

Restaurant Theme and Location in Downtown Toronto

Capstone Project – Final Report

For the IBM Data Science Professional Course

By Robert Marvin

Date: 17/07/2017

Contents

1. Introduction.....	3
1.1 The Client.....	3
1.2 Strategy.....	3
1.3 Data and Tools.....	4
3. Methods.....	5
3.1 Data collection.....	5
3.2 Data visualisation	6
4. Results.....	8
4.1 Restaurant Theme	8
4.2 Restaurant Location	8
5. Discussion & Conclusion	9

1. Introduction

1.1 The Client

The problem as outlined by the client is as follows. They are a national restaurant chain who have grown in size considerably due to their ability to listen to customer preferences. Their current strategy is based on a feedback system as follows:

1. Surveys are collected about which style of restaurant people would like to see added to the area they live in.
2. A restaurant is then themed on the highest chosen.

This current strategy requires additional staffing to conduct the surveys and often takes weeks or months to complete before a decision can be made. As the company is in a position to grow even faster, this process is a bottleneck in the company's' progression and as such, a new strategy for quickly determining the most successful theme of the restaurant to open is vital. Herein, we propose a strategy based on collecting data and utilizing location data to determine this parameter instantly.

1.2 Strategy

The purpose of this project is to determine which theme of the restaurant will be most successful in the Downtown Toronto area based on the selection of themes that are available in the whole of Toronto, Ontario. The most abundant styles of restaurant in Toronto will be compared with the styles available in the Downtown Toronto area. If a theme or selection of themes that are abundant in Toronto are less so in the Downtown Toronto area, this will help to guide the client when making a decision on which style of restaurant to open.

The top 3 restaurant themes will be extracted from the whole of Toronto and compared with the Downtown area. Determining the number of restaurants with these themes from the Toronto area, the 3 lowest abundant restaurants will be determined to help the client make a decision based on their knowledge of start-up costs for each of those 3 restaurant themes.

Once the decision has been made on which style of restaurant will be opened, location data will be used to determine the placement of the restaurant in the neighbourhood to minimise rivalry.

1.3 Data and Tools

The following data will be used for the investigation:

1. Number of each theme of restaurant in Toronto
2. Number of each theme of restaurant in the Downtown Toronto area
3. Placement of restaurant themes

The data will be scraped from available data on Wikipedia and sorted for analysis. This will be achieved with commonly used python libraries such as BeautifulSoup, numpy and pandas for data analysis and with foursquare API and geopy for location analysis.

The data was obtained from the following sources:

1. Postcode data:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
2. Coordinates: [Foursquare API](#)

Finally, data visualisation was carried out using matplotlib and folium.

3. Methods

3.1 Data collection

Firstly, postcode data for Toronto, Ontario was collected from the Wikipedia page. The BeautifulSoup package was used to extract data from the Wikipedia page which was in tabular form as shown in figure 3-1.

Toronto - 103 FSAs [edit]		
Note: There are no rural FSAs in Toronto, hence no postal codes should start with M0. However, the postal code M0R 8T0 is assigned to an Amazon warehouse in Mississauga, suggesting it may have reserved the M0 FSA for high volume addresses.		
Postal Code ↕	Borough ↕	Neighborhood ↕
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights
M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
M8A	Not assigned	Not assigned
M9A	Etobicoke	Islington Avenue, Humber Valley Village
M1B	Scarborough	Malvern, Rouge
M2B	Not assigned	Not assigned
M3B	North York	Don Mills
M4B	East York	Parkview Hill, Woodbine Gardens
M5B	Downtown Toronto	Garden District, Ryerson
M6B	North York	Glendale
M7B	Not assigned	Not assigned
M8B	Not assigned	Not assigned
M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grove, Islington, Cloverdale

Figure 3-1: List of postcodes, boroughs and neighbourhoods in Toronto, Ontario from Wikipedia page¹.

