Analysing the Restaurants in Canada

1) Introduction
    a. Background and problem

    When deciding on opening new restaurants, there are many things should be considered. Especially regarding the current locations of other restuarants. Analysing on how the restaurants are located and if there are other venues which are present near the restaurant can help the owners of the new restaurant to draw a conclusion.

    In this project, data from Foursquare will be leveraged to analyse and recommend the factors the new restaurant owners can consider before opening a new restaurant. Number of restaurant in the neighbourhood and other venues types near the restaurants had to be analysed before concluding.

    b. Interest

    Drawing a conclusion on opening a restaurant is quite vital as there might be huge financial decisions. Leveraging data science and data such as current location of restaurants, nearest venue types and identifying where these restaurants are clustered can help in making a solid decision.

2) Data
    a. Sources
        1) Main table

        Data scraped from
        https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

        This contains the data of boroughs in Canada. After scraping the data, it was placed in a data frame so that it can be used for analysis.
        2) Geopy

        Geopy was used to extract the latitude and longitude of every neighbourhood and restaurants in the data set. This was then passed to Foursquare to extract details on venues.
        3) Foursquare

        Foursquare is a location provider which can be leveraged to extract more details with the latitude and longitude provided. It was used for the below scenarios:
            a. Venues in each borough was extracted with the help of Foursquare
            b. After filtering the restaurants in each borough, Foursquare was leveraged again to obtain other venue types near the restaurant

    b. Cleaning and feature selection

    There were some cases that did not have cities or address. Those were removed. Some of the restaurant had a long name so splitting them by delimiter such as space

was not feasible. Thus, the category names that had restaurants were provided the type as restaurant.

In [*]:
```
# Filters the restaurant types
venues_res = venues_city[venues_city['category_name'].str.contains('Restaurant')]
# Drop the values which are null
venues_res = venues_res[~venues_res['location.address'].isnull()]
venues_res.head(5)
```

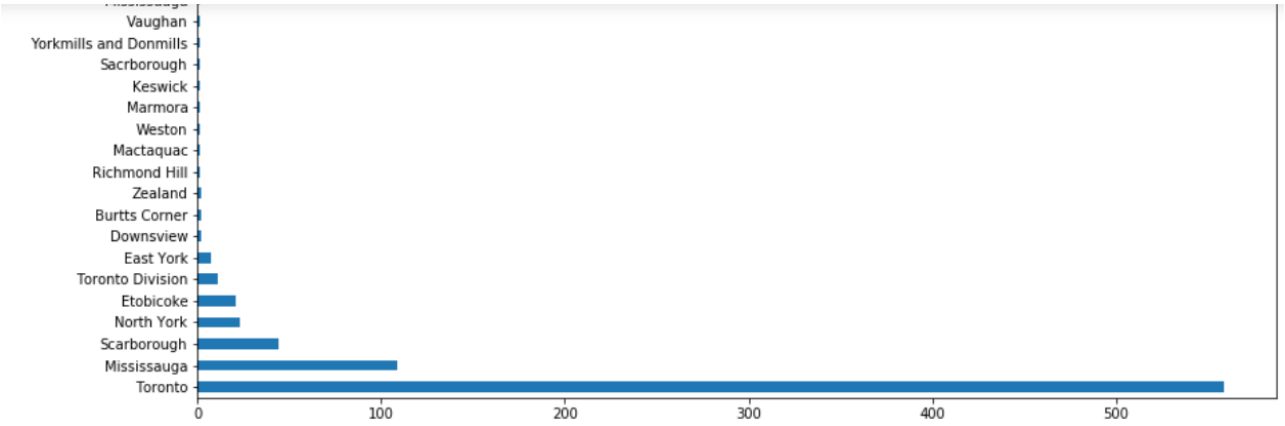
Out[12]:

