotlib and Seaborn for creating charts, Nominatim and Geocoder to retrieve geographical coordinates, and Folium library has been used to display geographical maps.

Bar charts and tables have been to display tabular data. Regression and coorelation analysis have been conducted to find out relationships between variables.

The results has been portrayed using tables and geographical maps.

Analysis

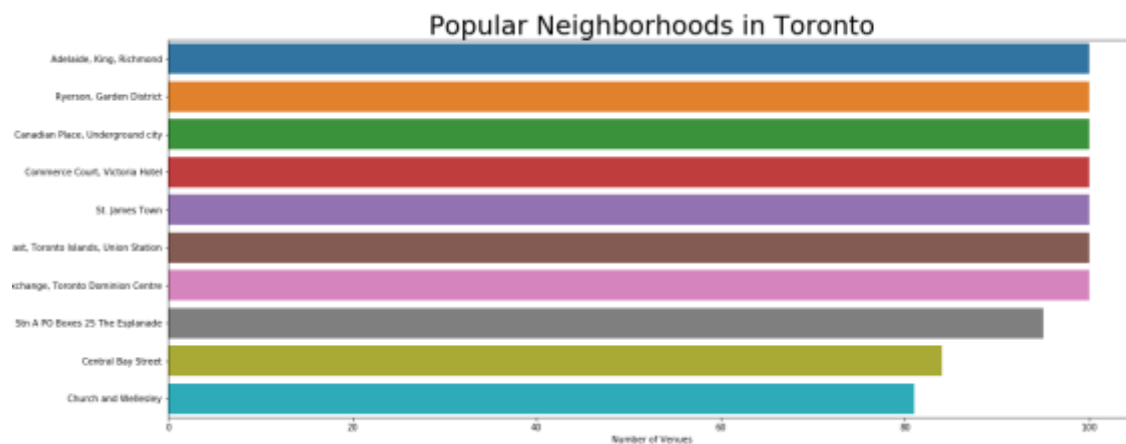
In first step, we get a list of neighbourhoods and postal addresses from Wikipedia page and then merge this with the geographical coordinates data downloaded through the Coursera website. For a visual understanding, we draw these neighbourhoods using a map for a visual representation.

Applied Data Science Capstone Project - Selecting Neighbourhoods in Toronto for Establishing a New Coffee Shop



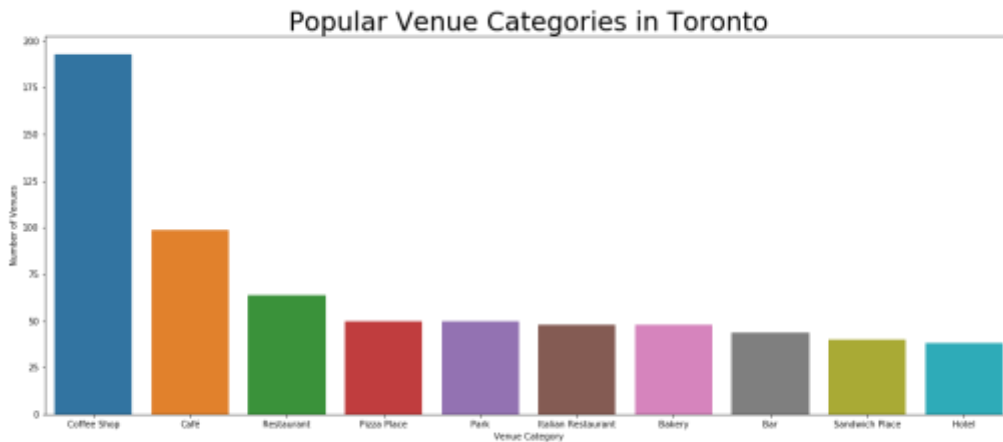
Next, we will collect data of all business venues in Toronto within radius of 1 km of the available neighbourhoods by making API calls to Foursquare website. We will convert the json data into data frame having columns Neighbourhood, Neighbourhood Latitude, Neighbourhood Longitude, Venue, Venue Latitude, Venue Longitude, Venue Category.

