Capstone Project

The Battle of Neighborhoods :

The best place to open an Asian restaurant in Toronto

Introduction/Business Statement

Where is the best location to open an asian cuisine restaurant in Toronto ?

The stakeholder or the target audience for this would be

- Any business owner who would want to venture into food industry
- Food chain business owner who would want to expand their businesses
- Marketing company in food business industry
- Fund manager who would want to fund a startup for food business
- Food lovers themselves who just wish to have a good asian cuisines
- Tourist who are looking for good asian cuisine

Data Source

- Postal code and neighborhoods of Toronto
  - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- Geospatial information
  - https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.cs
- Demography breakdown of Toronto area based on the ethnic
  - https://en.wikipedia.org/wiki/Demographics_of_Toronto
- Foursquare API for location data information

Methodology

- We scraped the data from the source and created a dataframe called df_neigh from scraped Toronto postal code , Borough and Neighboorhood.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.653654	-79.506944
99	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
100	M7Y	East Toronto	Enclave of M4L	43.662744	-79.321558
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999

```
Toronto_data.shape
```

```
(103, 5)
```

Methodology

- The next step is to populate the dataframe with latitude and longitude information. Data is joined with the same postalcode.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.653654	-79.506944
99	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
100	M7Y	East Toronto	Enclave of M4L	43.662744	-79.321558
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999

Methodology

- We would now want to add demographic data to reflect the population density of Asian ehtnic in Toronto neighborhood.

	Borough	Race1	Race2	Race3	Race4	Race5	Race6
0	Toronto,East York	White: 65.3%	Chinese: 8.9%	South Asian: 6.7%	Black: 5.6%	NaN	NaN
1	North York	White: 47.4%	Chinese: 14.0%	South Asian: 8.5%	Black: 5.2%	Filipino: 8.0%	West Asian: 5.3%
2	Scarborough	White: 26.5%	Chinese: 19.0%	South Asian: 25.4%	Black: 10.8%	Filipino: 8.4%	NaN
3	Etobicoke,York	White: 48.9%	Black: 15.7%	South Asian: 11.9%	, Latin American: 5.6%	NaN	NaN

```
df_demographic['Borough']
```

```
0    Toronto,East York
1           North York
2          Scarborough
3       Etobicoke,York
Name: Borough, dtype: object
```

Methodology

- Master Toronto neighborhood data is then filtered with this high ethnic Asian populated area. Snippet of the scipts below

```
#will search for Asian restaurant nearby neighborhood
search_query = 'Asian'
radius = 1000  # set to 1000m
LIMIT = 100 # set to 100 venues
print(search_query + ' .... OK!')
for i in range(len(Toronto_data)):
    url_x = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&oauth_t
    results_x = requests.get(url_x).json()
    try:
        num_venues = len(results_x['response']['venues']) #get number of venues
    except:
        num_venues = 0
    print("<>",i,"<>",num_venues,"<>",Toronto_data.loc[i,'Latitude'],"<>",Toronto_data.loc[i,'Longitude
    Toronto_data.loc[i,'AsianRestaurant'] = num_venues
    print('There are {} popular spots around {} in {}.'.format(num_venues,Toronto_data.loc[i,'Neighborh
#dataframe2 = json_normalize(results_x['response']['venues'])
```

```
gh.
<> 56 <> 0 <> 43.7527583 <> -79.4000493
There are 0 popular spots around York Mills West in North York.
<> 57 <> 0 <> 43.696319 <> -79.5322424
There are 0 popular spots around Westmount in Etobicoke.
<> 58 <> 1 <> 43.7500715 <> -79.2958491
There are 1 popular spots around Wexford, Maryvale in Scarborough.
```

Methodology

- Now we get the get the list of Asian restaurant in Asian ethnic populated neighborhood

```
#will search for Asian restaurant nearby neighborhood
search_query = 'Asian'
radius = 1000  # set to 1000m
LIMIT = 100 # set to 100 venues
print(search_query + ' .... OK!')
for i in range(len(Toronto_data)):
    url_x = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&oauth_t
    results_x = requests.get(url_x).json()
    try:
        num_venues = len(results_x['response']['venues']) #get number of venues
    except:
        num_venues = 0
    print("<>",i,"<>",num_venues,"<>",Toronto_data.loc[i,'Latitude'],"<>",Toronto_data.loc[i,'Longitude
    Toronto_data.loc[i,'AsianRestaurant'] = num_venues
    print('There are {} popular spots around {} in {}.'.format(num_venues,Toronto_data.loc[i,'Neighborh
#dataframe2 = json_normalize(results_x['response']['venues'])
```

```
gh.
<> 56 <> 0 <> 43.7527583 <> -79.4000493
There are 0 popular spots around York Mills West in North York.
<> 57 <> 0 <> 43.696319 <> -79.5322424
There are 0 popular spots around Westmount in Etobicoke.
<> 58 <> 1 <> 43.7500715 <> -79.2958491
There are 1 popular spots around Wexford, Maryvale in Scarborough.
```