	categories	category_name	hasPerk	id	location.address	location.cc	location.city	location.country	location.c
469	[{'id': '4bf58dd8d48988d115941735', 'name': 'M...	Middle Eastern Restaurant	False	4c9e7135542b224b753ff49f	564 wilson heights blvd.	CA	Toronto	Canada	
470	[{'id': '4bf58dd8d48988d14e941735', 'name': 'A...	American Restaurant	False	4b4a37b6f964a520387f26e3	531 Wilson Heights Blvd	CA	Toronto	Canada	
471	[{'id': '4bf58dd8d48988d115941735', 'name': 'M...	Middle Eastern Restaurant	False	4bc3acbdabf495210a22c493	660 Sheppard Avenue West	CA	Toronto	Canada	
472	[{'id': '4bf58dd8d48988d1d2941735', 'name': 'S...	Sushi Restaurant	False	4c9fd84e542b224b3f2701a0	638 Sheppard Ave W	CA	Toronto	Canada	
473	[{'id': '4bf58dd8d48988d1c4941735', 'name': 'R...	Restaurant	False	4b65c410f964a52050fe2ae3	634 Sheppard Avenue W.	CA	North York	Canada	

3) Exploratory analysis

a. Distribution of restaurants in Canada

It is quite interesting to note that many of the venues such as restaurants, shopping malls and shops are in Toronto



b. Geo analysis

This Geo analysis was performed on data within Canada. After viewing the map, it can be noted that the restaurants seem to be clustered together.

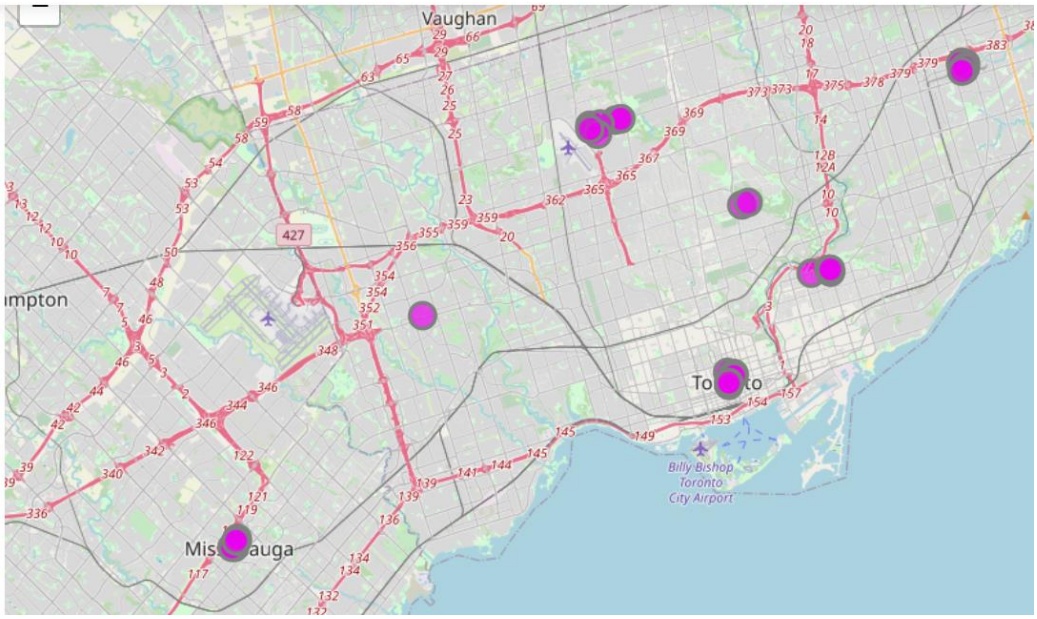


Figure 1 Map of Canada

Since Toronto seem to have the highest number of venues, when zooming in on Toronto, many restaurants are present in Downtown. However, this do not provide more details. So, the map was further zoomed in.