Next, summarize the data to identify the popular neighbourhoods in Toronto and draw a bar chart of these.



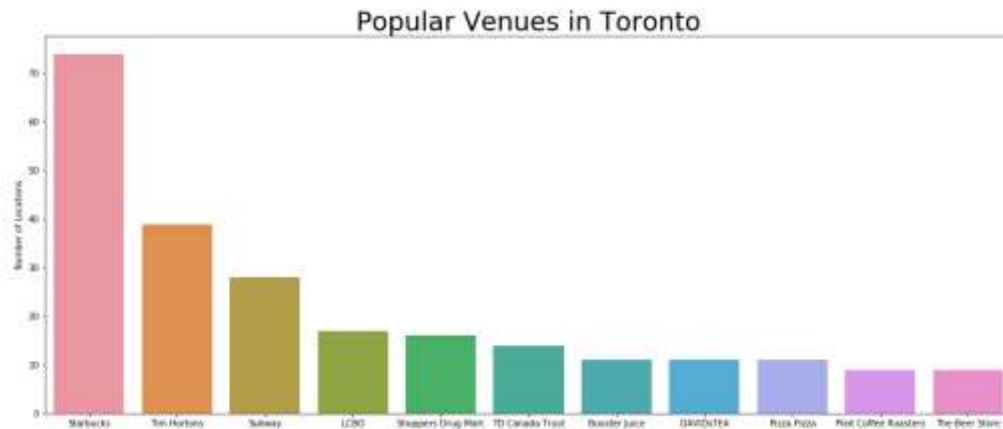
Next, we will summarize the data for number of venue categories in each neighbourhood for this we shall count the number of venue categories in the sample, reset index and then rename columns.

Applied Data Science Capstone Project - Selecting Neighbourhoods in Toronto for Establishing a New Coffee Shop



We learn from the above graph that **coffee shops** are indeed the most popular category followed by cafe. This answers the audience's questions whether sufficiently large consumer market of coffee exists and whether it could be a viable business.

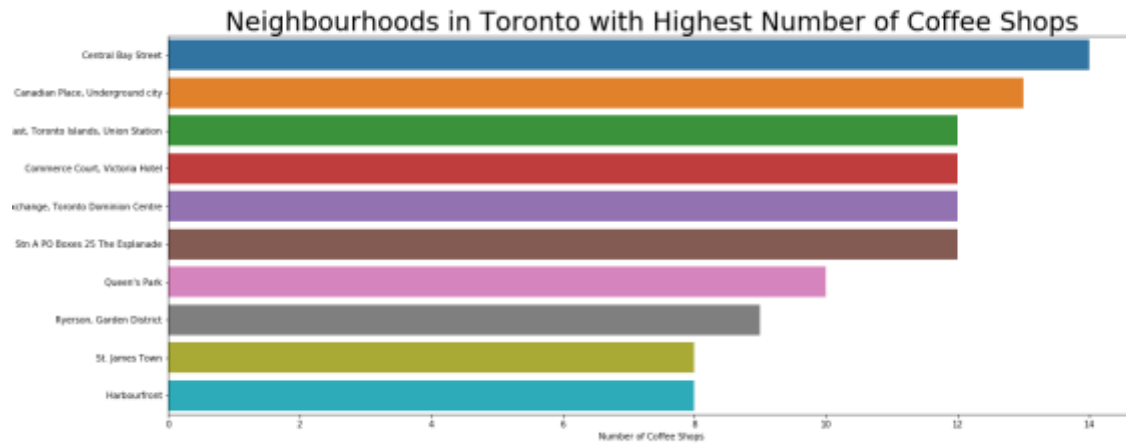
Next, we analyse data to find out popular business venues in terms of number of their existences or locations.