Methodology

- We would use Foursquare API for collecting location data from the web.
  - url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&oauth_token={}&v={}&query={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, neighborhood_latitude, neighborhood_longitude, ACCESS_TOKEN, VERSION, search_query, LIMIT)

Methodology

- Now we get the get the list of Asian restaurant in Asian ethnic populated neighborhood. The information then combined with the master Toronto data

```
#final_df = df.sort_values(by=['2'], ascending=False)
Toronto_data.sort_values('AsianRestaurant',ascending=False).head(20)
```

	PostalCode	Borough	Neighborhood	Latitude	Longitude	AsianRestaurant
22	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	12.0
77	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280	10.0
28	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568	10.0
37	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	9.0
8	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	9.0
41	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	8.0
65	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049	8.0
79	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	7.0
14	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	7.0
72	M5W	Downtown Toronto	Enclave of M5E	43.646435	-79.374846	6.0
63	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049	6.0
33	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752	4.0
62	M1S	Scarborough	Agincourt	43.794200	-79.262029	4.0

Methodology

- Geopy.Nominatim is being used to locate the latitude and logitude on Toronto

```
address = 'Toronto, Ontario'

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Toronto are {}, {}.'.format(latitude, longitude))
```

```
The geograpical coordinate of Toronto are 43.6534817, -79.3839347.
```

Methodology

- We use K-means clustering algorithm to cluster Toronto venues category .For this project, we would use K equal to 5 and train it with Toronto_group dataframe. This algo will choose the similarity(closeness) of the neighborhood based on the category and cluster them together.

	Neighborhood	Venue Category_Accessories Store	Venue Category_Adult Boutique	Venue Category_Airport	Venue Category_Airport Food Court	Venue Category_Airport Gate	Ve Category_Air Lou
0	Agincourt	0.000000	0.000000	0.000000	0.0000	0.0000	0.0
1	Alderwood, Long Branch	0.000000	0.000000	0.000000	0.0000	0.0000	0.0
2	Bathurst Manor, Wilson Heights, Downsview North	0.000000	0.000000	0.000000	0.0000	0.0000	0.0
3	Bayview Village	0.000000	0.000000	0.000000	0.0000	0.0000	0.0
4	Bedford Park, Lawrence Manor East	0.000000	0.000000	0.000000	0.0000	0.0000	0.0

Result & Findings

- 7 identified borough with high density of Asian ethnic . Scarborough being the most dense with more than 50% of ethnic Asian population. Table shows high density of Asian ethnic in Toronto