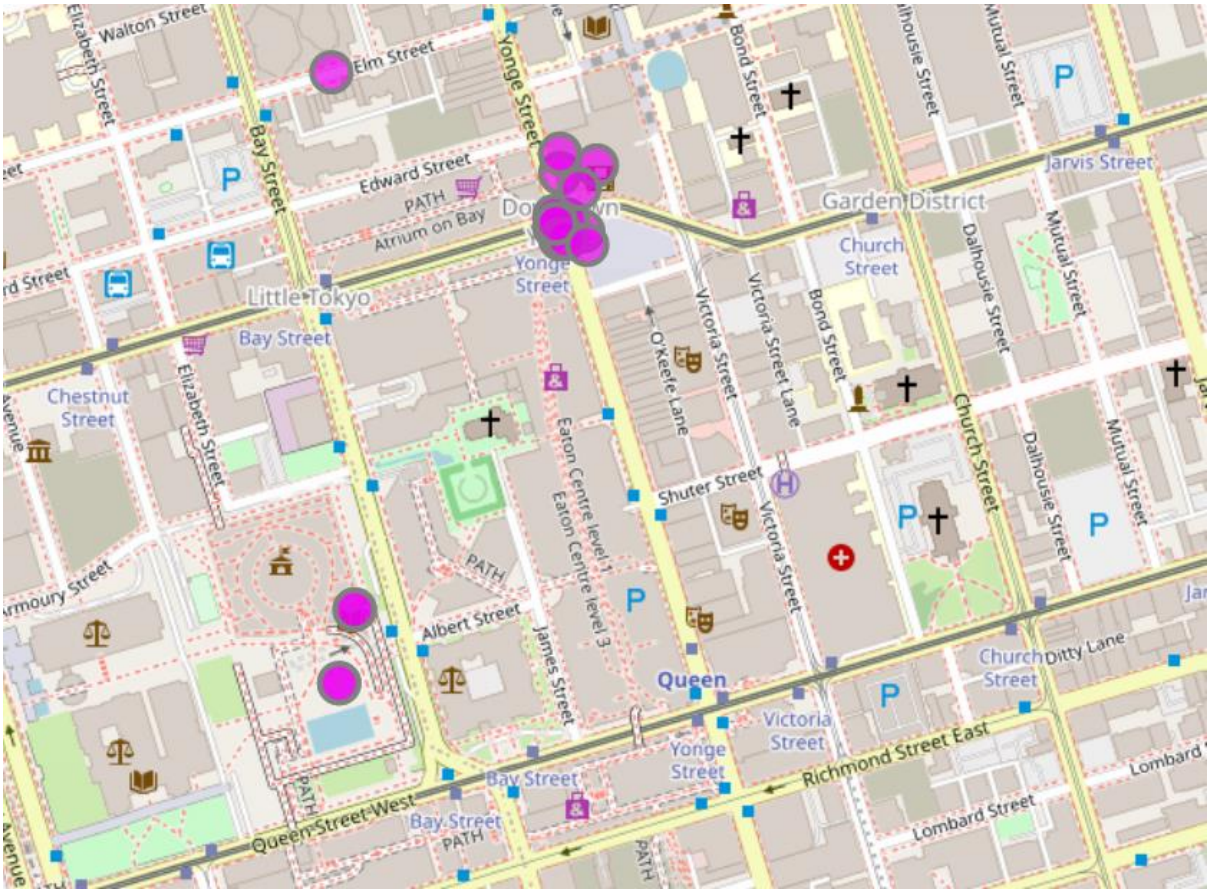
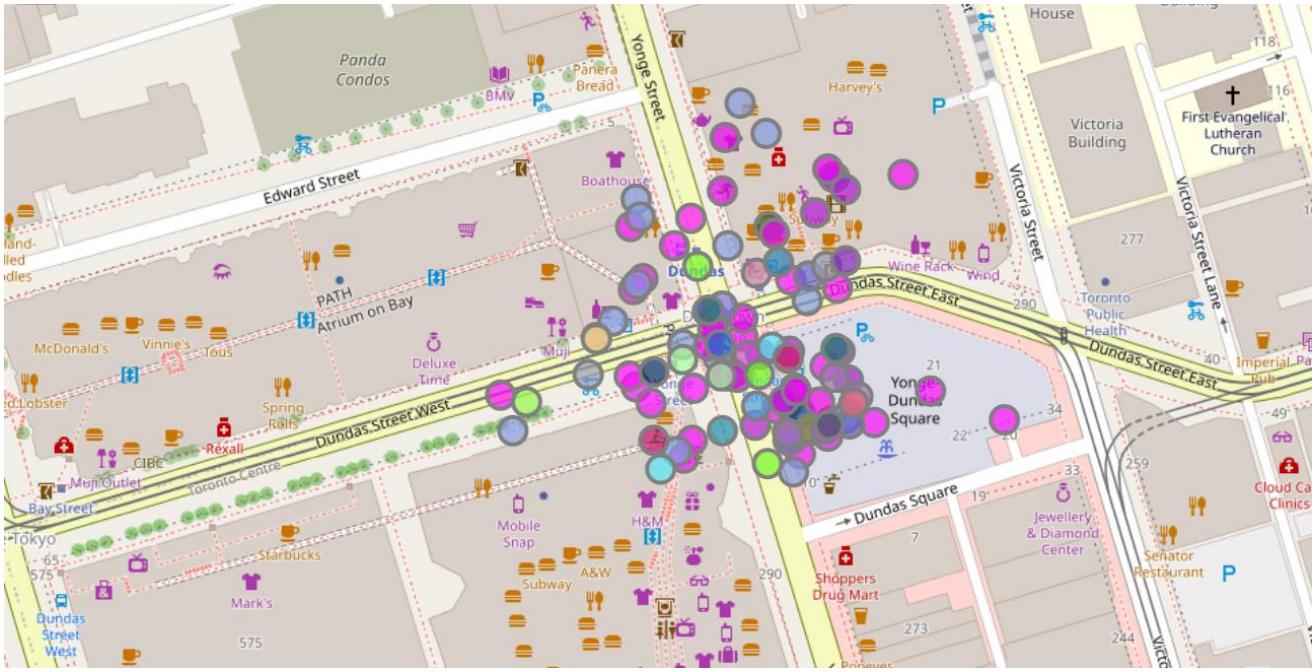