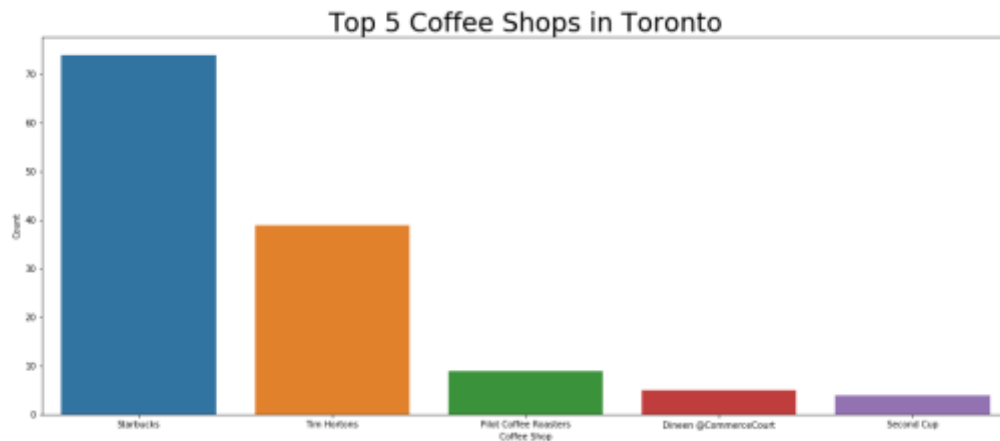
The above graph shows a variety of venues among the top 10 popular venues with **coffee shops** occupying the top two positions.

Thus far, we have identified the popular neighbourhoods, categories and venues in Toronto. Now, let us which locations have the most potential for our audience who want to open a new coffee shop. For that we will identify the top 10 neighbourhoods in terms of concentration of coffee shops.

Applied Data Science Capstone Project - Selecting Neighbourhoods in Toronto for Establishing a New Coffee Shop



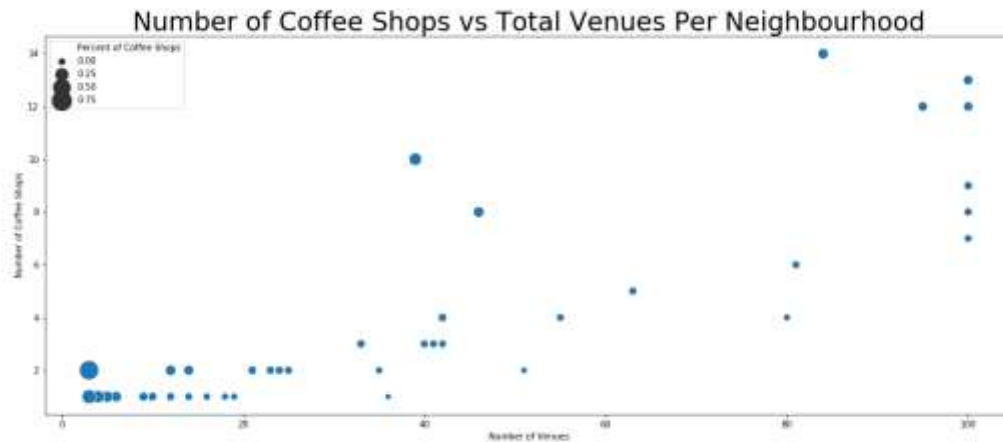
At the same time, our audience is interested in knowing which are the popular coffee brands popular in Toronto.



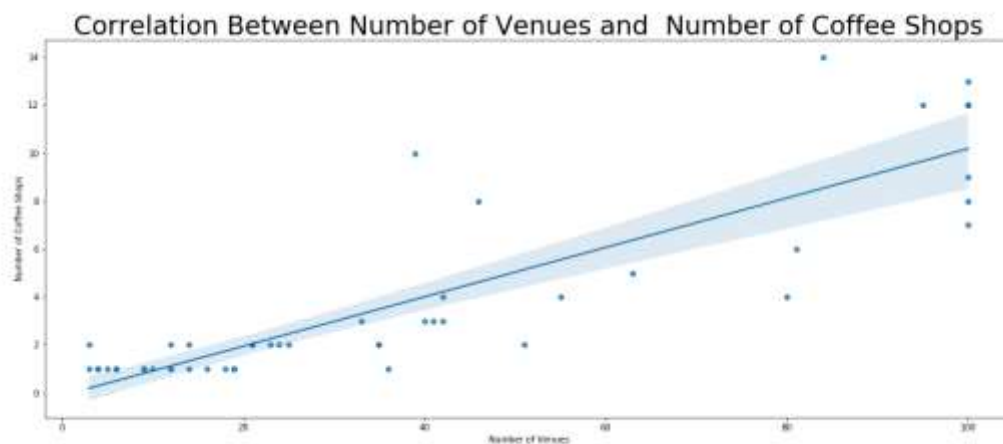
We learn that in Toronto Starbucks is the most popular **coffee shop** venue followed by Tim Hortons.