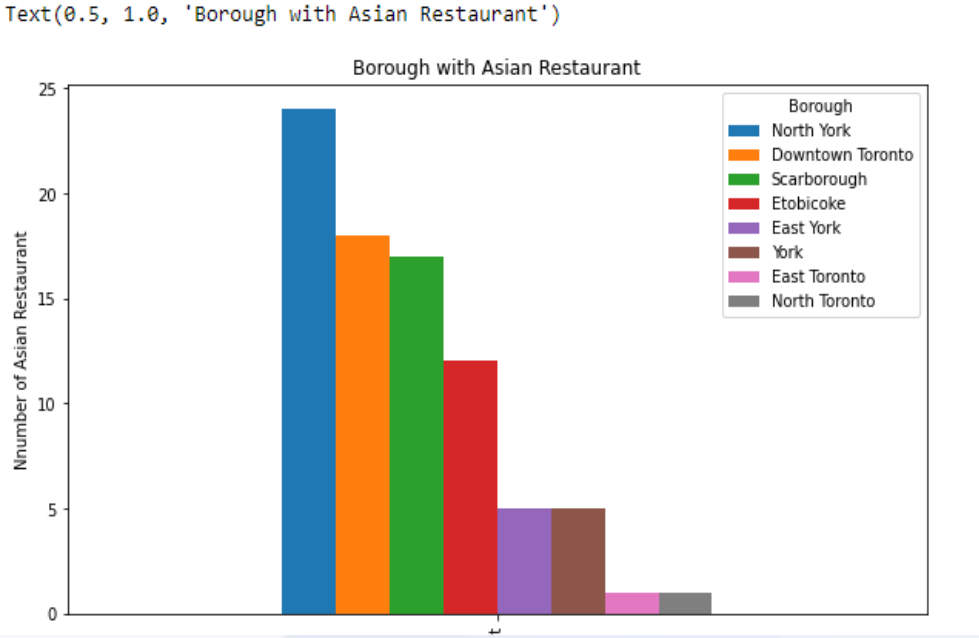
	Borough	Race1	Race2	Race3	Race4	Race5	Race6
0	Toronto,East York	White: 65.3%	Chinese: 8.9%	South Asian: 6.7%	Black: 5.6%	NaN	NaN
1	North York	White: 47.4%	Chinese: 14.0%	South Asian: 8.5%	Black: 5.2%	Filipino: 8.0%	West Asian: 5.3%
2	Scarborough	White: 26.5%	Chinese: 19.0%	South Asian: 25.4%	Black: 10.8%	Filipino: 8.4%	NaN
3	Etobicoke,York	White: 48.9%	Black: 15.7%	South Asian: 11.9%	, Latin American: 5.6%	NaN	NaN

```
df_demographic['Borough']
```

```
0    Toronto,East York
1           North York
2          Scarborough
3       Etobicoke,York
Name: Borough, dtype: object
```

