

# Data Science Capstone Report

---

## 1.0 Introduction

### 1.1 Background

This project will make use data science analytics to guide a new business project.

### 1.2 Business Problem

A client would like open an Italian restaurant in the downtown area of Toronto. They would like to know which neighbourhood would be best to open a new restaurant. The client is opening a more upscale restaurant so would not like to compete with other Italian restaurants in the area. They would also like to be in an area where there is a good choice of nightlife attractions in the hope customers would most likely look for some place to eat first and then go out afterwards while in the area.

In summary the customer is looking for:

- Downtown Toronto location.
- Area with other types of restaurants.
- Near nightlife venues.
- Limited competition for Italian restaurants.

## 2.0 Data acquisition and cleaning

### 2.1 Data Sources

We will use the Four Square API and its a data sources to obtain information on the overall areas the client is interested. In this case the downtown Toronto neighbourhoods. We will make use of the venue categories to narrow our search criterea. The following catgeories will be used:

- Food
- Italian Restaurant
- Nighlife Spot

The Food category will be used to obtain an overall idea of the number of restaurants in a given neighbourhood. The Italian Restaurant category will by used to determine if a neighbourhood already has a good selection to choose from.

We will merge the Nightlife and Restaurant location data to get an idea for a possible location.

We will also use the Wikipedia data on postal locations for the Toronto area to map the neighbourhoods actual location.

#### 2.1.1 References

Links to information on the data sources used:

[Foursquare Search Endpoint Details](#)

[Foursquare Venue Categories](#)

[Wikipedia Toronto Postal Codes](#)

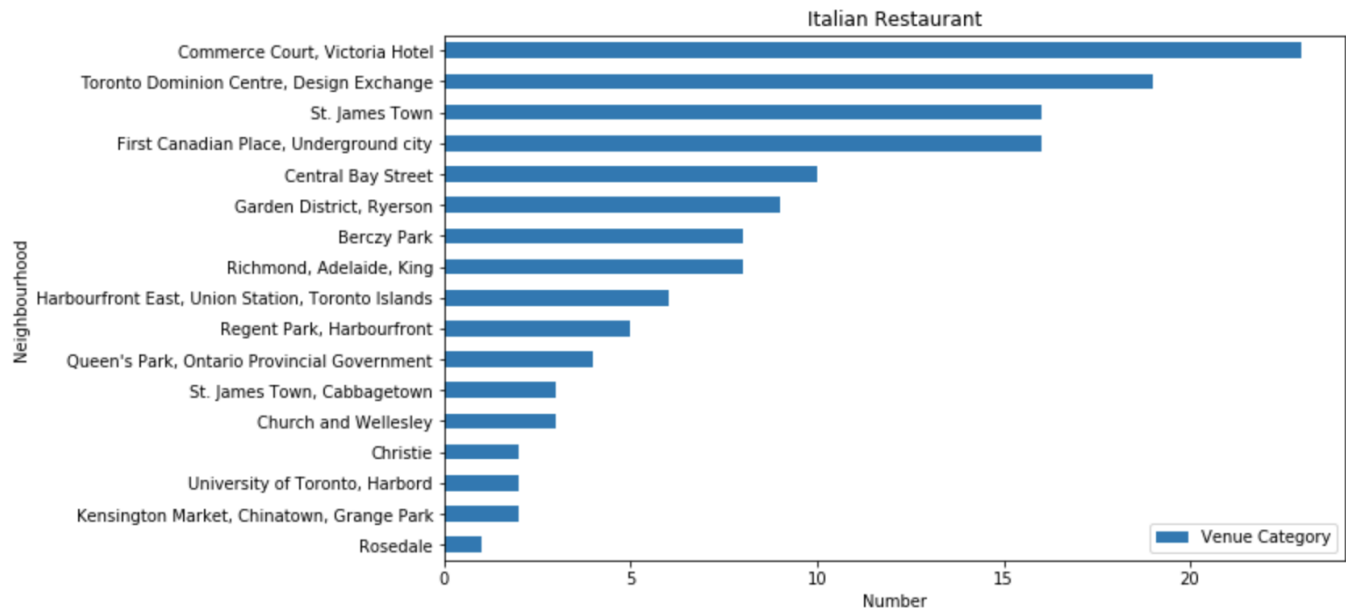
2.2 Data Cleaning

We will narrow the search from the Toronto area to the downtown area only. We will also throw out any P.O. Box locations as this is not a real neighbourhood.

