Battle of Neighbourhood

1. Introduction / Business understanding

Toronto is the largest city in Canada and an economically viable place in terms of business, finance, entertainment, technology and so on. With a population of about 3 million people, the city is ranked as one of the most liveable cities in the world according to the report by the Economist Intelligence Unit (City of Toronto, 2019). Within the city of Toronto, Downtown Toronto can boast of the city's main business areas and the most culturally diverse city in Toronto, Ontario. My client wants to shift to unknown city in Toronto and require help to find the best neighbourhood to start a restaurant. My client wants a neighbourhood that is like his current borough because he feels that would help him settle easily.

2. Target

I will analyse and recommend a place in any of the overall neighbourhoods for my client planning to open a restaurant. The location of interest must be a densely populated area with few or no restaurant. I will utilize the data scientist skills to explore Toronto neighbourhood datasets and extract the needed information. Advantages of each area will then be clearly expressed so that the best possible final location can be chosen by stakeholders

3. Data Acquisition and Cleaning.

3.1 Data Acquisition:

Three different datasets were used, someone of which was scraped from Wikipedia and Canada open data portal. You can view the data through the link shown below;

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

Population Information: https://www12.statcan.gc.ca/censusrecensement/2016/dppd/hlt-fst/pd-pl/Table.cfm?Lang=Eng&T=1201&SR=1&S=22&O=A&RPP=9999&PR=0

Geocoordinates of Canada by Postal codes: http://cocl.us/Geospatial_data

3.2. Data Cleaning:

Some data downloaded or scraped from the website has some missing data. For example, the data in the below Figure was an Html data scraped from Wikipedia. After the dataset was read into a data frame, it was realized that some Boroughs and Neighbourhoods are not assigned.

```
[2]: #scrape Canada html data from web
     html_doc = "https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"
     df_Canada = pd.read_html(html_doc, header=0)
     df_Can = df_Canada[0]
     df_Can.head()
```

Out[2]:

	Postcode	Borough	Neighborhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Harbourfront

NEIGHBOURHOOD VISUALIZATION MAP

TORONTO MAP WITH NEIGHBOURHOODS IN WEST AND DOWNTOWN TORONTO

4. Data Analysis

Using the Foursquare API, the top 100 venues for each neighbourhood were extracted and please find the neighbourhoods stood out as one with the highest number of venues.

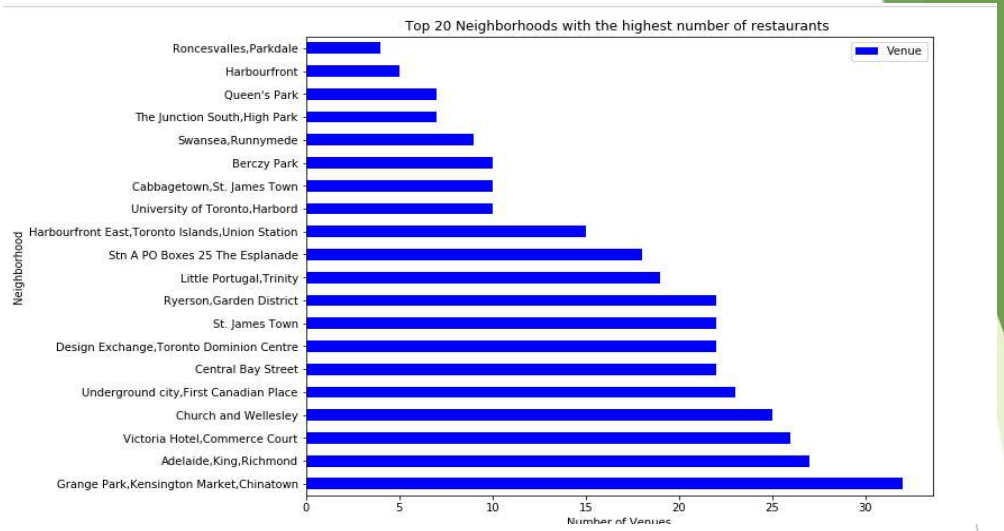
Visualize the restaurants on a Toronto map as displayed below;