We shall see through a scatter plot whether a relationship exists between the number of business venues and the number of coffee shops in a neighbourhood.

Applied Data Science Capstone Project - Selecting Neighbourhoods in Toronto for Establishing a New Coffee Shop



We would also like to identify a correlation between the overall number of venues and number of coffee shops in that neighbourhood. This is to answer the question that whether opening a coffee shop in a busy neighbourhood would be a good choice.



The above graph depicts that a positive correlation exists between overall number of venues and number of **coffee shops** in that neighbourhood.

Results and Discussion

List of Potential Neighbourhoods

Towards the last part of our analysis, we shall filter out a list of potential neighbourhoods for opening a new coffee shop.

Our criteria for selection of a potential neighbourhood for opening a new coffee shop are:

Applied Data Science Capstone Project -
Selecting Neighbourhoods in Toronto for Establishing a New Coffee Shop

1. The number of total venues in the neighbourhood must be 50 or more. This is to ensure that ample commercial activity and customer base exists in the neighbourhood.
2. The number of existing coffee shops should be less than 20% of the total venues in that neighbourhood. This is to avoid entering a market which is already crowded by competing coffee shops.

After applying the above criteria, fourteen neighbourhoods have been identified as potential areas for opening a new coffee shop venue.

Sr.	Neighbourhood	Number of Venues	Number of Coffee Shops	Percent of Coffee Shops
0	St. James Town	100	8	8%
1	Ryerson, Garden District	100	9	9%
2	Commerce Court, Victoria Hotel	100	12	12%
3	First Canadian Place, Underground city	100	13	13%
4	Design Exchange, Toronto Dominion Centre	100	12	12%
5	Harbourfront East, Toronto Islands, Union Station	100	12	12%
6	Adelaide, King, Richmond	100	7	7%
7	Stn A PO Boxes 25 The Esplanade	95	12	13%
8	Central Bay Street	84	14	17%
9	Church and Wellesley	81	6	7%
10	Chinatown, Grange Park, Kensington Market	80	4	5%
11	Fairview, Henry Farm, Oriole	63	5	8%
12	Berczy Park	55	4	7%
13	Little Portugal, Trinity	51	2	4%

Lastly, we will mark the identified potential neighbourhoods on the map for visual representation.

Applied Data Science Capstone Project -
Selecting Neighbourhoods in Toronto for Establishing a New Coffee Shop



Conclusion

From the above analysis, we learn that:

1. A number of neighbourhoods in Toronto have high business activity in terms of existing venues. The top ten neighbourhoods in terms of business concentration have more than 80 venues each.
2. Coffee shop is a viable venue or business category in Toronto. In fact, it is the largest venue category comprising of 193 coffee shops with café venues with 93 locations at the second position.
3. The top ten popular business venues in Toronto comprises of coffee shops, fast food chains, drugstores, retail stores and bank branches. Coffee shops are at top of the list these venue categories.
4. Highest number of coffee shops are located in Central Bay Street, First Canada Place, Underground City, Harbourfront East, Toronto Islands, and Union Station.
5. A list of potential neighbourhoods for opening a new coffee shop has been identified which have number of total venues in the neighbourhood must be 50 or more and the number of existing coffee shops be less than 20% of the total venues in that neighbourhood.