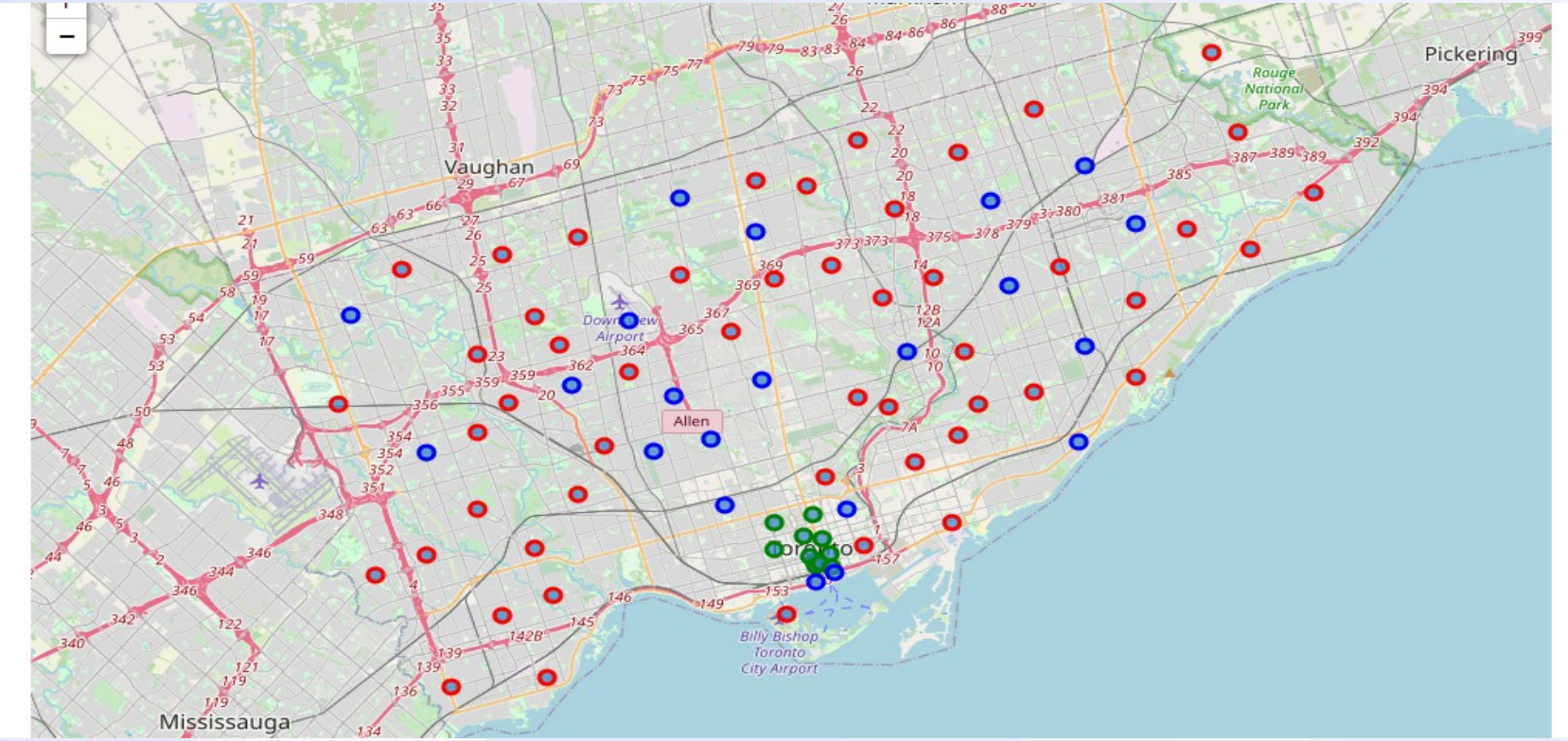
Result & Findings

- 83 total neighborhood with high density of Asian ethnic based on available demographic data. North York having the highest number of dense asian ethnic population with 24 neighborhood followed by burough of Downtown Toronto and Scarborough.



Result & Findings

- Visualisation of Asian in the neighborhood.*Map show the neighborhod with number of Asian restaurant. Green dot represent > 5, blue > 0 & < 6 while red = 0*



Result & Findings

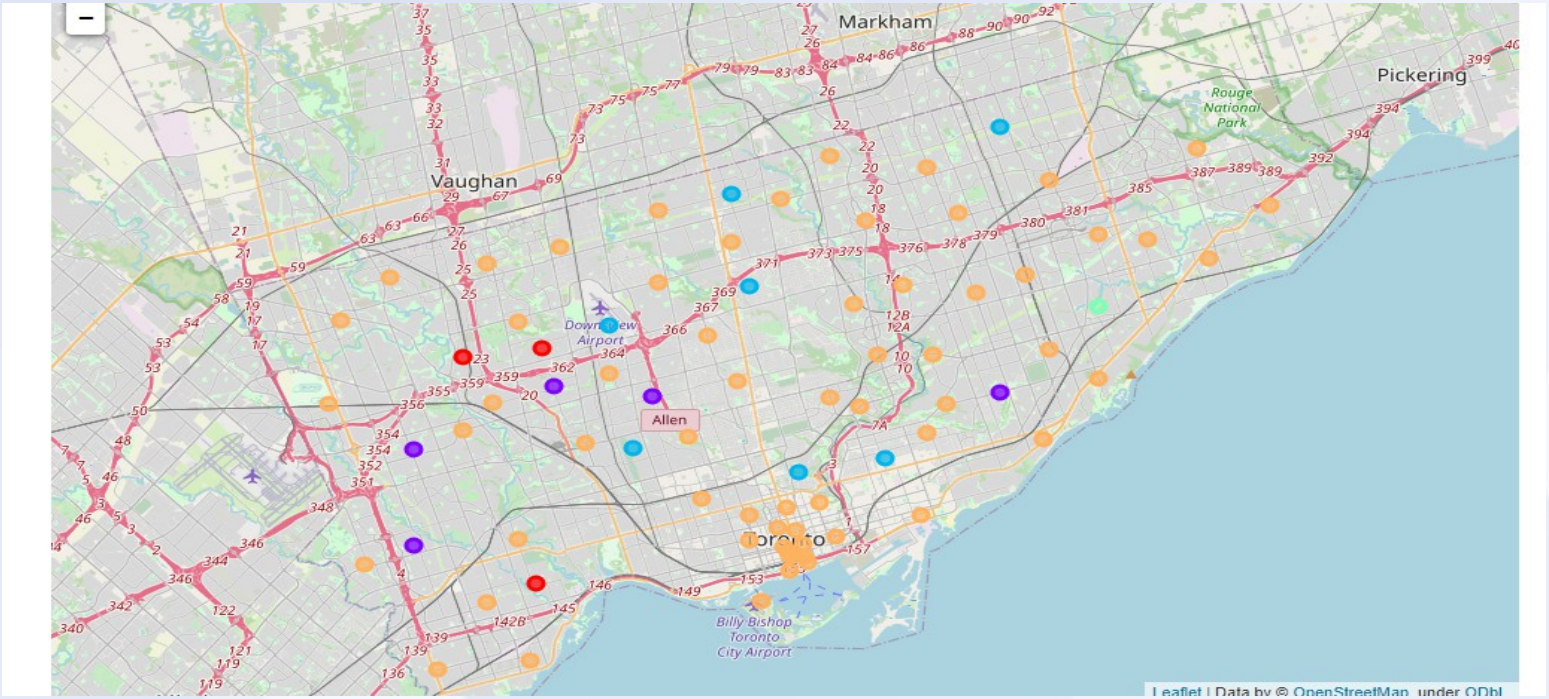
- Downtown Toronto borough being the most dense Asian restaurant with Central Bay, First Canadian Place, Undergound City, Richmond, Adelaide & King neighborhood having 10 or more restaurant.

```
#final_df = df.sort_values(by=['2'], ascending=False)
Toronto_data.sort_values('AsianRestaurant',ascending=False).head(20)
```