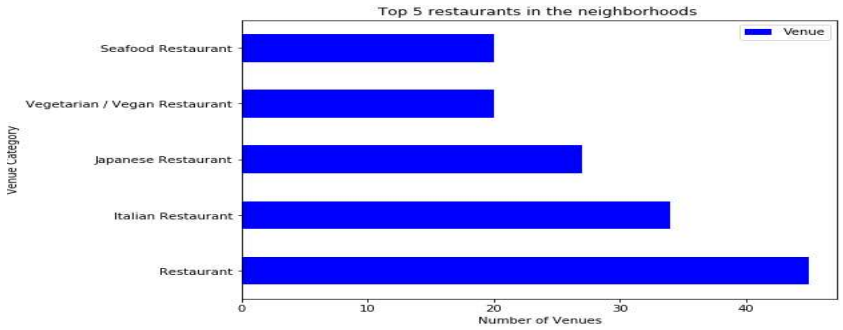
```
#Sort dataframe by venues in descending order
West_Downtown_most_venues_sorted= West_Downtown_most_venues.sort_values('Venue', ascending=False)
# Since the max is 100 , we print dataframe with venues equal to the maximum values
West_Downtown_most_venues_sorted= West_Downtown_most_venues[West_Downtown_most_venues['Venue']==100]
West_Downtown_most_venues_sorted
```

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Adelaide,King,Richmond	100	100	100	100	100	100
Design Exchange,Toronto Dominion Centre	100	100	100	100	100	100
Harbourfront East,Toronto Islands,Union Station	100	100	100	100	100	100
Ryerson,Garden District	100	100	100	100	100	100
St. James Town	100	100	100	100	100	100
Underground city,First Canadian Place	100	100	100	100	100	100
Victoria Hotel,Commerce Court	100	100	100	100	100	100

The top 5 restaurant types in the whole neighborhoods was extracted and shown

The top 20 neighborhoods were ranked according to the number of restaurants





The top 10 venues in each Neighborhood was extracted and shown

The top 5 venues and its frequencies are extracted and shown

```
# add clustering labels
neighborhoods_venues_sorted.head(5)
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adelaide,King,Richmond	Coffee Shop	Café	Bar	Steakhouse	Restaurant	Asian Restaurant	Hotel	Sushi Restaurant	Thai Restaurant	Burger Joint
1	Berczy Park	Coffee Shop	Cocktail Bar	Farmers Market	Beer Bar	Seafood Restaurant	Bakery	Steakhouse	Cheese Shop	Café	Butcher
2	Brockton,Parkdale Village,Exhibition Place	Café	Breakfast Spot	Coffee Shop	Gym	Grocery Store	Pet Store	Performing Arts Venue	Nightclub	Italian Restaurant	Intersection
3	Cabbagetown,St. James Town	Coffee Shop	Restaurant	Pub	Park	Pizza Place	Bakery	Café	Italian Restaurant	Sandwich Place	Breakfast Spot
4	Central Bay Street	Coffee Shop	Café	Italian Restaurant	Burger Joint	Japanese Restaurant	Sandwich Place	Ice Cream Shop	Bar	Gym / Fitness Center	Bakery

```
----Adelaide,King,Richmond----
          venue  freq
0   Coffee Shop  0.08
1    Steakhouse  0.04
2           Bar  0.04
3          Café  0.04
4         Hotel  0.03

----Berczy Park----
                venue  freq
0         Coffee Shop  0.09
1        Cocktail Bar  0.05
2          Steakhouse  0.04
3  Seafood Restaurant  0.04
4      Farmers Market  0.04

----Brockton,Parkdale Village,Exhibition Place----
           venue  freq
0           Café  0.12
1  Breakfast Spot  0.08
2     Coffee Shop  0.08
3             Bar  0.04
4    Burrito Place  0.04
```


York
East York
Toronto

5. Results and Discussion

Recommended location for a restaurant.

- Chinatown, Kensington Market and Grange Park neighbourhood altogether have the highest number of restaurants.

- I would choose Union Station, Toronto Islands, Harbourfront East as the area and I would like to explore for the restaurant opportunities.

- One of the reasons is that it has fewer restaurants and a busy place. We could see that despite having more venues within the neighbourhood s, restaurants are fewer. This could mean that the neighbourhood s is a busy type and have high viability for business.

6. Conclusion

I was able to identify a neighbourhood with a similar feature by k-means clustering. With the client also considering a little densely populated neighbourhood, I considered the next-in-line city after downtown Toronto in terms of population and further did an exploratory analysis of its neighbourhoods. The project is intended to recommend an optimal location for my client hoping to open a restaurant in any of the neighbourhood s in either West or Downtown Toronto. I considered the top 20 neighbourhood s with the highest number of restaurants.

Also, I tried to explore further to determine the top 5 restaurants in the neighbourhood s. I realized a distinct neighbourhood location with an opportunity, because despite having more venues the most neighbourhood , there are fewer restaurants. Based on my result, I suggested that the client should consider a general-purpose restaurant and not the one made for demographics. I believe that will increase the restaurant's chance of exploiting consumer opportunities.