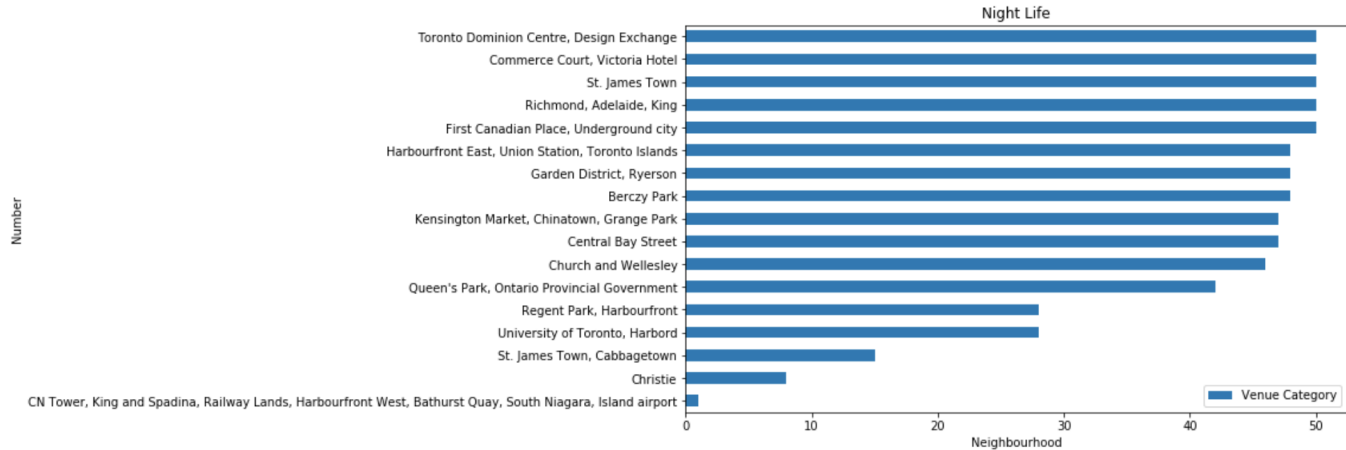
We will also clean up some of the Italian Restaurants. Included in the search were Italian bakeries and Pizza places which are not the type of restaurant the client is looking to open so these will be dropped.

3.0 Exploritory Data Analysis

The first step was to find the Italian restraurants in each neighbourhood. This was done using the Foursquare data to search for Italian restraurants in the downtown Toronto area. The following chart shows the result of that search in each neighbourhood:



The next step was to repeat the process but for nightlife venus in each neighbourhood using the Foursquared data. The following chart shows the number of nightlife venues in each neighbourhood:



This data was then merged with the Italian restaurant selections.

The last step was to search for all restaurants in each neighbourhood to determine if the neighborhood is a popular food destination. This data was then merged with the other two. The following table shows all three criteria merged together:

	Italian Restaurant	Restaurants	Nightlife
Neighbourhood			
Kensington Market, Chinatown, Grange Park	2	50	47
University of Toronto, Harbord	2	49	28
Christie	2	27	8
Church and Wellesley	3	50	46
St. James Town, Cabbagetown	3	50	15
Queen's Park, Ontario Provincial Government	4	50	42
Regent Park, Harbourfront	5	50	28
Harbourfront East, Union Station, Toronto Islands	6	50	48
Richmond, Adelaide, King	8	50	50
Berczy Park	8	50	48
Garden District, Ryerson	9	50	48
Central Bay Street	10	50	47
First Canadian Place, Underground city	16	50	50
St. James Town	16	50	50
Toronto Dominion Centre, Design Exchange	19	50	50
Commerce Court, Victoria Hotel	23	50	50