	PostalCode	Borough	Neighborhood	Latitude	Longitude	AsianRestaurant
22	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	12.0
77	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280	10.0
28	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568	10.0
37	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	9.0
8	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	9.0
41	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	8.0
65	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049	8.0
79	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	7.0
14	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	7.0
72	M5W	Downtown Toronto	Enclave of M5E	43.646435	-79.374846	6.0
63	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049	6.0
33	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752	4.0
62	M1S	Scarborough	Agincourt	43.794200	-79.262029	4.0

Result & Findings

- Clustering of Toronto neighborhood based on categorical venues as below. Cluster 4 being dominant with most number of neighborhood.



Result & Findings

- Most common venues for Scarborough, North York, East York and East Toronto.

```
]: Toronto_merged[Toronto_merged['Borough'] == "Scarborough"]['1st Most Common Venue'].unique()

]: array(['Venue Category_Fast Food Restaurant', 'Venue Category_Bar',
       'Venue Category_Donut Shop', 'Venue Category_Coffee Shop',
       'Venue Category_Fried Chicken Joint', 'Venue Category_Playground',
       'Venue Category_Hobby Shop', 'Venue Category_Bakery',
       'Venue Category_Motel', 'Venue Category_College Stadium',
       'Venue Category_Indian Restaurant', 'Venue Category_Auto Garage',
       'Venue Category_Skating Rink', 'Venue Category_Park'], dtype=object)
```

```
Toronto_merged[Toronto_merged['Borough'] == "North York"]['1st Most Common Venue'].unique()

array(['Venue Category_Bus Stop', 'Venue Category_Portuguese Restaurant',
       'Venue Category_Clothing Store',
       'Venue Category_Japanese Restaurant', 'Venue Category_Restaurant',
       'Venue Category_Fast Food Restaurant', 'Venue Category_Bank',
       'Venue Category_Massage Studio', 'Venue Category_Airport',
       'Venue Category_Grocery Store', 'Venue Category_Basketball Court',
       'Venue Category_Pizza Place', 'Venue Category_Park',
       'Venue Category_Food Truck', 'Venue Category_Sandwich Place',
       'Venue Category_Baseball Field', 'Venue Category_Ramen Restaurant',
       'Venue Category_Athletics & Sports', 'Venue Category_Pharmacy'],
      dtype=object)
```

Result & Findings

- East York and East Toronto having the least of common venues and zero Asian restaurant in their locality.

```
: Toronto_merged[Toronto_merged['Borough'] == "East York"]['1st Most Common Venue'].unique()

: array(['Venue Category_Pizza Place', 'Venue Category_Spa',
       'Venue Category_Coffee Shop', 'Venue Category_Indian Restaurant',
       'Venue Category_Park'], dtype=object)

: Toronto_merged[Toronto_merged['Borough'] == "East Toronto"]['1st Most Common Venue'].unique()

: array(['Venue Category_Light Rail Station'], dtype=object)
```

Discussion

- Downtown Toronto having the most of Asian restaurant based on Foursquare location data with 101 then followed by Scarborough with 10. The rest of the neighborhood much lesser with East Toronto and East York having none of Asian restaurant around.
- Scarborough and NorthYork and York having around 10 and less Asian restaurant are suitable location for setting up new restaurant
- As for the data that we have, East Toronto is the least prefered place to setup a restaurant due to single common venues.

Recommendation

- Based on the data, findings and discussion highligted above, the recommend of the best locations for new Asian restaurant are in Scarborough and North York due to many common venues and with none or less having an asian restaurant.