The data was then formatted into a data frame ready for cleaning. All boroughs and neighbourhood values with 'not assigned' as the cell value were removed from the data frame.

Next, the coordinates obtained from geopy were uploaded from a readily formatted CSV file and merged with the previously mentioned data frame to obtain a data frame with all the information needed for analysis as shown in figure 3-2.

	Postal Code	Borough	Neighborhood	Latitude	Longitude
98	M9N	York	Weston	43.706876	-79.518188
99	M9P	Etobicoke	Westmount	43.696319	-79.532242
100	M9R	Etobicoke	Kingsview Village, St. Phillips, Martin Grove ...	43.688905	-79.554724
101	M9V	Etobicoke	South Steeles, Silverstone, Humbergate, Jamest...	43.739416	-79.588437
102	M9W	Etobicoke	Northwest, West Humber - Clairville	43.706748	-79.594054

Figure 3-2: Final data frame with relevant information for analysis.

A separate data frame containing the same information for Downtown Toronto was also obtained from the above data frame using pandas to select the Downtown Toronto data rows.

3.2 Data visualisation

Foursquare API was then used to obtain nearby venues and their categories. Once the venues had been placed into a data frame with their corresponding coordinates, all restaurant venues were placed into a new data frame. This data frame was then used to evaluate the number of restaurants and their themes in Toronto and the Downtown Toronto area, as shown in figures 3-3 and 3-4.

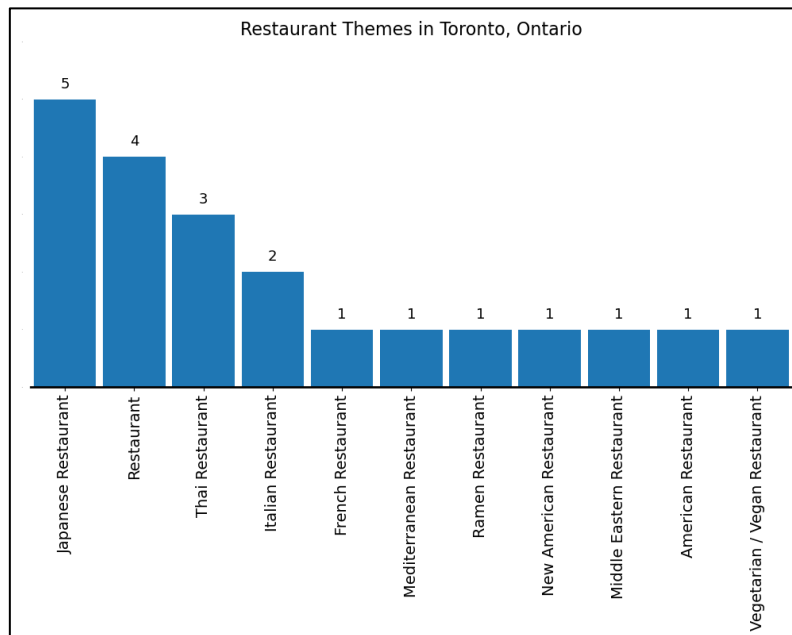


Figure 3-3: Number of each themed restaurant in Toronto, Ontario.