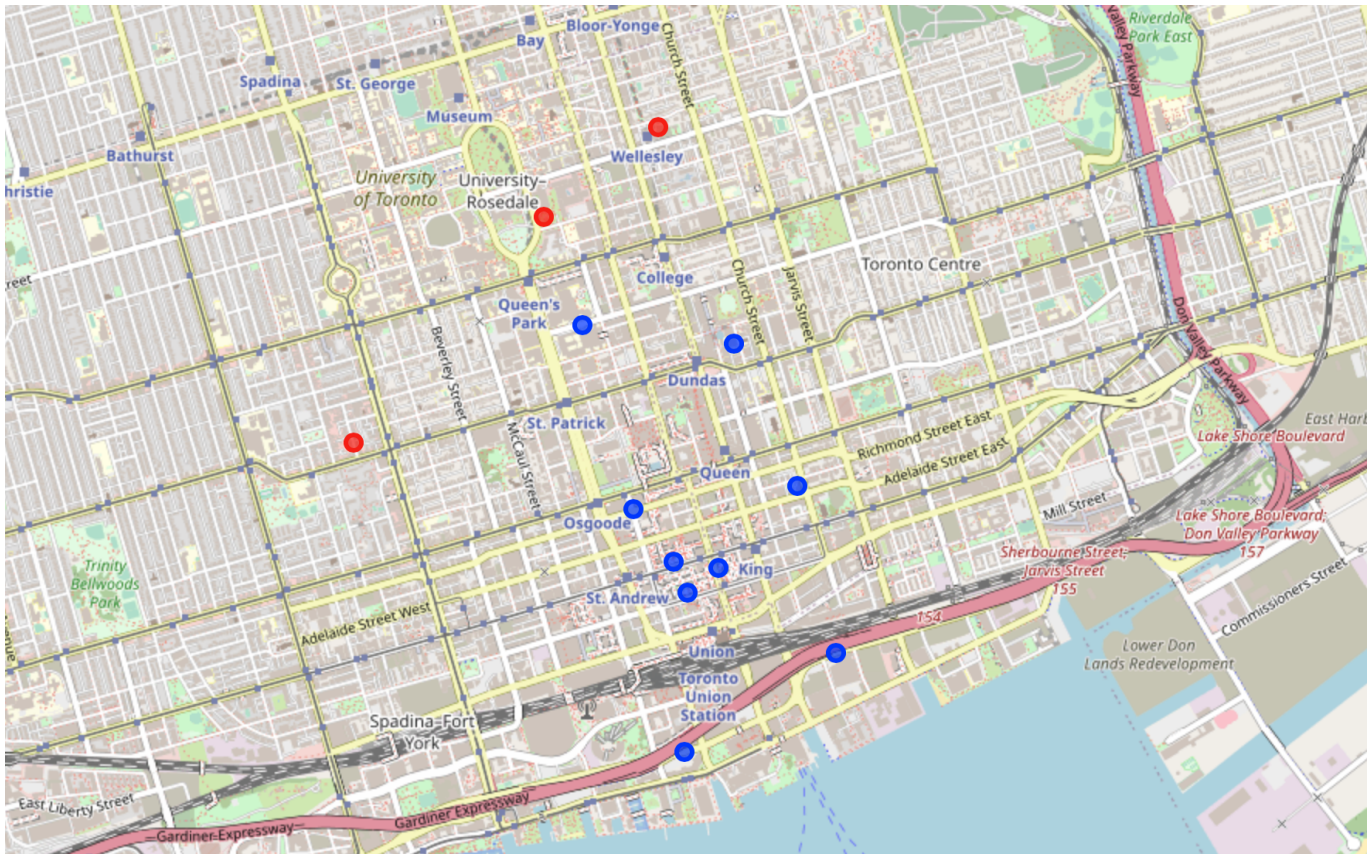
From this table locations that have less than 30 Food or Nightlife venues in that neighbourhood were filtered out as the criteria was to be in a popular area. The following table is the final list that meets the original criteria:

	Italian Restaurant	Restaurants	Nightlife
Neighbourhood			
<b>Kensington Market, Chinatown, Grange Park</b>	2	50	47
<b>Church and Wellesley</b>	3	50	46
<b>Queen's Park, Ontario Provincial Government</b>	4	50	42
<b>Harbourfront East, Union Station, Toronto Islands</b>	6	50	48
<b>Richmond, Adelaide, King</b>	8	50	50
<b>Berczy Park</b>	8	50	48
<b>Garden District, Ryerson</b>	9	50	48
<b>Central Bay Street</b>	10	50	47
<b>First Canadian Place, Underground city</b>	16	50	50
<b>St. James Town</b>	16	50	50
<b>Toronto Dominion Centre, Design Exchange</b>	19	50	50
<b>Commerce Court, Victoria Hotel</b>	23	50	50

The neighbourhoods were then mapped to their latitude and longitude so that they can be plotted on a map.