Figure 2 zoomed version of Toronto

After zooming in more, it can be seen that it is near a hospital. This raised the question whether the presence of other venues such as hospital can indicate that there might a restaurant

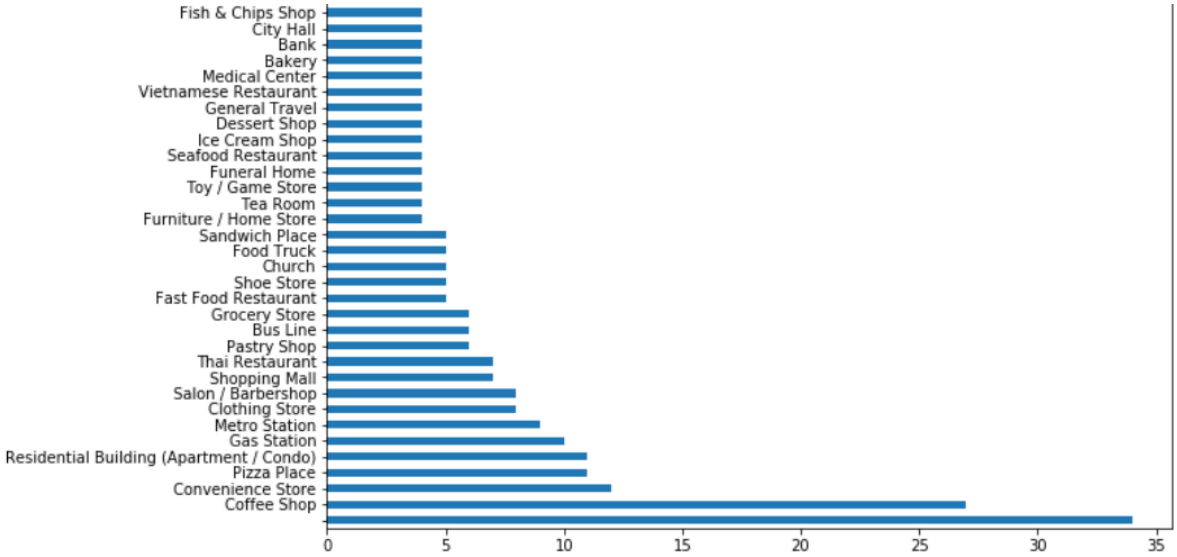
Figure 3 Further zoomed in version of Toronto

As can be seen from below, the same image of hospital with other venues seem to be more congested. Hence better view of analysing the data was required to understand which venue types can be seen near a restaurant more than others.



c. Nearest locations of restaurants

As can be seen from above, it was intriguing to note that the presence of a restaurant might be influenced by other venues also. So, the data from other venues near the restaurants were analysed. It was interesting to note that many restaurants were located near venues such as coffee shops, other restaurants, Metro stations and Gas Stations.



4) Conclusion

As can be seen from the maps, restaurants tend to cluster together with other restaurants. Most of the restaurants seem to be located near Gas station,coffee shops and other restaurants. The clusters seem to be more near the city central rather than the outliers and more restaurants seem are near shopping malls