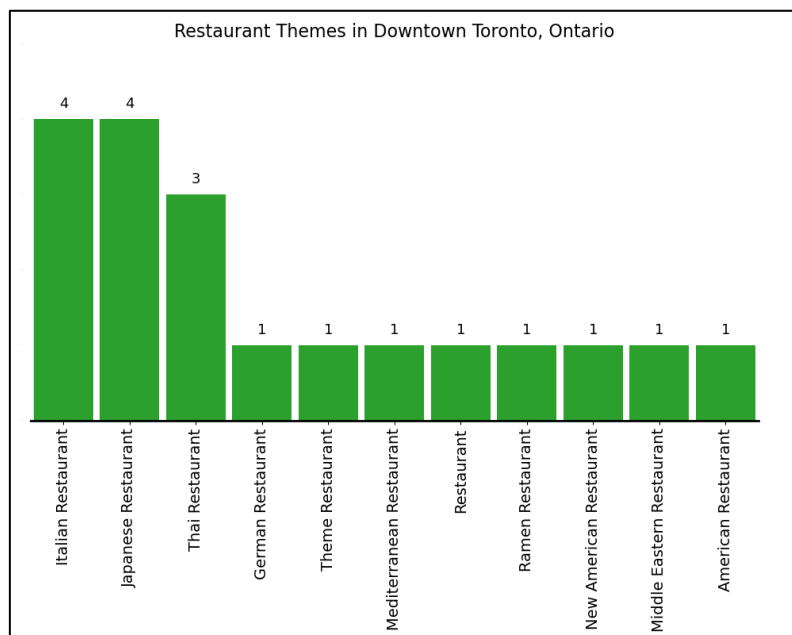


Figure 3-4: Number of each themed restaurant in the Downtown Toronto Borough.

Maps were also plotted for the top restaurants in the Downtown Toronto borough as shown in figure 3-5.

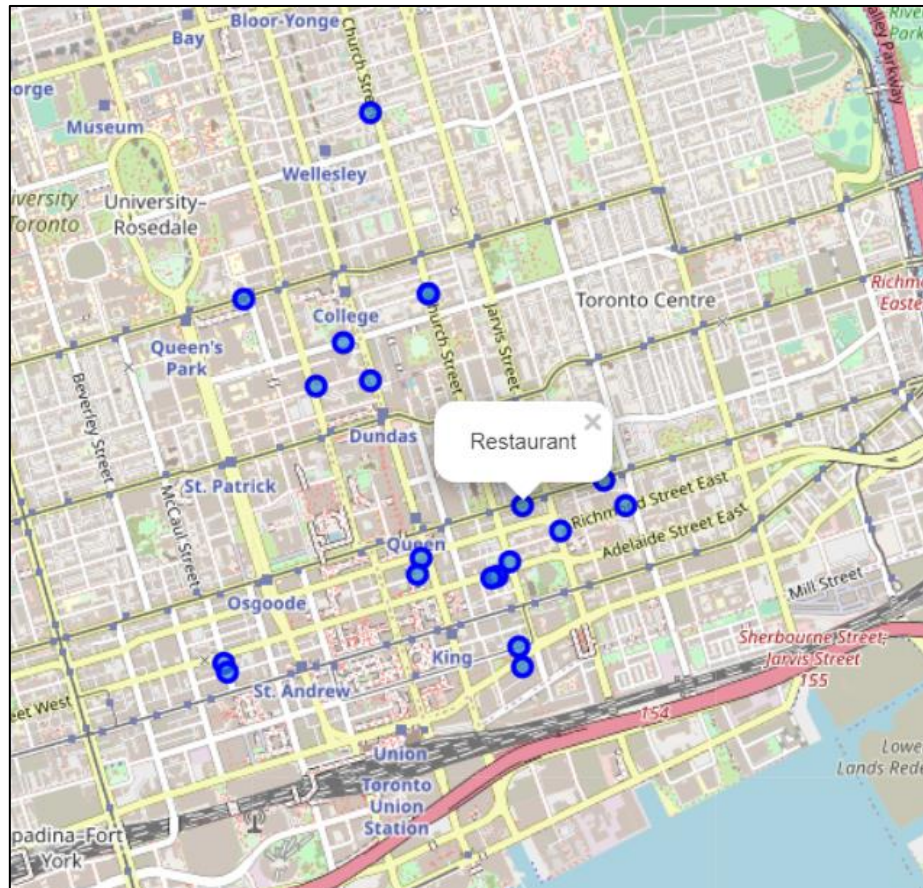


Figure 3-5: Map of Downtown Toronto showing top restaurants in the area.