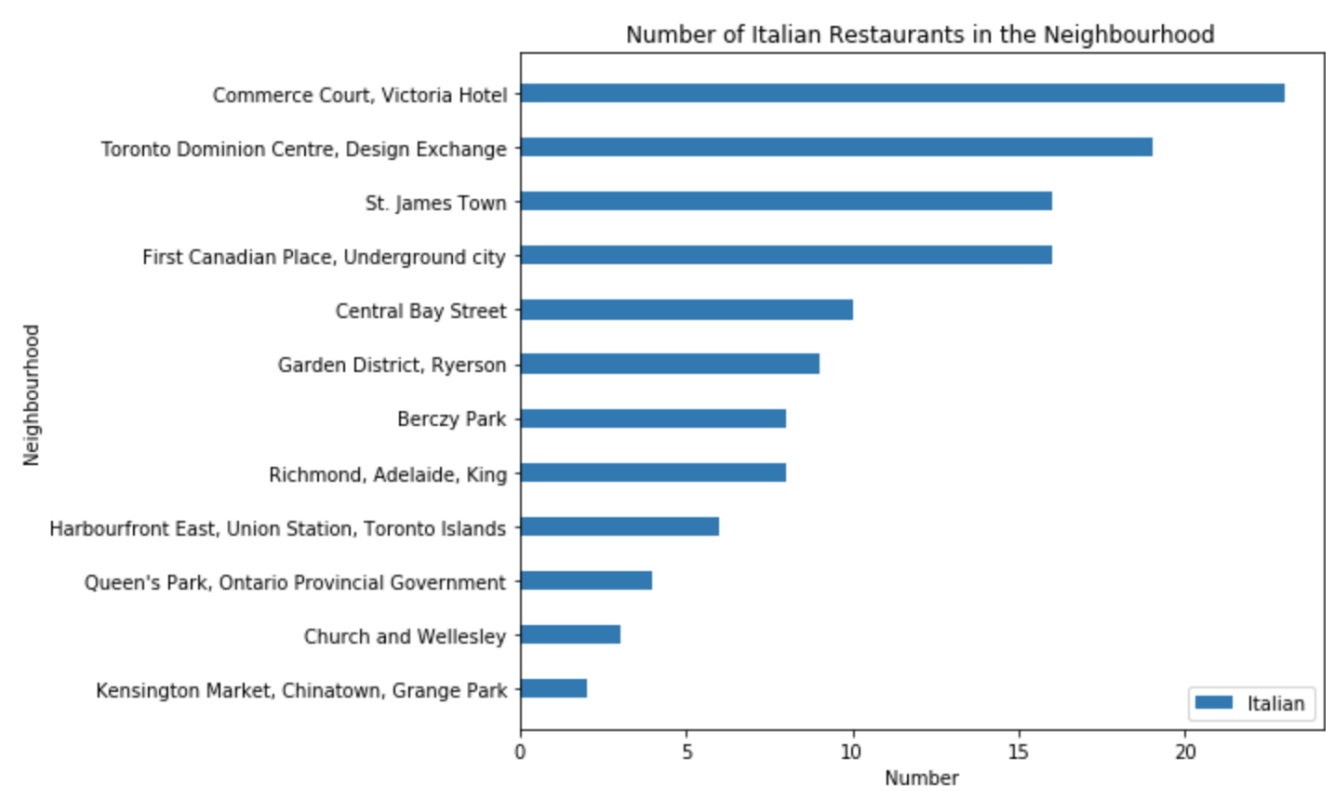
## 4.0 Conclusion

The following map illustrates the potential locations by neighbourhood:



The sites that are in red are the top three choices.

The following chart is the narrowed down choice based on the original criteria:



From the chart we can see there are three potential locations that have little competition but have an active nightlife and restaurant scene. The Kensington Market, Chinatown, Grange Park area would be the first choice. As it has low competition from other Italian Restaurants.