4. Results

4.1 Restaurant Theme

From the plot shown in figures 3-3, it can be seen that the most popular restaurants in Toronto are:

1. Japanese
2. General restaurant
3. Thai

Next, this data is compared with the number and themes of restaurants in Downtown Toronto, as shown in figure 3-4.

From these two datasets we can see that the most abundant restaurants in Toronto and Downtown Toronto are Japanese restaurants which is no surprise. Following this, the next most abundant in Toronto is a general food restaurant which there seems to be only 1 of in Downtown Toronto. Finally, the Thai restaurant is the third most popular in Toronto which is also true for downtown Toronto. So, based on the findings for the top 3 most popular restaurants in both areas, there appears to be a gap in the market for a **general food restaurant** in the Downtown Toronto area so this is what will be suggested to the client.

4.2 Restaurant Location

To further assist the client, the location of the general restaurant in the Downtown Toronto area was plotted on a map, as shown in figure 4-1. This gives the client the ability to make an informed decision about where to place the new restaurant.

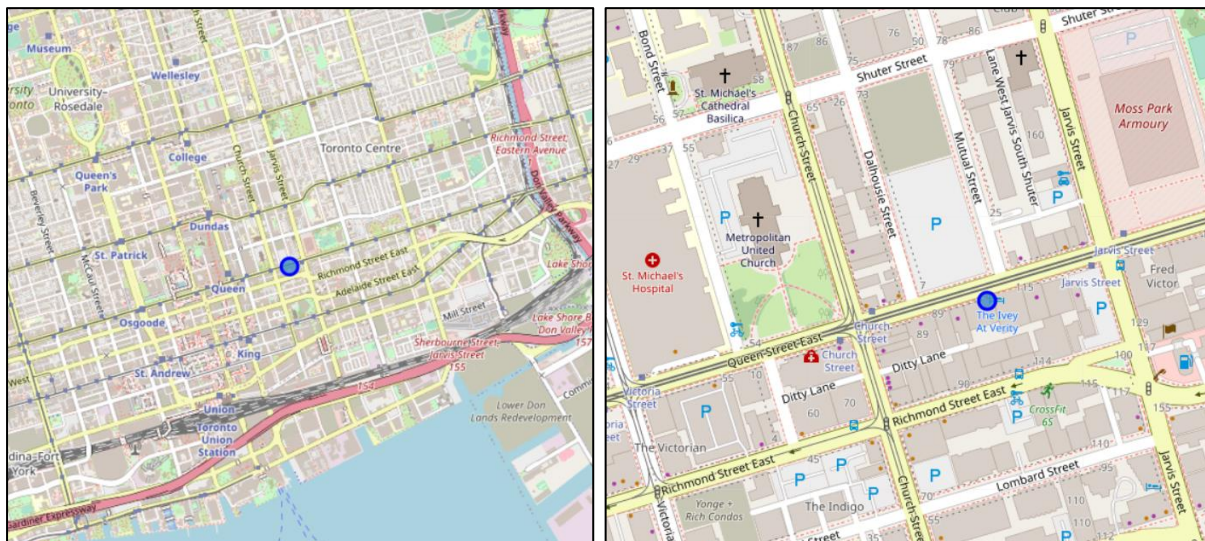


Figure 4-1: Location of the restaurant in Downtown Toronto (left), Zoomed in (right).

5. Discussion & Conclusion

Based on the findings, our client was able to make an informed decision on which style of restaurant to open in the Downtown Toronto borough and where to begin to look for a venue to open the restaurant. This saved time and resources for the client. Moreover, this process can be applied to future endeavours for the company all around the country to find locations of themed restaurants.

Further analysis could be done on demographic data which may have a large impact on the potential success of a particular theme of restaurant.

Additional machine learning techniques could even extend this analysis. For example, regression techniques could be applied to past data to predict the success of a particular theme of restaurant.

References

1. Wikipedia. List of postal codes of Canada: M. *Wikipedia*
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M (2020).